

Bad Neighborhoods of the Electromagnetic Spectrum:
A Method for Predicting the Deployment of Unlicensed Devices

Christian Sandvig (csandvig@uiuc.edu)
University of Illinois at Urbana-Champaign

August 2007

Submitted to the 35th Research Conference on Communication, Information, and Internet
Policy, Arlington, Virginia, USA.

Bad Neighborhoods of the Electromagnetic Spectrum: A Method for Predicting the Deployment of Unlicensed Devices

Abstract

An uncontroversial prediction about the future of electromagnetic spectrum use is that the developed world will soon be filled with far more license-exempt low powered devices. Today's widespread deployment of Wi-Fi and Bluetooth exemplify this trend. Although in many ways a consumer's Wi-Fi enabled portable device couldn't be more different from a licensed television broadcast tower serving a major city, the mental frameworks used in policymaking to reason about the use of spectrum remain in thrall to high-powered broadcast towers. For example, some proposals for differential power limits divide the nation into two kinds of places: rural and urban. Often policy action about spectrum is never based in any empirical knowledge about use or congestion. This paper demonstrates the problems with these approaches by reporting on a longitudinal empirical study of low-powered unlicensed spectrum use (specifically, Wi-Fi deployment) in Illinois and California over four years.

From 2004 to 2007, researchers visited selected census block groups ($N = 60$) in urban, suburban, and rural areas in two states and used modified packet sniffers to identify Wi-Fi access points, collecting several million observations. By combining the packet sniffer with a GPS receiver, we estimated the location of each access point we discovered. Then, we employed the extensive diffusion of innovations literature to capture demographic features salient to Wi-Fi diffusion (Rogers, 1995), and gathered these data from the US Census. Next, we developed a demographic regression model that predicts Wi-Fi deployment in our study areas ($R^2 = .84; p < .001$). Our method is significantly better than existing alternatives (R^2 change = .48; $p < .001$), and is also reliable (tested via split-half / cross validation) at modeling new data ($.77 \leq R^2 \leq .92$).

Finally, we constructed a subset of the relevant Census tables for the entire US and applied our regression model to construct a national unlicensed spectrum prediction engine. Given any address in the US, our service -- named the "RED Project" for "Rendering Electromagnetic Distributions" -- predicts Wi-Fi deployment and (using a Macromedia Flash front end) dynamically renders the regression equation predictions on a street map or satellite photo at a variety of zoom levels.

Aside from the possible utility of this prediction service (for example, at predicting congestion), surprisingly, our findings also overturn a number of habits of thought about spectrum. For a variety of reasons, some densely populated city blocks are devoid of Wi-Fi and some rural areas are approaching congestion failure. As the diffusion of innovations literature predicts, adoption cannot be predicted simply by income or by population density. Some "bad" neighborhoods in US cities are "good" neighborhoods of the spectrum--that is, they are devoid of congestion--and vice-versa. Some neighborhoods that communication companies think of as a very desirable urban market are almost devoid of signals. We conclude by trying to come to terms with the variegated new spectrum topography of low-powered consumer devices.

A radical shift in the public policy approach for managing the electromagnetic spectrum of the US is now underway. Advances in wireless technology and the present market success of “Wi-Fi” (IEEE 802.11a/b/g) wireless data networking have combined to convince many that unlicensed spectrum—spectrum that does not require permission to transmit—is now a viable way to manage a much larger amount of our wireless communication. However, by definition use of unlicensed spectrum is unplanned and decentralized. Public policy statements and the scientific literature to date have characterized unlicensed as “messy” (Abernathy, 2002) and unpredictable, even when they are supporting it (see Noam, 1995; Satapathy & Peha, 1997).

This enthusiasm for unlicensed spectrum has put administrative law communication regulators in an awkward position. While the idea of license-exempt spectrum is consonant with the neoliberal enthusiasm for deregulation and it ought to find political support among consumer products manufacturers and computer manufacturers, there remain many arguments against it. Entrenched industries already control the great swaths of the spectrum available for non-governmental use. Even if new allocations for unlicensed can be found that don't directly challenge their own licenses, they rightly feel that any increase in unlicensed services undermines the investments in spectrum that they have already made. Spectrum allocation by licensing provides vast sums for the national treasury via auctions. Less discussed is the very real opposition to unlicensed on the grounds of regulatory culture and process.

That is to say, unlicensing spectrum presently creates few benefits that can be measured by regulators, but it imposes great risks for regulators. Regulatory officials take licensing and spectrum allocation as their reason for existing at all. Although unlicensed devices must accept interference and have no regulatory protection from it, nonetheless an unlicensed band that has become congested threatens the competence of the national

regulator, who will be publicly blamed regardless of any statements about interference posted on certification notices or in administrative rules. Unlicensed spectrum is a challenge for them because when spectrum is unlicensed it is very difficult to predict how or how much it will be used. While this makes it a boon to innovation, in regulatory culture it may feel like the creation of a license-exempt spectrum park is irrevocably throwing that part of the spectrum away.

This paper describes a research project that hypothesizes the opposite of some of these conceptions about license-exempt spectrum. It hypothesizes **that unlicensed spectrum is very predictable**. Specifically, extensive social scientific research on the diffusion of innovations (Rogers, 1995) and on the development of communication infrastructure (see below) already provides a clear picture of how and which people use communication technology. This project uses the current example of Wi-Fi as a way to analyze the development of unlicensed, decentralized communication systems in the field. This paper presents a fairly simple method that allows social scientists to map Wi-Fi by sampling the electromagnetic spectrum directly (instead of using, for example, a survey questionnaire given to users). By mapping the diffusion of Wi-Fi geographically, the project visualizes the spread of this technology across the US. By linking this spatial data to the publicly available US Census it is possible to create a predictive model that can be used to understand diffusion generally and to pragmatically demonstrate that unlicensed transmissions are predictable. This is a new way of analyzing and thinking about both unlicensed spectrum and differentials in access to communication technology.

Why Wi-Fi?

The purpose of considering unlicensed Wi-Fi communication networks is three-fold. First, “network” is not a new word or a complex one—it is about 400 years old, it appears in the bible, and it refers to any “work” (something that is made), that resembles a “net,” as in the mesh used to snare fish. Thought of this way, calling a telephone system or a computer system a network is poetic, but misleading. If we held aloft the telephone wires of the world, held true to the shapes we find in the earth, the topographical pattern we find would not catch fish. This is because networks are not used to link all places—they are not fine sieves that drape the land but selective connections that link “valuable places in a non-contiguous pattern” (Castells, 1998: 144). This selectivity is the central problem for several strains of research on communication technology; it animates studies of *diffusion* (e.g., Rogers, 1995), *inequality* (e.g., Schiller, 1996), *universal access* (e.g., Anderson et al., 1995) *universal service* (e.g., Mueller, 1997), and the “*digital divide*” (e.g., Norris, 2002). The spatial approach to unlicensed is a way of considering this central problem of access. It allows the combination of geographic analysis, engineering, and previous social scientific research. A model of unlicensed Wi-Fi use could, for instance, forecast “holes” where populations are likely to be underserved by communication technology across the US.

