The Web of Law

THOMAS A. SMITH*

TABLE OF CONTENTS

I. INTRODUCTION TO NETWORK THEORY ............................................................... 316
   A. From Random Graphs to Scale-Free Networks ........................................... 317
   B. Fitness and Network Evolution ................................................................. 322

II. THE WEB OF LAW .............................................................................................. 324
   A. The Distribution of Legal Citation Frequency .......................................... 324
   B. All Federal and State Cases as a Group ..................................................... 325
   C. Particular Federal Courts ........................................................................... 329
   D. State Courts ............................................................................................... 334
   E. The Citation Distribution of Legal Scholarship ....................................... 335
   F. The Analogy Between the Web of Law and the Physics Literature Network .......................................................... 336

III. CONSEQUENCES, APPLICATIONS, AND FUTURE RESEARCH ......................... 341
   A. Is the Web of Law a “Small World”? ......................................................... 343
   B. Legal Clusters ............................................................................................ 345

* Professor of Law, University of San Diego. Special thanks to LexisNexis without whose very generous help and skills this article would not have been possible. Jane W. Morris, Director, Customer Programs, and Mike Colangelo, Senior Software Engineer, both of LexisNexis, made the empirical part of this project possible with their invaluable assistance. I am also grateful to Professor Kwan-Liu Ma and Michael Ogawa of the Department of Computer Science, University of California, Davis, and Tim Yost, Consulting Software Engineer, LexisNexis, for their ongoing work on the visualization dimension of this project. Thanks to Larry Solum for many helpful comments and encouragement, and to Bob Hillman for his suggestions and tough questions. Thanks to Albert-László Barabási, Michael Rappaport, Ethan Katch, Laurence Lessig, Judge Richard A. Posner, David Post, Brett McDonald, Lior Strahilevitz, and Cosma Shaliz for useful comments. The opinions, inventions, conjectures, and conclusions, as well as any errors in this article, are solely the author’s own and not those of LexisNexis or any of those who provided suggestions or comments. This article is dedicated to TI. © 2004-05 Thomas A. Smith.
Understanding networks is crucial to many sciences, from physics and biology to economics and sociology. However, one of the largest, most accessible, and best documented human-created networks in existence—the centuries-old network of cases and other legal authorities where lawyers discover the law on any given topic—has received much less attention than other less consequential networks. The legal citation network, what I call the “Web of Law,” consists of cases, statutes and

2. W. Souma et al., Complex Networks and Economics, 324 PHYSICA A 396 (2003).
other legal authorities, and the citations that link them together. By studying its overall shape, we can learn new things about how law is organized and evolves.

This essay is called “The Web of Law” because the overall topology, or mathematical structure, of the Web of Law closely resembles that of the World Wide Web. Both the World Wide Web and the Web of Law are “directed” networks,\(^5\) have grown organically to a large size, and evince striking features of self-organization. Applying network analysis to the Web of Law yields insights into the overall structure of law that are of significant jurisprudential interest.

Some of the most interesting recent work on networks concerns the overall shape of the World Wide Web. Physicists such as Albert-László Barabási have argued that the Web\(^6\) is a “scale-free” network, a term

---

\(^5\) A directed network is one in which the links run one way only. In a citation network, the citing work cites the cited work, but not the other way around. A collaboration network is undirected. If A collaborates with B, then B also collaborates with A. See Albert-László Barabási, Linked 165–67, 169 (2002).

\(^6\) In this essay, I use “Web” to refer to the World Wide Web. I refer to the legal citation network generally and the American legal citation network in particular, as the “Web of Law.” The phrase “The Web of Law” has been used before, though I did not
explained in more detail below. Whether one calls the Web’s shape “scale-free” or prefers some other characterization, it is clear that some of the striking features of the Web are also conspicuous in the Web of Law, as may be seen in the evidence this essay will present. While previous legal citation studies have looked at particular questions of judicial influence, or how doctrines change over time, the study discussed in this essay begins to examine the overall shape of the American legal citation network. This approach takes its inspiration from the work of statistical physicists, who have applied their techniques to study the shape of many different kinds of networks, including the citation networks of physics papers. This approach is not without controversy. Sociologists pioneered important aspects of network theory and seem greatly to resent the intrusion of physicists into their realm. Some mathematicians, on the other hand, consider the approach of the physicists too rough and ready. These criticisms are not entirely unjustified. However, exploring the global structure of the Web of Law must begin somewhere, and the moniker “essay” here signals the preliminary nature of this attempt to do so. And the preliminary results are sufficiently interesting to merit presenting them to a general legal readership.