Second, the empirical measurement of actual Wi-Fi diffusion generates new datasets and methods in an area where little is known. While network mapping is used in network engineering practice (Byers & Kormann, 2003), so far it is rarely used in social science (e.g., Matei, 2004). While the diffusion of innovations literature and the infrastructure development literature are both well-developed, neither have often had a spatial component. The increasing sophistication of geographic information systems (GIS) and the increasing ability to perform desktop analysis of large data sets (such as the entire US Census) open new approaches to social science that this project will pioneer.

Third, the project includes an Internet mapserver as a way to disseminate these predictions about unlicensed activity in the electromagnetic spectrum. Similar to a *Mapquest* or a *Google Maps*, a “Spectrum Mapserver” of a small portion of the unlicensed spectrum allows users to be able to enter an address and evaluate the model’s predictions for an area of their choice. This facilitates public understanding of complicated scientific and technical phenomena and validates or disproves the notion of this form of prediction in a compelling way. In addition, it is a new approach to the logic behind a “spectrum server” (e.g., see Ileri, Samardzija, & Mandayam, 2005).

The History and Consequences of License-Exempt Spectrum

Twenty years ago, the FCC embarked on an experiment where three small slices of “junk” (poorly penetrating) frequencies were allowed to be used without permission. Consumer electronics manufacturers argued that very low power could be combined with new spread spectrum technology to give protection enough from interference, and they hoped to sell short range devices like cordless phone handsets. Over the next two decades, the use of these bands increased to support a diverse array of devices, including today’s Wi-Fi. In 2002, unlicensed systems were so pervasive and commercially successful that policymakers called broadly for spectrum allocation reform and the expansion of unlicensed bands, with Wi-Fi the exemplary successful application (Spectrum Policy Task Force, 2002). That is, “Wi-Fi is an existence proof for the validity of the open spectrum argument” (Werbach, 2002).

The idea of unlicensed or “open” spectrum runs counter to eighty years of government licensing, but the thinking behind this change has been building for over a decade (for an introduction and overview, see Peha, 1998, 2005). For at least the last twenty years, advances in digital communication have led engineers to predict an almost complete freedom from scarcity in the spectrum (Pool, 1983). But whether spectrum is scarce or not,

unlicensed or open spectrum has been found to be a viable approach (Lehr, 2004).

Essentially, unlicensed operation is not unregulated: it regulates transmission power level by certifying equipment instead of limiting access or the location of users (47 CFR Part 15).

Earlier efforts to use market mechanisms to control access to the spectrum were dominated by auctions (Milgrom, 2004), but unlicensed spectrum is also a market-driven allocation mechanism – it simply shifts the relevant market to equipment manufacturers (Benkler, 2002).

As any user of an older, static-filled cordless phone will realize, large numbers of low-powered devices may produce localized interference, but users are required to tolerate it. Power limits are intended to keep this interference from becoming widespread. Due to the low power of transmissions, the case of Wi-Fi may ostensibly seem to be about only short range communication. However, researchers have also seen it as a bellwether for more extensive networks – a decentralized, short range beginning to what could become a network with longer links (Sawhney, 2003) possibly through advanced routing and cooperation rather than improvements in range for individual transmitters (Benkler, 1998, 2002). The viability of these proposals rests on matters of routing and social organization beyond the scope of this proposal, but suffice to say many researchers have seen Wi-Fi as important.

The same potential that made unlicensed low-power wireless technologies like Wi-Fi attractive because it is ideologically consonant with the political goal of deregulation also leads to a perception that it is risky and unpredictable. In public comments to the FCC on the expansion of unlicensed, some filers claimed that large unlicensed systems are impossible. Without a license's legal guarantee of freedom from interference, no one could predict what would happen in unlicensed spectrum and no company would invest in a system that was so "high-risk."

Although the use and the efficiency of the electromagnetic spectrum might seem to be a problem for engineers, in the past arguments about the spectrum have been filled with

assumptions about human behavior. Spectrum allocation regimes have privileged some human uses over others, and even some forms of social organization over others (Streeter, 1996). In the center of today's debates about spectrum are predictions about humans, not radios. For instance, whether an expanded unlicensed regime will work or not depends crucially on whether unlicensed slices of the spectrum are "full" or "empty," where "full" might indicate congestion failure. This is controlled by device efficiency, but also by what imagined users will do with the devices and how tolerant these users (and uses) will be to quality degradation due to interference.

There is an extensive, excellent economic literature on spectrum allocation, and some attention has been paid to unlicensed. However, the general approach has been to limit interference and increase efficiency by modifying the behavior of devices and protocols—such as by introducing rules of spectrum sharing etiquette (e.g., see Peha, 2000). This is valuable, but this project approaches the topic from a very different direction by testing whether or not the user's behavior (not the device or protocol) is predictable.

Technology Diffusion and Space

One of the beginnings of research on the diffusion of innovations famously began with a diagram of a neighborhood showing the spread of new agricultural ideas between farmers (Rogers, 1995), yet the spatial dimensions of diffusion are rarely assessed quantitatively. Instead, the S-curve originally drawn by Gabriel Tarde (1903) has been used to plot the rate of individual adoption of a new idea or technology over time. Individual characteristics have predicted the adoption of new technology (Rogers, 1995). The use of demographics to predict the use of specific communication technologies has a long history (see Schramm & Roberts, 1971). In this project, instead of measuring the societal rate of adoption for Wi-Fi or differentiating groups of adopters from each other, we use a geographic unit of analysis to model and visualize the progress of Wi-Fi across space, and

therefore the use of the electromagnetic spectrum across space. There is some precedent for this effort.

While scholars have typically viewed diffusion as an individual-level phenomenon and have used approaches like *network communication* models to show influence between potential adopters (Valente, 1995), there are a few exceptions that have used *spatial* models (see Allaway et al., 1994; Zhang et al., 2002). Internet diffusion has already been usefully analyzed geographically via the study of domain name registrations (e.g., Grubestic, 2002) and network topography has been analyzed spatially for the purposes of predicting network survivability (e.g., Grubestic et al., 2003). Indeed, spatial diffusion studies already exist but they often proceed under the heading of “penetration” (Mueller & Schement, 1996) “geographic distribution” (Forman et al., 2003; Greenstein et al., 2003) or “economic geography” (Greenstein, forthcoming). Some of these situate themselves in diffusion theory (see Greenstein & Prince, 2004). The increasing capability of GIS systems and the availability of new spatial data are creating opportunities for spatial approaches to studies of technology diffusion and even communication itself (e.g., Matei et al., 2001). These studies typically employ the spatial dimensions of their data to investigate hypotheses about space, e.g., agglomeration and clustering, or the difference between regions or areas.

The rise of the Internet has produced a concomitant interest in maps of communication infrastructure and virtual spaces (Dodge & Kitchin, 2000, 2001). Some of the diffusion studies cited above use visualizations and some do not. Many scholars see the visualization of the otherwise invisible geography of infrastructure as essential. These visualizations employ new cartographic techniques to visually “reveal the uneven and unequal distribution of infrastructure” (Dodge & Kitchin, 2003). This builds on earlier research that employed striking maps to reveal previously unknown patterns in, for example, domain name registration (Moss & Townsend, 1997), global Internet inequality

(Holderness, 1998), Internet content production (Zook 2000, 2001), and the topography of the wired Internet backbone (Moss & Townsend, 2000).

A number of scholars have studied or promoted efforts to redress spatial inequality in infrastructure. After identifying infrastructure, service, or application “holes,” they have advocated the creation of organizations like wireline-based community networks (Schuler, 1994, 1996) that provide Internet service in needy areas (Rogers, Collins-Jarvis & Schmitz, 1994; Schmitz, Rogers, Phillips & Pascal, 1995; Silver, 2004). Others have similarly investigated the role of neighborhood projects (Hampton & Wellman, 1999, 2000, 2003), and public access centers (Straubhaar, LaPastina, Lentz, Main, & Taylor, 2000). For a more detailed review, see Harrison & Stephen (1999) and O’Neil (2002). Much of this tradition of scholarship relies on quantitative studies of inequality and forecasts of diffusion like this project.

The Case of Wi-Fi: The Wi-Fi Cloud as a Morphological Puzzle

When choosing one technology to reason about “unlicensed technology,” we must first justify our choice of case (Yin, 1994). The Wi-Fi “movement” (after Kling & Iacono, 1995) that arose to take advantage of the IEEE standard for wireless communication often defy existing conceptions of a technological system. While Wi-Fi was originally envisioned as a home, classroom, and small office technology to replace a few feet of Ethernet cable, it has been deployed in many other unexpected ways by unexpected actors that include individuals, corporate IT departments, cafes, telecommunications companies, governments, and universities.

At a technology conference in late 1999, Steve Jobs introduced the Apple AirPort, the first Wi-Fi access point (AP). Jobs promised users, “it’s a liberating experience to surf the Internet...while freely moving about your home or classroom.” Users, however, have had

other ideas about where they wanted to freely move. In the last five years the technology behind the AirPort and similar devices—now called Wi-Fi (for “wireless fidelity”)—has become wildly popular. Wi-Fi allows computers to communicate without wires at high speeds over short distances (see Institute of Electrical and Electronics Engineers, 1999).

Initially conceived as a technology to solve small networking problems in the home, it and its successor (WiMax) have instead become a contender for the next generation of Internet infrastructure (Lehr & McKnight, 2003). Wi-Fi has even been deployed as a city and county-wide communication system (Gillett, Lehr, & Osorio, 2004; Scott et al., 2005). In its first five years, Wi-Fi use has grown faster than previous fast-growing technologies such as Ethernet. Last year between 44 and 99 million users (Bar & Galperin, 2005; Legon, 2005) around the world spent \$2.4 billion (Hales, 2005) to use some of the over 2000 different kinds of devices that use Wi-Fi (Wi-Fi Alliance, 2005).

In a beautiful metaphor, Alessandro Ovi has called Wi-Fi networks “water lilies” (Negroponte, 2002). In this metaphor the lily pad is the coverage of the network: the lilies grow independently, but they can eventually cover the surface of a pond (the Wi-Fi cloud). The lily pads overlap and some lilies can stifle others (interference). The stems of most lilies lead to the Internet through a fast corporate Ethernet or a home-office/small-office broadband connection. These Wi-Fi access points (APs) are widely available, easy to configure and inexpensive.

The entities that now provide the Wi-Fi cloud each produce connectivity for themselves by deploying Wi-Fi. Yet, third parties can also use this connectivity with the intent of the network’s owner, through the owner’s ignorance, by design, or by security failure. That means that the presence of a well-developed Wi-Fi cloud in an affluent apartment building often promotes both formal and illicit sharing of Internet connections. A number of nonprofit and activist organizations have formed to assist in this sharing (Sandvig, 2004). A debate about the security standards and defaults for Wi-Fi is ongoing in

the trade press (Shah & Sandvig, under review). The notion that an unsecured Wi-Fi cloud is a community property has also been embraced by some local government agencies for reasons of economic development.

While we often think of modern communication systems as being provided by large telecommunications companies, initial investments in Wi-Fi “hotspots” that allow Internet use in public places for a fee were disappointing, causing some to predict a “Wi-Fi bubble.” However, adoption continues to grow, and even some hotspot providers such as T-Mobile (the largest) now report that they have a sustainable base of paying users (T-Mobile reports 450,000 in a three month period, Associated Press, 2005). While the “hotspot” is emblematic of Wi-Fi, these numbers are so small in comparison to the installed base of Wi-Fi that the story of Wi-Fi to date does not seem to be largely about centralized rollout of large systems by dominant carriers. The FCC estimates that 3,000 wireless Internet Service Providers are also currently operating, and most of these are small entrepreneurs.

The Wi-Fi cloud is then a mixture of “lone” APs purchased by individuals and networks constructed by firms or cooperating groups (like tenant associations) that may be informal, formal, corporate, governmental, or some hybrid. It is not particularly surprising that Wi-Fi has so often been labeled disorganized or chaotic by scholars (this author included, Sandvig, 2004).

Research Design

This project represents original collection of wireless signal data, combined with secondary data from the US Census for analysis, then disseminated via a Mapserver. The unit of analysis is the Census block group, the smallest geographic unit for which the Census Bureau tabulates sample data. The Census divides the US into block groups where the minimum population is 600 and is maximum is 3,000 with an optimum population of 1,500. Block group boundaries must follow visible and identifiable landscape features such as roads,

railroad tracks, rivers, or coastlines, therefore researchers can readily identify them in the landscape with a key. The block group is a smaller unit than the more familiar *tract*, which contains from 1 to 9 block groups. This unit of analysis allows our primary wireless data to be combined with secondary Census data. Block groups are defined by population and can vary dramatically in size. In pretests, it took researchers a few hours to carefully traverse all public streets in one block group.

Representative sampling is a problematic notion in geographic applications generally, but even standard approaches to generalization (such as randomized X,Y coordinate selection weighted by population density) are impossible for this kind of research. First, the typical approach attempts to weight by population density as a claim to represent the number of humans. However, we aim to investigate use of the spectrum, and empty spaces with few humans (and thus, few transmissions) may be just as interesting for the purposes of our research. In addition, for a proof-of-concept academic project the collection of Wi-Fi data is too time-consuming and a representative sample of block groups is prohibitively expensive. As a comparison, a company promoting a “Wi-Fi Positioning System,” Skyhook Wireless, privately surveyed major roads in 25 US cities. It claims to have taken three years to develop a database of 1 million APs. This was undoubtedly at great expense (the financial backing of the company is not public), and yet compared to our very small pretests we find this number of APs to be extremely low.

Instead, we employ diversity sampling to generate variance in our proposed independent variables of interest. Because the independent variables are largely demographics from the Census, we can select block groups for wireless (dependent variable) measurement after first considering the composition of block groups. Preliminarily, we considered income, education, race, student status, occupation, age, dominant construction materials (e.g., wood vs. brick buildings), structure density, population density (esp. urban/suburban/small town/rural), neighborhood type (commercial/residential/mixed).