Thus, this essay presents highlights from a large citation study done by me and employees of LexisNexis, one of the leading electronic legal research providers. In an effort to capture the overall shape of the Web of Law, this study covers nearly all U.S. federal and state cases, more...
Preliminary results strongly suggest that the American case law network has the overall structure that network theory predicts it would: a structure that visually and in general terms appears much like that of the Web and other citation networks, such as those of scientific papers. It shows that this structure, however its precise mathematical structure may ultimately be characterized, is present at virtually every jurisdictional level of our legal system, from the U.S. Supreme Court to the lower state courts. Virtually all of the jurisdictions I have examined have the very highly skewed distributions of citation frequency that are classically found in studies of citation networks, such as those of scientific papers, and many other real networks. While these highly skewed distributions are familiar to students of scientific citations, my study is the first (to my knowledge) that shows that this phenomenon is ubiquitous in American common law.

Whether one looks at U.S. Supreme Court cases, federal appellate or district court cases, bankruptcy or National Labor Relations Board cases, or state supreme or intermediate appellate court cases, American common law has everywhere this recognizable, highly skewed distribution of citations. This should have significant consequences for how we think about American law and the American legal system, and indeed other precedent-based, common law systems as well. Many lawyers and legal scholars will be surprised to learn, for example, how relatively very few cases get the vast majority of all citations, while most cases are never or rarely cited. This is part of what it means that the citation frequency distributions I examined were “highly skewed.” Thus, of approximately four million federal and state cases, about 400,000 cases are not cited at all. Another 773,000 are cited only once. The number of cases with a given number of citations falls off rapidly as the number of citations increases. As measured by citation frequency, therefore, precedential authority is extremely concentrated in a relatively very small core of cases and secondary authorities. This sort of pattern (though not necessarily so skewed) may be found in many other real networks—not

---

13. This study was undertaken with the help of LexisNexis, using its Shepard’s citation service.
14. The only jurisdictions that do not evince the highly skewed distributions discussed below are those with too few cases to make a distribution.
15. See database on file with author.
only citation networks, but also the Web, networks of cellular proteins, film actors, and many others. 16

The network properties of the Web of Law are not just curiosities (though they are that). On reflection, it is clear they are jurisprudentially important as well. 17 Important features of the legal system can be inferred from their study, features that reveal themselves only by viewing law’s global network structure from a macro-level perspective. Common law systems are networks and must obey the mathematics of networks. By analogy, physical forces such as gravity determine the way massive objects behave and do so in ways describable by mathematical laws. These laws dictate that stars and planets, for example, will be spheres, not cubes, and planetary orbits elliptical, not circular. Similarly, common law legal systems—systems based on cases and other authorities that grow in number over time, and in which cases cite earlier cases and authorities—will also inevitably have a characteristic structure. Every federal Article III jurisdiction and every state supreme court jurisdiction, for example, has the same basic pattern of citation frequency distribution even though these courts operate largely independently of each other. In each jurisdiction there are relatively very few cases that are cited very frequently, and a large majority of rarely or never cited cases. Not only does each jurisdiction have this division between the often and rarely cited, but the curves describing these relationships are very similar in shape from jurisdiction to jurisdiction, as the reader may see for herself in Part II.

While this highly skewed distribution of authority is conspicuous in the Web of Law, there are other interesting properties that other real networks evince and which future research may well discover in the

16. The link degree distribution graphs of parts of the Web of Law shown in Part II are typical of the link degree distributions found in many networks. Scientists and mathematicians disagree, however, about how the shape of these curves is best described. Some network scientists, mostly physicists, follow the Barabási approach and call them “scale-free” distributions. See Barabási, supra note 5, at 86–87. Others, often mathematicians and statisticians, argue that the distribution is better characterized as a stretched exponential, or in some other way. See P.L. Krapivsky et al., Connectivity of Growing Random Networks, 85 PHYSICAL REV. LETTERS 4629, 4629 (2000). A recent article, which appears to very carefully evaluate the citation network of physics papers, argues that the distribution is best characterized by two power law equations. See S. Lehmann et al., Citation Networks in High Energy Physics, 68 PHYSICAL REV. E. 026113-1 (2003). Part II discusses this controversy in more detail. However, this technical controversy should not detract from the importance of discovering that this highly skewed distribution, with its characteristic shape, permeates the Web of Law at every level. Important jurisprudential consequences follow regardless of how the shape is ultimately characterized.

17. By “jurisprudentially important” I mean significant to our understanding of law and how it works at a high level of generality. I do not mean to suggest that the network properties of law raise any special philosophical questions.
Web of Law. For example, many real, highly skewed networks, such as scientific collaboration networks and the Web, are “clustered.” That is, the nodes are not all uniformly linked to each other, but rather, some are densely connected to each other in clusters, and these clusters only more loosely connected to one another. If this is also true for the Web of Law, as seems likely, it would not be accurate to say that the law is, in Frederic Maitland’s famous phrase, a “seamless web,” at least not if seamless means smooth. Law would instead be an uneven, clumpy web, with some parts thickly connected within themselves, but only loosely connected to other parts. If common law systems organize themselves into clusters in this way, then they would have an organic structure that would be discoverable through network analysis. This should be of great interest to scholars of legal systems and may have important practical implications as well. Another feature many real, highly skewed networks display is a surprising degree of integration, the so-called “small world” phenomenon. The Web of Law, though large, is probably


20. The origin of this phrase is unclear. Ethan Katsh explains that:


21. It is very probably also true of the Web of Law generally. Clustering analysis needs to be done on the Web of Law as a whole to prove this.


rather well integrated, at least in the sense of being a “small world.” The research proving this remains to be done, but given that the Web of Law appears to be like other highly skewed, real networks in other respects, it seems likely it will have this feature as well. Measuring the “diameter” of the Web of Law, and different parts of it, could tell us how well integrated it is overall, and how well integrated different parts of it are with each other. The relative fewness of important cases (the extreme skewedness of the network), clustering, and integration are illustrations of actual or probable network properties of the Web of Law that deserve more research.

This essay has three parts. Part I introduces some basic concepts of network science, such as nodes, links, random graphs, evolving networks, scale-free networks, small worlds, the preferential attachment or “rich get richer” dynamic, node fitness, and clusters. These ideas are explained intuitively. Part II presents highlights of the study alluded to above, demonstrating that the network of American case law is a highly skewed network,24 very similar to the World Wide Web25 and others.26 Part III explores some implications of the study, and provides some possibly fruitful areas for future research.

I. INTRODUCTION TO NETWORK THEORY

A network is just a set of items, termed nodes or vertices, with connections among them, termed links or edges. Networks are mathematical objects, but there are concrete examples everywhere. There are social networks of friends and acquaintances, economic networks of producers

26. The study of scale-free networks in particular and complex networks in general has generated a large literature. A helpful bibliography of network-related literature generally may be found in Cosma Shalizi’s online notebook on complex networks. See Cosma Shalizi, Complex Networks Notebook (July 20, 2007), http://www.cscs.umich.edu/~crshalizi/notebooks/complex-networks.html#bulletin; S. N. Dorogovtsev & J. F. F. Mendes, Evolution of Networks 31–54, 80–81 (2003) (summarizing studies of networks for degree distribution, including citation networks). The more than four million node U.S. legal citation network analyzed in the Smith/LexisNexis study appears to be the largest by number of nodes by some margin. However, our study measured only “in degree.”