These are all independent variables that either appear in the Census or can be logged by researchers who visit the site.

Unfortunately, the sample size is largely dictated by the time-consuming nature of the data collection, and is not driven by the requirements of statistical power. Nonetheless we have obtained statistically significant results from our first attempt at a longitudinal analysis, presented below.

Instrumentation and Measurement for Wireless Data Collection

Researchers drove every public street in a selected study area while carrying specialized equipment. This is conceptually similar to what happens when any laptop user with a wireless card opens a laptop to see if wireless networks are present, in that we examine public information (the management frame). However, we use specialized passive detection equipment that is superior to the average laptop.

Researchers carried refurbished Sharp *Zaurus* personal digital assistants with expanded storage (via SD drive) where the standard operating system has been replaced with a Linux variant (*OpenZaurus*) and the wireless network sniffer *Kismet* has been installed. *Kismet* is a standard passive packet sniffer with additional features that allow it to distinguish and report on Wi-Fi networks. Using an expansion sleeve, the *Zaurus* can be modified to accept a PCMCIA network card and connected to a GPS receiver. The network card can also be connected to an external antenna (in this case, a 7.4dBi omnidirectional wand).

While the researcher moves through the research site at street level with this equipment, the GPS device records a stream of latitude and longitude coordinates and a timestamp. At the same time, the network card feeds data to *Kismet* which identifies summary information about each network by reading the public management frame and logs summary information about each network. The variables of interest are: signal-to-noise

ratio, frequency, reported identifier, security protocol in use (e.g., WEP), Wi-Fi product identification (model and manufacturer). The *content* of the wireless communication is not recorded.

After a survey, the log files can be uploaded to a network server where the latitude and longitude recorded by the GPS device allow researchers to merge Wi-Fi data with the US Census in *ArcGIS*. That is, researchers can both draw a map of a wireless cloud and perform analyses that link the visualized wireless to any variable from the Census.

Reliability and Validity Testing

As these methods are new (or are at least new to this purpose of prediction), it is important to pause and ensure that these methods actually measure what they claim to. First, it is critical to note that this is a study of reception. While one might use the location of detected signals to infer that the transmitter is nearby, this would be a conceptual mistake. Due to the nature of urban geography and the ease with which Wi-Fi signals can be detected if they have line-of-sight, it is entirely possible to encounter situations like those depicted in figure 1.

Even within the framework of reception, second we should ask do these methods reliably locate reception? Clearly some sources of error are well understood, such as the reliability of GPS. However, high-frequency wireless signals are very sensitive to a variety of constantly changing conditions (weather, humidity, interference, presence of large metal objects like cars), it may be that one pass through a sample area would produce an entirely different pattern of reception than a second pass minutes later, even though the enduring pattern of reception had not changed. This change could be explained by (for example) the constant reconfiguration of traffic on the streets changing the reflections of the signals. In addition, it may be that a different driver or a different route through the sample area would produce entirely different measurements. This is both a reliability and validity problem.

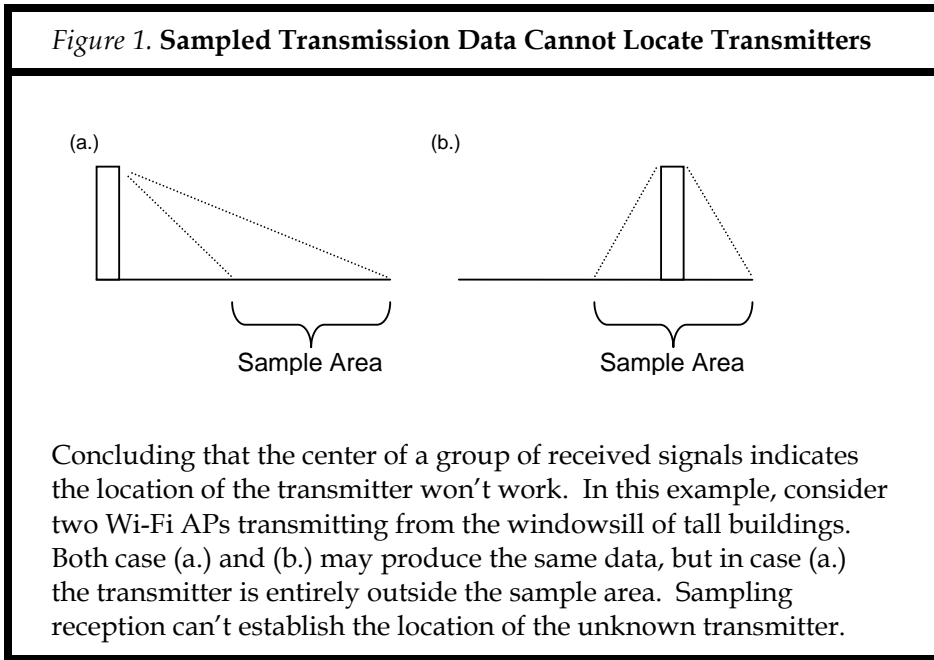
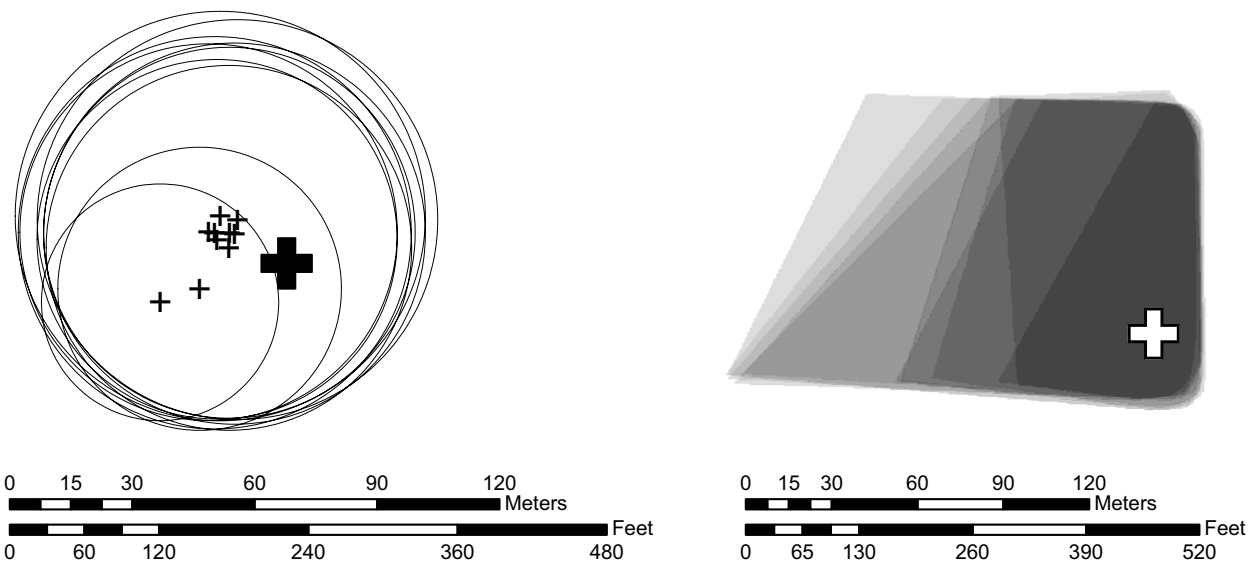


Figure 2. Reliability and Validity Testing: The Dispersion of Observations and Centers



We explored these problems by repeatedly measuring our own access points whose location was known. We used different sets of (the same) equipment, different researchers, different routes, and different speeds. Overall, as figure 2 indicates, these methods were reasonably robust. Along the each side of figure 2, we plotted ten passes measuring one known network using the methods elaborated by Byers and Kormann (2005). The large black and white crosses indicate the actual location of the AP. Along the left, we used the centroid visualization technique common in mapping programs like *Kismet*. We found that different drivers and sets of equipment did not produce different results.