316
and customers, and scientific networks of research collaborators. Networks also abound in nature. Blood vessels form a distributional network, the brain contains a neural network, and predators and prey in ecosystems form food networks, to name just a few examples. The Web of Law is the network that consists of cases and other legal authorities, such as statutes, treatises, and law review articles (the nodes), and the citations that link them to one another.

A. From Random Graphs to Scale-Free Networks

Mathematicians have studied networks, in the form of graph theory, at least since Leonard Euler’s solution of the Koningsberg Bridge problem in 1736. The renaissance of contemporary graph or network theory can be dated to 1959, when Erdos and Renyi began publishing a series of eight important papers on random graphs. Random graphs are different from the networks we are concerned with, but are a good place to begin the exposition. To construct a random graph, we can begin with fifteen nodes, as shown below. From these fifteen nodes, we select two at random, and establish a link between them. We carry on this procedure for some specified number of links, each time picking randomly the two nodes to be connected.

29. In this essay, I concentrate on cases and not other legal authorities. However, this analysis could be extended to include other legal authorities including statutes and legal scholarship. I hope to do this in future research.
32. Barabási & Oltvai, supra note 1, at 105.
Erdos and Renyi proved that in a large random graph, each node will have approximately the same number of links, or, to use network terminology, will be of approximately the same “link degree.” Link degree in a random graph, they showed, follows a Poisson distribution, which roughly resembles the familiar normal, or bell curve, distribution. The mean of this distribution tells us how many links a typical node in a random graph or network has.

Many important networks, however, are far from random. Instead of nodes having approximately the same number of links, a few nodes have many links, while most nodes have only a few. This structure emerged dramatically when Barabási and Albert (BA) studied the network structure of the World Wide Web. Beginning their investigation, they expected to find that the number of links running into each web page followed a Poisson distribution, as in a random network. Instead, they found that the number of links followed a power law distribution.

---

33. Barabási & Oltvai, supra note 1, at 105.
34. See Barabási, supra note 5, at 21–22.
35. For background on Poisson distributions, see generally John J. Heldt, Quality Sampling and Reliability: New Uses for the Poisson Distribution (1998).
BA coined the term *scale-free* to describe this sort of network. Scale-free alludes to the fact that there is no typical scale of link degree—as no node can be said to have the typical number of links—as there is in a random network (where it is the mean of the Poisson distribution). The explanation hinged on two features that distinguish scale-free networks, such as the World Wide Web, from random networks. First, random networks begin with a fixed number of nodes. Links are then added to connect randomly selected nodes. Thus each node in the random network has a certain probability of being chosen as one of the next pair that will get a new link between them. In the World Wide Web, however, and in many other both human-created and natural real networks, the number of nodes is not fixed. The World Wide Web grew over time and new web pages are created every day. (In the Web of Law, by analogy, cases were decided over decades and centuries and new cases are decided every day.) Scale-free networks are created by a dynamic process in which the number of nodes grows over time.

---

Figure 3. Scale-free networks form when networks grow over time and links are formed by preferential attachment.\textsuperscript{39}

The second important feature of scale-free networks, BA observed, was that new links are not added randomly. In random networks, each node has the same probability of getting a new link added to it in the link-creation process. In a scale-free network, the probability that a node will acquire a new link depends on how many links it has already. Nodes with more links have a greater chance than nodes with fewer links of acquiring additional links as the network grows. In terms of links, the “rich get richer,” a mechanism BA called “preferential attachment.” They proved that the combination of these two features, network growth and preferential attachment, produces the scale-free network structure, with its power law distribution of node degree, and the domination of the network by a few nodes with very many links.\textsuperscript{40}

Many regarded this as a significant discovery in network science, while others thought it merely confirmed known processes. However one assesses its originality, however, BA’s work on the structure of the World Wide Web helped shift the study of networks away from the Erdos-Renyi random graph model to real networks and their properties. Following BA’s lead, researchers soon discovered many other real networks had a scale-free or other highly skewed link degree distribution, among other interesting properties. Graphs for several of these networks are shown below.