However, traffic patterns produced persistent sources of bias. (For example, a traffic light or stop sign will slow traffic and result in more measurements of a nearby network.) Along the left side of figure 2, the center of the observations has been dragged toward a nearby traffic signal. While cleaning the data by eliminating overlapping observations can help reduce this source of error, this is still not a control for the speed of the researcher. Instead, we eliminated this source of error by interpolating the presence or absence of a network to a larger geographic unit (the block group).

The right side of figure 2 presents semitransparent convex hulls drawn around each pass of observations. The darkest grey indicates the most agreement about the area of reception. All passes agreed about a substantial portion of the reception envelope. Note that the grid pattern of streets strongly influences the shape of the convex hull representing reception—it is semi-rectangular. Note also that the area of reception varies substantially in size on some passes. We corrected for this reliability problem by interpolating to a fairly large geographic unit, thus minimizing the errors caused by variations in reception envelope near borders.

Analysis: Characterizing Wireless Diffusion

Multiple regression can test for significant relationships between characteristics of census block groups and the density of wireless networks that are found within them. Our first longitudinal analysis is presented here. We should stress that we do not expect startling results from these regressions. While each technology has particularities, many findings from the diffusion literature are remarkably consistent across technologies. The point of this project as elaborated above is not to prove that diffusion of innovations theory works, again. Instead, the advance is that we are applying this well-developed body of research to a new problem (spectrum allocation) and using it predictively. We use regression so that the resulting regression equation can be used as the basis for predictions about Wi-Fi use in areas we have not studied: a system for the prediction of unlicensed spectrum use.

The Spectrum Mapserver

As a way to advance our argument that the spectrum can be predicted, we use a public Web-based mapserver application to display our predictions about the electromagnetic spectrum across the US. While clearly there will be significant uncertainty in our predictions as we are extrapolating from a small set of block groups to the entire country, we believe that if these predictions are presented in an engaging manner they will spark intellectual and public policy debates. We employ a *PostGres* with *PostGIS* database backend to generate on-the-fly geographical querying capacity from a subset of the 2000 US Census. This is linked to a custom-written *Macromedia Flash* interface that visualizes the results on top of output from *Yahoo! Maps*. The end result is a public Web site where users can type in an address and be presented with our predictions for Wi-Fi use in that neighborhood. In effect, our mapserver allows the public to “fly around” the spatial predictions from our regression equation.

This effort is also a way to enlist the public and scientific community to help check our results, and to present them in an engaging manner. We allow users of the mapserver to rate the accuracy of predictions based on their own experience, engage in discussion about them, and to visually and statistically compare our predictions to actual measurements (not predictions) produced by our research group. It is also possible that entrepreneurs might use a Web service like this one to decide whether or not investing money in a communication system that uses unlicensed frequencies is a good idea, and it is possible that regulators might use such a service to understand the dynamics (and possibilities for congestion) presented by unlicensed spectrum.

An Example: Considering Wireless in Two Chicago Neighborhoods

As an introductory example, consider Wi-Fi use in two areas of central Chicago, West Humboldt Park and Lake View. Figure 3 shows a visualization of Wi-Fi use in a few blocks of both neighborhoods (to protect the privacy of residents, street names are omitted). Figure 3 gives both network density (the number of Wi-Fi networks per unit of space) and population density for both neighborhoods, and shows wireless networks at two time points. The methods above allow the drawing of polygons that approximate the coverage of each unique network. Each polygon has been made semitransparent so that overlapping networks will produce darker shading.

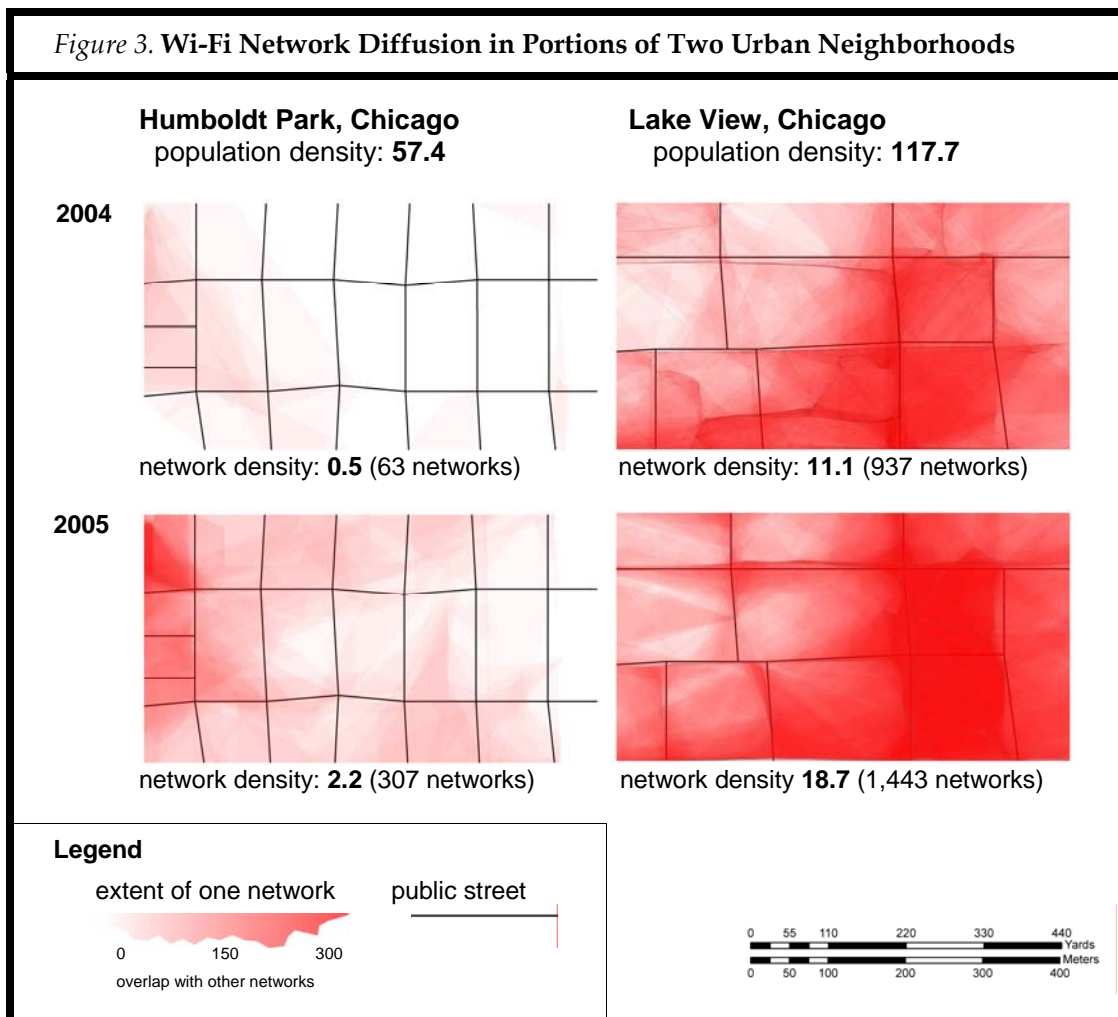
Population density is included here because it is often used as a shorthand to reason about spectrum use. For example, in what is possibly a hangover from broadcasting, in casual talk engineers and spectrum planners are likely to refer to less populated areas as places where the spectrum is “empty.” As one example that also illustrates university rivalries, Sahai (2005) of UC Berkeley refers to Champaign-Urbana as “middle of nowhere” in a graphic about spectrum scarcity (p. 6).