\textsuperscript{39} Barabási & Bonabeau, \textit{supra} note 36, at 55.
\textsuperscript{40} Barabási, \textit{supra} note 5, at 86–87.
Figure 4(a) (left). Degree distribution of out-degree (links running out from web pages) from a portion of the web consisting of 200 million web pages, in a log-log format. Figure 4(b) (right) shows the distribution of in-links (links running into web pages) in a sample of the Web. Note that the y-axis shows proportion of web sites.41

Figure 4(c) (below left). Degree distribution of links between film actors. Actors are nodes in the network. Two actors are linked if they both appeared in the same film.42 Log-log format. Figure 4(d) (below right) shows degree distribution in the “language network.” Log-log format.43

42. Barabási & Albert, supra note 25, at 510.
43. Ramon Ferrer Cancho & Ricard V. Solé, Two Regimes in the Frequency of Words and the Origins of Complex Lexicons: Zipf’s Law Revisited, 8 J. QUANTITATIVE LINGUISTICS 165, 165. See also Albert & Barabási, supra note 10.
In the four networks shown above, namely, (a) out-links (that is, links running out of) web sites, (b) in-links (links to) websites, (c) film actor collaborations, and (d) word frequency, the highly skewed distribution of links is conspicuous. All of the graphs are in a log-log format. There is very lively controversy over whether these distributions are best described with power law formulas, stretched exponentials, or with some other formula. These technical issues, while important, need not detain us at this stage. The main point is to see the general shape these graphs share and to compare that with the shape of Web of Law distributions in Part II. In the upper left of each of the graphs in Figure 4, we see that there are many nodes that get few links, while in the lower right, relatively few nodes with many links. The BA preferential-attachment-plus-growth mechanism is one plausible explanation for naturally occurring highly skewed networks. However, there are features of many real networks, including the Web of Law, which it does not explain. Nevertheless, it has the advantage of being simple, elegant, and powerful, and so it serves well to introduce the application of network theory to the Web of Law.

B. Fitness and Network Evolution

While BA’s model of scale-free networks resembles the Web in overall structure, it is unrealistic in important respects. The model’s mechanism of preferential attachment assures that the oldest nodes will always have the most links. Yet for the Web and many other networks, this is often not true. There are many prominent exceptions, such as Google, one of the most widely used Web search engines. Google is much newer than many Web sites, yet far more popular than most. Some nodes obviously acquire links through some mechanism or mechanisms other than preferential attachment. Various mechanisms may explain this. In a model suggested by Barabási and Bianconi, these unusually attractive nodes are more “fit”

44. In the log-log format, the values on the respective axes increase logarithmically—so, 10, 100, 1000, and so on. This exposition makes the power law and similar curves much easier to study.
45. See infra Part II.F (briefly discussing the model that seems to be the best candidate for fitting the distributions of the citations graphs).
46. The plot appears “fat” in some graphs in the lower right because the logarithmic scale puts the same number of nodes that have, for example, over 1000 links on the same line.
47. It is important to note that many different processes can produce power law distributions. See Mark Newman, The Power of Design, 405 Nature 412, 412–13 (2000). Growth plus preferential attachment is just one of them.
than others. Bianconi discovered that the mathematical model physicists use to describe a quantum physical phenomenon called Bose-Einstein condensation (BEC) also describes the evolution of networks in which nodes have different levels of fitness. Using this model, one can describe how much the distribution of links in the network is due to rich nodes getting richer (preferential attachment), and how much is due to the fit nodes getting richer (more competitive nodes garnering more links). The ability to measure whether the winning nodes in a network are winning because they are old, or because they are fit, or some combination of the two, would be very useful. Network scientists are developing network evolution models that take differences in node fitness into account.

* * *

Part I introduced some basic concepts of network theory. Part II below presents representative highlights of a large citation study I and LexisNexis employees performed which confirms the ubiquity of a recognizable highly skewed citation distribution in nearly all jurisdictions in the American legal citation network. How this distribution, which is readily visible in the graphs, is best characterized mathematically, remains for future work. A pure power law, scale-free distribution, such as BA suggested for in-link (citation) degree in the Web, appears not to be a very good fit for the Web of Law. A two power law distribution (discussed in more detail below), which some scientists argue characterizes the high energy physics paper citation network, seems, on the other hand, quite promising. However, the main point of the following part is to show that an important empirical regularity permeates the American legal citation network, a fact that has not been demonstrated before this study. Part III suggests some of the consequences that follow from the structure of the American legal network that this evidence reveals and suggests some promising avenues for future research.


49. See, e.g., Alain Barrat et al., Weighted Evolving Networks: Coupling Topology and Weight Dynamics, 92 PHYSICAL REV. LETTERS 228701 (2004); Hyun-Joo Kim et al., Weighted Scale-Free Network in Financial Correlations, 71 J. PHYSICAL SOC’Y JAPAN 2133 (2002); S.H. Yook et al., 86 PHYSICAL REV. LETTERS 5835 (2001); Dafang Zheng et al., Weighted Scale-Free Networks with Stochastic Weight Assignments, 67 PHYSICAL REV. E. 040102 (2003).
II. THE WEB OF LAW

Part I introduced some basic concepts of network theory and explained some of the properties different networks have. This part presents evidence about the network of American case law.50

A. The Distribution of Legal Citation Frequency

BA’s investigation of the structure of the World Wide Web revealed that the distribution of links to Web pages (“in degree”) did not fit a Poisson curve as predicted by the random graph model, but rather had something much closer to a power law distribution. While the best exact characterization of the degree distribution of the World Wide Web and other large, real networks remains controversial, it is reasonably clear that the Web and many other large, real networks, such as those shown in Figure 4, have power law tails, such that most nodes have only a few links and a few nodes have most of the links.51 In the BA model, nodes are more likely to get additional links depending on how often they have been linked to before, a mechanism BA calls “preferential attachment.”52 This is analogous to the concept of legal authority. Cases that have been cited approvingly by judges in the past are seen as authoritative and are therefore more likely to be cited in the future. While sometimes new cases gain authority quickly, preferential attachment seems, as a first approximation, a plausible model of how cases accumulate legal authority (measured as number of citations).

With this in mind, I proposed to LexisNexis, the owner of the well-known Shepard’s citation service,53 that we conduct a study to determine the citation frequency of some large sample of U.S. cases, and hypothesized that it would follow a power law distribution. It proved possible to measure the citation frequency of virtually all U.S. cases, state and federal, both as a whole and as organized into the approximately three hundred different state and federal jurisdictional categories that

50. As Part II.F, infra, explains in more detail, the most promising statistical description of the Web of Law appears at present to be a two power law model, similar to that of the high energy physics paper citation network as described by Lehmann et al., supra note 16. This may be thought of as a network with two regions, one full of dead nodes, which are cases that have been rarely cited and will not be cited again, and another region of “living” precedents containing cases that have been cited relatively many times and are gaining more citations through preferential attachment. See infra text accompanying notes 69–72.
51. Part III discusses some respects in which the degree distribution departs from a power law distribution.
52. See Bianconi & Barabási, supra note 48.
Shepard’s tracks. The following pages set out graphs showing the citation frequency distributions of all state and federal cases as a group, of all state cases and federal cases respectively, and of some important or representative federal and state jurisdictions. The graphs are intended largely to speak for themselves. Readers will see that in every jurisdiction shown, a very similar pattern of citations may be seen. Part II.E discusses some possible interpretations and implications of this ubiquitous pattern. However, it should be stressed that the discovery of a ubiquitous distribution (or closely similar distributions) in the American legal citation network is a remarkable empirical phenomenon very much worthy of note in its own right.