The wireless visualization itself yields immediate results. First, Lake View appears likely to have areas of extreme congestion. One street corner on the lower (2005) map of wireless in Lake View contains 281 overlapping Wi-Fi networks. This is already beyond some predictions of the carrying capacity of unlicensed (for example, Sahai notes that 6 overlapping networks makes spectrum “feel” scarce; Sahai, 2005: 2). Nonetheless, Wi-Fi appears to work in Lake View, likely because Internet traffic is bursty and the backhaul is likely to be relatively narrow compared to the capabilities of Wi-Fi (e.g., a cable modem).

Second, the rate of diffusion appears to be rapid, even though only two time points are shown. Third, when considering the wireless maps with the population density figures the visualization suggests the lack of explanatory power to be gained from the shorthand of population density alone. Humboldt Park appears partly “empty” and yet both of these neighborhoods are urban. Additionally, while Figure 3 shows that Lake View is about twice as densely populated as Humboldt Park, at one time point it has almost ten times the density of wireless networks.

Characteristics of Residents. Of course, the answers to these puzzles reside in consideration of the people in these neighborhoods. Both neighborhoods are alike in that they are mixed-use areas with commercial avenues topped and backed by multi-story housing and apartment buildings, with the occasional single-family house. Although only a portion is shown, both census tracts considered here are home to 8,000-9,000 people (although Lake View is more dense, with high-rise apartments), and both neighborhoods overwhelmingly feature brick or stone construction. Both neighborhoods have few students (< 9%). However, they diverge in most other demographics.

Along the left of the figure, West Humboldt Park is 97% black with high unemployment (13%) and has a per capita income of \$9,060 and a median age of 25.1. 56% of residents have a high school diploma.



Note. densities in millions per decimal degree.

In contrast, along the right of the figure, Lake View is 86% white and is in the Chicago Community Area (CA) where those employed in Information Technology and Management are most likely to live. It has low unemployment (2%) with a per capita income of \$54,280. The median age is ten years higher than Humboldt Park, at 35.1. Almost all residents (95%) have a high school diploma.

Predicting Unlicensed Use via Multivariate Regression

The diffusion literature emphatically stresses the requirement for a multivariate approach in this situation. While at first Figure 3 might suggest that Wi-Fi diffusion is a

simple matter of population, the previous section used demographics to show that this is not the case. As a second attempt, the reader might conclude that Wi-Fi diffusion is a matter of population density and poverty, this is still not the case. In the study of communication technology there are many examples of technologies that are popular among the poor (the best current example being the cellular phone) and there are populations with low incomes (demographers might say they are living in “situational poverty”) that form some of the groups that are most likely to adopt new technologies – such as college students.

Table 1. The Diffusion of Unlicensed Wireless in 30 Block Groups, 2004-2005

Ordinary Least Squares Regression on the Density of Wireless Systems

	<u>Independent Variable</u>	<u>$\beta_{std} (t)$</u>
Model 1 (the traditional view)	population density	.62 (5.44)***
	R^2_{adj} : .37; F(1, 48): 29.5***	
Model 2 (proposed approach)	population density	.66 (9.71)***
	socioeconomic status (income, education)	.31 (2.28)*
	median age	-.32 (-4.51)***
	race (% majority [white])	.22 (1.73)†
	months since introduction of technology	.29 (4.09)***
R^2_{adj} : .84; F(4, 44): 54.1***		

$N = 55$ (30 block groups, some sampled twice, 1 year apart)

R^2 change between models: .48***

† = $p < .1$; * = $p < .05$; ** = $p < .01$; *** $p < .001$

We conducted a regression on 30 block groups, some of which were measured more than once ($N = 55$ total block group measurements), presented as Table 1. These data were gathered in Illinois (both in Chicago and small towns) and California from 2004 to 2007. We gathered block group data from the Census on population density, socioeconomic status (standardized per capita income + % with high school diploma), median age, all standard variables from diffusion studies. We tested two models, the first being a straw man but still provocative model employing only population density – arguably a common shorthand way

of reasoning about the spectrum (Model 1). In the second model, we added the variables predicted by the diffusion literature.

The results are surprisingly strong: a change in R^2 of .48, with all factors but race significant at $p < .05$ and the direction consistent with diffusion theory. Results from the regression are given in Table 1.

Testing the Model's Ability to Predict Wi-Fi

While table 1 presents a regression result that is interesting in itself (that is, model 2), the larger value here is to develop a regression model that will reliably predict wireless density in areas that were not measured. We tested the value of the factors in model 2 via split-half / cross validation. The dataset was randomly divided in half, then regression models with the same factors were run, and the fit of each of the resulting equations was tested on the other half of the data. We found an adjusted R^2 of .77 and .92.

This paper presents only our first attempt at prediction. We intend to gather more data and continue to refine our model. However, even this first attempt does a surprisingly good job at predicting Wi-Fi. Aside from the discussion of cross validation above, it is possible to use the online mapserver to “fly through” the model's predictions for areas that the user knows in the US. With the visualization of wireless and some local knowledge, it is possible to pinpoint and reason about the spread of Wi-Fi.

Conclusions, Limitations, and Further Work

Overall, this paper has shown that it is possible to use the diffusion of innovations predictively with the new technology of Wi-Fi. In the area of spectrum regulation, this is a valuable goal because it holds the promise of reducing the uncertainty about creating blocks of license exempt spectrum. Further work extending this paper would use the model to predict congestion failure of Wi-Fi deployments in unlicensed spectrum, thus forecasting the health of the overall allocation block for unlicensed.

Still, these results have a number of limitations that must be mentioned. As the diffusion literature suggests that each technology's diffusion curve has a different shape, to reliably employ this method to reason about all unlicensed spectrum, all unlicensed technologies should be modeled, not just Wi-Fi. This would require a larger campaign of data collection for a variety of technologies. In these data, it is worth noting again that the time variable is "stacked" and not normally distributed. While the regression should be resistant to this source of bias, autocorrelation from repeated measure of the same geographic areas may inflate the reported R^2 values.

The Bad Neighborhoods of the Electromagnetic Spectrum

To return to the title of this paper, a major lesson from this research has been the large amount of local variation in unlicensed spectrum. While to some degree we should expect local variation when we are dealing with low powered devices, this insight has not yet reached much of the policy discourse about unlicensed. For example, some recent proposals for differential power limits divide the US into two broad kinds of geography, "urban" and "rural." This is clearly not a very helpful tactic.

Indeed, the lesson here is that for the electromagnetic spectrum, many "bad" neighborhoods are actually the "good" neighborhoods of unlicensed. That is, rather than granting an easement or a higher transmission power level to rural people, one could be granted to neighborhoods with low socioeconomic status or high age, as they much less likely to have congestion, even though they have a lot of people.

Sources Cited

- Abernathy, K. (2002). My Vision of the Future of American Spectrum Policy: Remarks of FCC Commissioner Kathleen Q. Abernathy before the 6th annual conference on Technology and Society. Washington, D.C., USA.
<http://www.fcc.gov/Speeches/Abernathy/2002/spkqa228.html>
- Allaway, A. W., Black, W. C., Richard, M. D. & Mason, J. B. (1994). Evolution of a Retail Market Area: An Event-History Model of Spatial Diffusion. *Economic Geography*, 70(1): 23-40.
- Associated Press, The. (2005, June 13). "T-Mobile: 450,000 people paid for Wi-Fi." *MSNBC*.
<http://msnbc.msn.com/id/8202465/>
- Bar, F. & Galperin, H. (2004). Building the Wireless Internet Infrastructure: From Cordless Ethernet Archipelagos to Wireless Grids. *Communications & Strategies* 54(2): 45-68.
- Benkler, Y. (1998). Overcoming Agoraphobia: Building the Commons of the Digitally Networked Environment. *Harvard Journal of Law and Technology*, 287.
- Benkler, Y. (2002). Some Economics of Wireless Communications. *Harvard Journal of Law & Technology* 16: 25.
<http://jolt.law.harvard.edu/articles/pdf/v16/16HarvJLTech025.pdf>
- Byers, S., & Kormann, D. (2003). 802.11b Access Point Mapping. *Communications of the ACM*, 46(5): 41-46.
- Castells, M. (1998). *End of Millennium. The Information Age: Economy, Society and Culture*, v. 3. Oxford: Blackwell.
- Dodge, M. & Kitchin, R. M. (2000). *Mapping Cyberspace*. London: Routledge.
- Dodge, M. & Kitchin, R. M. (2001). *An Atlas of Cyberspace*. London: Routledge.

- Dodge, M. & Kitchin, R. M. (2003). The role and value of maps of Internet infrastructure. In *Moving People, Goods and Information in the Twenty First Century: Urban Technology, the New Economy, and Cutting-Edge Infrastructure*, pp. 159-185. New York: Spon Press.
- Forman, C., Goldfarb, A., & Greenstein, S. (2003). The Geographic Dispersion of Commercial Internet Use. In S. Wildman & L. Cranor (eds.), *Rethinking Rights and Regulations: Institutional Responses to New Communication Technologies*, pp. 113-145. Cambridge: MIT Press.
- Gillett, S. E., Lehr, W. H., & Osorio, C. (2004). Local Government Broadband Initiatives. *Telecommunications Policy* 28(7/8): 537-558.
- Greenstein, S. (forthcoming). The Economic Geography of Internet Infrastructure in the United States. In M. Cave, S. Majumdar, I. Vogelsang (eds.) *Handbook of Telecommunications Economics, Volume II: Technology Evolution and the Internet*. Elsevier. <http://www.kellogg.northwestern.edu/faculty/greenstein/images/htm/Research/articles/geography.pdf>
- Greenstein, S. & Prince, J. (2004). The Geographical Diffusion of the Internet in the United States. In M. Singh (ed.) *The Practical Handbook of Internet Computing*, pp. 56.1-56.17. Boca Raton, FL: CRC Press.
- Greenstein, S., Lizardo, M., & Spiller, P. (2003). The Evolution of Advanced Large Scale Infrastructure in the United States. In A. Shampine (ed.), *Down to the Wire: Studies in the Diffusion and Regulation of Telecommunications Technologies*, pp. 3-36. New York: Nova Science.
- Grubestic, T. H. (2002). Spatial dimensions of Internet activity. *Telecommunications Policy* 26(7/8): 363-387.
- Grubestic, T. H., O'Kelly, M. E., & Murray, A. T. (2003). A geographic perspective on commercial Internet survivability. *Telematics and Informatics* 20(3): 51-69.

- Hales, P. (2005, March 29). "Wireless networks set to go through the roof: Worth billions soon, says report." *The Inquirer: Research and Markets*.
<http://www.theinquirer.net/?article=22198>
- Hampton, K. N., & Wellman, B. (1999). Netville On-Line and Off-Line: Observing and Surveying a Wired Suburb. *American Behavioral Scientist*, 43(3), 475-492.
- Hampton, K. N., & Wellman, B. (2000). Examining Community in the Digital Neighbourhood: Early Results from Canada's Wired Suburb. In K. Isbister (ed.), *Digital Cities: Technologies, Experiences, and Future Perspectives*, pp. 194-208. Heidelberg, Germany: Springer-Verlag.
- Hampton, K. N., & Wellman, B. (2003). Neighboring in Netville: How the Internet Supports Community and Social Capital in a Wired Suburb. *City and Community* 2(4): 277-311.
- Harrison, T. M., & Stephen, T. (1999). Researching and Creating Community Networks. In S. Jones (Ed.), *Doing Internet Research: Critical Issues and Methods for Examining the Net* (pp. 221-241). Thousand Oaks, CA: Sage Publications.
- Ileri, O., Samardzija, D. & Mandayam, N. B. (2005). Demand Responsive Pricing and Competitive Spectrum Allocation via Spectrum Server. *Proceedings of the First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*.
<http://www.winlab.rutgers.edu/pub/docs/Omer-Dyspan.pdf>
- Kling, R., & Iacono, C. S. (1995). Computerization Movements and the Mobilization of Support for Computerization. In S. L. Star (Ed.), *Ecologies of Knowledge: Work and Politics in Science and Technology* (pp. 119-153). Albany, NY: State University of New York Press.
- Lehr, W. (2004). The Economic Case for Dedicated Unlicensed Spectrum Below 3GHz. Paper presented to the New America Foundation Conference on Unlicensed Spectrum. Washington, D.C., USA.
http://itc.mit.edu/itel/docs/2004/wlehr_unlicensed_doc.pdf