**B. All Federal and State Cases as a Group**

Figure 5 below shows the citation frequency (link degree) distribution for the set of all U.S. federal and state (including D.C.) cases in the Shepard’s database, about four million cases in all.\(^{54}\) Simple inspection reveals that the distribution bears a marked resemblance to the distributions that appear in Figure 4 above, and also, and especially, to the distribution of the high energy physics paper network, shown in Part II.E below.\(^{55}\) As explained in Part II.E below, the Web of Law distribution appears to resemble more closely the physics paper citation distribution, which is well described by a two power law distribution.

The Web of Law data evinces the highly skewed distribution often seen in citation networks, such as has been described in scientific paper citation networks.\(^{56}\) However, it has never been shown before to be present virtually universally in American law. This is jurisprudentially significant. It shows how, whatever the jurisdiction, relatively few cases are cited frequently, while the large majority are infrequently cited. The extent of the concentration of legal authority (at least as measured by citation frequency) in only a relative few cases is remarkable. For all

---

\(^{54}\) Database on file with author.

\(^{55}\) The solid trend line shows a power law distribution, and in the upper right of the chart, the R-squared fit of the actual distribution to the power law trend line is given as approximately 0.832. I acknowledge that the mere calculation of the R-squared to a power law trend line of a link degree distribution gives merely a crude fit. This crude fit suggests some refined model could do better. Visually, the distributions show significant curvature, and the similarity to the physics paper distribution suggests a two power law curve would give a much better fit. I plan to do this in a future article.

\(^{56}\) See infra Part II.E.
federal and state cases that have ever been cited, for example, approximately 
four million total cases, 773,000 are cited only once. The number of 
cases falls off rapidly as the number of citations increases. A total of 
1,210,766 cases, or about 30\% of the total of cited cases, have been cited 
eten or more times. About 130,949 cases, a scant 3\% of the total of cited 
cases, have been cited fifty or more times. About 1322 cases, only .03\% 
of the total of cited cases, have been cited 1000 or more times. In terms 
of the number of cites garnered by the most highly cited cases, the 
distribution is very skewed. Putting aside problems with treating citation 
frequency as a measure of jurisprudential influence, this indicates that 
the vast majority of influence is concentrated in a relatively small number 
of cases.

57. For technical reasons, cases receiving zero citations are not shown in the 
graphs. Their number is calculated by subtracting the total number of cases with 
citations from the total number of cases.

58. For problems with using citations to measure impact of articles in the social 
sciences, see Antony J. Chapman, *Assessing Research: Citation-Count Shortcomings*, 

59. Indeed, the distribution is so skewed it raises the question of whether the Web 
of Law might have undergone, or be close to undergoing, what Barabási and Bianconi 
describe as an analog of Bose-Einstein condensation in a network. When this 
“condensation” occurs, the network develops one or a few dominating nodes. *See*
Bianconi & Barabási, *supra* note 48, at 5634; *see also infra* Part III.A (discussing 
integration and disintegration of networks).
Figure 5. The citation frequency distribution of American case law (federal, state, and D.C. cases in the Shepard’s database).

A visually very similar, highly skewed citation frequency distribution is also present in the sets, respectively, of all federal and all state cases, as Figures 6 and 7 show below.

60. An important article that noted power law-like distributions in citations is Post & Eisen, supra note 4. Post and Eisen explain how common law style decision-making would tend to produce citation frequency distributions with a power law form. See id. This is an important insight. Interestingly, important new research in complex networks suggests a deep link between some scale-free networks and fractal geometry. See Chaoming Song et al., Self-Similarity of Complex Networks, 433 NATURE 392 (2005). Whether the Web of Law is such a scale-free fractal network awaits future research.
Power Law Distribution for All Federal
(2,208,400 cited cases)

\[ y = 96385.207x^{-1.476} \]
\[ R^2 = 0.754 \]

Figures 6 and 7. Citation frequency distribution for all federal and state cases, respectively.
C. Particular Federal Courts

Federal courts consist of the Supreme Court at the apex, courts of appeal, district courts, and finally various specialized federal courts, such as federal bankruptcy courts. Interestingly, as the graphs below indicate, visually very similar, highly skewed citation frequency distributions appear at every level of federal jurisdiction. Figure 8 below shows the distribution for the U.S. Supreme Court.