- Lehr, W. & McKnight, L. W. (2003). Wireless Internet Access: 3G vs. Wi-Fi? *Telecommunications Policy* 27(5/6): 351-370.
- Legon, J. (2005, March 13.) "Get ready to tune in to wireless Net." *CNN*.
<http://www.cnn.com/2003/TECH/ptech/03/12/wifi.growth/>
- Matei, S. (2004). Wireless Networking Takes Off in Case-Study Town, Reflects National Trends. Unpublished manuscript. West Lafayette, IN: Purdue University.
http://matei.org/research/wireless/matei_wireless_pressrel.PDF
- Matei, S., Ball-Rokeach, S. J., & Qiu, J. L. (2001). Fear and Misperception of Los Angeles Urban Space: A Spatial-Statistical Study of Communication-Shaped Mental Maps. *Communication Research* 28(4): 429-463.
- Milgrom, P. (2004). *Putting Auction Theory to Work: From the Pioneer of Spectrum Auction Design*. New York: Cambridge University Press.
- Moss, M. L. & Townsend, A. (1997). Tracking the Net: Using Domain Names to Measure the Growth of the Internet in US Cities. *Journal of Urban Technology* 4(3): 47-59.
- Moss, M. L. & Townsend, A. (2000). The Internet Backbone and the American Metropolis. *The Information Society* 16: 35-47.
- Mueller, M. L. (1997). *Universal Service: Competition, Interconnection and Monopoly in the Making of the American Telephone System*. Washington, D.C: AEI Press.
- Mueller, M. L. & Schement, J. R. (1996). Universal service from the bottom up: a study of telephone penetration in Camden, New Jersey. *Information Society* 12(3): 272-292.
- Negroponte, N. (2002, October). Being Wireless. *Wired* 10(10):
<http://www.wired.com/wired/archive/10.10/wireless.html>.
- Noam, E. M. (1995). Taking the Next Step Beyond Spectrum Auctions: Open Spectrum Access. *IEEE Communication Magazine* 33(12): 66-73.
- Norris, P. (2002). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. New York: Cambridge University Press.

- O'Neil, D. (2002). Assessing Community Informatics: A Review of Methodological Approaches for Evaluating Community Networks and Community Technology Centers. *Internet Research*, 12(1), 76-102.
- Peha, J. M. (1998). Spectrum Management Policy Options. *IEEE Communications Surveys* 1(1): 1-8.
- Peha, J. M. (2000). The Path Toward Efficient Coexistence in Unlicensed Spectrum. Contributed to The *IEEE 802.16 WirelessHUMAN Study Group* (802.16hc-00/03).
http://www.ieee802.org/16/human/contrib/80216hc-00_03.pdf
- Peha, J. M. (2005). Approaches to Spectrum Sharing. *IEEE Communications Magazine* 43(2).
<http://www.comsoc.org/ci1/Public/2005/Feb/cireg.html>
- Pool, I. de Sola. (1983). *Technologies of freedom*. Cambridge, MA: Belknap Press.
- Rogers, E. M. (1995). *Diffusion of Innovations*. (4th ed.) Chicago: Free Press.
- Rogers, E., Collins-Jarvis, L., & Schmitz, J. (1994). The PEN Project in Santa Monica: Interactive Communication Equality and Political Action. *Journal of the American Society for Information Science*, 45, 401-410.
- Sahai, A. (2005). Some Fundamental Limits on Cognitive Radios and Their Implications. Paper presented to the Quello Center Research Conference on the Economics, Technology, and Policy of Unlicensed Spectrum. East Lansing, MI, USA.
- Sandvig, C. (2004). An Initial Assessment of Cooperative Action in Wi-Fi Networking. *Telecommunications Policy* 28(7/8): 579-602.
- Satapathy, D. P. & Peha, J. M. (1997). Spectrum Allocation Without Licensing: Opportunities and Dangers. In: G. L. Rosston & D. Waterman (eds.), *Interconnection and the Internet*, pp. 49-75, Mahwah, NJ: Lawrence Erlbaum Associates.
- Sawhney, H. (2003). Wi-Fi networks and the rerun of the cycle. *Info*, 5(6), 25-33.
- Schiller, H. (1996). *Information Inequality*. New York: Routledge.

- Schmitz, J., Rogers, E., Phillips, K., & Paschal, D. (1995). The Public Electronic Network (PEN) and the Homeless in Santa Monica. *Journal of Applied Communication*, 23(26-43).
- Schramm, W. & Roberts, D. F. (eds.). (1971). *The Processes and Effects of Mass Communication*. (rev. ed.). Urbana, IL: University of Illinois Press.
- Schuler, D. (1994). Community Networks: Building a New Participatory Medium. *Communications of the ACM*, 37(1), 39-51.
- Schuler, D. (1996). *New Community Networks: Wired for Change*. New York: Addison-Wesley.
- Scott, B., Chesley, R. E., Lakshmipathy, N., Ramachandran, K., Barranca, M. (2005). Profiles of Municipal and Community Broadband Networks. *New America Foundation Spectrum Policy Program Publications*. Washington, D.C.: New America Foundation.
http://www.newamerica.net/Download_Docs/pdfs/Doc_File_2245_1.pdf
- Shah, R. & Sandvig, C. (2005). Software Defaults as De Facto Regulation: The Case of Wireless Routers. Paper Presented to the 33rd Annual Telecommunications Policy Research Conference (TPRC) on Communication, Information, and Internet Policy, Arlington, Virginia, USA.
- Spectrum Policy Task Force. (2002, November). *Final Report of the Spectrum Policy Task Force*. Washington, DC: Federal Communications Commission.
http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-228542A1.pdf
- Straubhaar, J., LaPastina, A., Lentz, B., Main, S., & Taylor, J. (2000). *Structuring Access: The Role of Public Access Centers in the "Digital Divide"* (Report of the Telecommunications and Information Policy Institute). Austin, TX: The University of Texas.
- Streeter, T. (1996). *Selling the air: A critique of the policy of commercial broadcasting in the United States*. Chicago: University of Chicago Press.
- Tarde, G. (1903) *The Laws of Imitation*. E. C. Parsons (tr.). New York: Henry, Holt and Co.
- Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton Press.

Werbach, K. (2002, October). Open Spectrum: The New Wireless Paradigm. *Spectrum Series*

Working Paper #6. Washington, D.C.: New America Foundation.

http://werbach.com/docs/new_wireless_paradigm.htm

Wi-Fi Alliance, The. (2005). "Wi-Fi Backgrounder." [http://www.wi-](http://www.wi-fi.org/OpenSection/backgrounder.asp?TID=5)

[fi.org/OpenSection/backgrounder.asp?TID=5](http://www.wi-fi.org/OpenSection/backgrounder.asp?TID=5)

Zhang, X., Fan, S., & Cai, X. (2002). The path of technology diffusion: Which neighbors to

learn from? *Contemporary Economic Policy* 20(4): 470-478.

Zook, M. A. (2000). The Web of Production: The Economic Geography of Commercial

Internet Content Production in the United States. *Environment and Planning A* 32: 411-426.

Zook, M. A. (2001). Old Hierarchies or New Networks of Centrality? The Global Geography

of the Internet Content Market. *American Behavioral Scientist* 44(10): 1679-1696.

Yin, R. K. (1994). *Case Study Research: Design and Methods* (2nd ed.). Thousand Oaks, CA:

Sage.