![Power Law Distribution for U.S. Supreme Court](image)

Figure 8. Citation frequency distribution for the U.S. Supreme Court.

On inspection, the distribution for the Supreme Court appears to be a highly skewed distribution similar to those for all federal and state cases.61 About 28,000 Supreme Court decisions have been cited only...

---

61. Professor Daniel Farber argues that the power law distribution of citations in certain sets of cases is consistent with what he terms a “tectonic” model of statutory interpretation. Daniel A. Farber, Earthquakes and Tremors in Statutory Interpretation: An Empirical Study of the Dynamics of Interpretation, 89 MINN. L. REV. 848, 857–58
Using the data from Shepard’s provided by LexisNexis, a total of 2,839,156 “citing references” are divided among 107,874 Supreme Court cases that are cited at least once. (If Case A cites Case B one or several times, that counts as one citing reference.) If cases that receive more citing references are thought of as more authoritative, we can see that authority is concentrated in a relatively few opinions, and that most opinions have relatively little authority. Almost 68% of cited opinions

(2005). The reference comes from the Gutenberg-Richter power law, which describes the distribution of earthquake magnitudes. There are many earthquakes of small magnitude and a few of large magnitude, and the distribution of earthquakes fits a power law distribution. The mechanism that produces this distribution is thought by some to be phenomena of self-organized criticality, which is also evidenced by power law distributions. See Mark Buchanan, Ubiquity 43–62 (2000). Farber argues that power law distributions in statutory interpretation cases may indicate that a kind of rupturing, rather than gradual process, characterizes how legal interpretations of important texts, such as statutes and constitutions, change over time. See Farber, supra, at 876. However, Farber’s thesis, while it may be correct, is not supported by the data he presents. In fact, as Figure 8 shows, all U.S. Supreme Court cases are approximately power law distributed, as indeed are cases in virtually all jurisdictions. It is thus difficult to come up with any subset of Supreme Court cases or those of any other court that is not approximately power law distributed. Simply counting by hand, I have not, for example, been able to come up with any search, including nonsense searches such as random numbers, which do not produce search results that have highly skewed citation frequency distributions, that appear on inspection to be roughly power law or exponentially distributed. That Farber’s results showed statutory interpretation searches had similar results does not show that they are in any way special.

It may be true that shifts in constitutional and statutory interpretation usually occur, as a historical matter, more often as ruptures or paradigm shifts than as a result of a gradual process. However, the fact that the citation frequencies of the entire Web of Law are approximately power law distributed, and more particularly, that the in-degree of Supreme Court cases appears to be power law distributed, implies that the evidence Farber adduces does not in fact support his tectonic model any more than it supports a gradualist model. Recall that growth plus preferential attachment is a gradualist model and it generates a power law distributed citation frequency.

Network theory does suggest a way to test empirically whether some cases or interpretative approaches acquire authority suddenly. It is possible to measure the rate at which cases acquire citations, and compare that to the rate at which they would have done so, just by virtue of preferential attachment. Some cases might be seen to take off dramatically in terms of in-degree, analogously to the way in which the Google web site garnered links far faster than a mere preferential attachment mechanism could explain. Paradigm-shifting cases would appear therefore, in my model, as cases of extraordinarily high fitness, as that concept is explained above. See supra text accompanying notes 48–49 and infra text accompanying notes 95–96. If Farber’s hypothesis is true, some appropriately identified set of cases about constitutional interpretation should have significantly higher fitness than does a suitable set of control cases (such as all Supreme Court cases). A similar approach could be used to measure the fitness of legal scholarship, such as law review articles. I believe that it is likely that certain cases and law review articles would appear to have much greater fitness than most of their kind and would qualify as paradigm-shifting authorities, but only actually measuring fitness would reveal this for certain.

62. In future analyses, it may make sense to eliminate such routine matters, and see what the effect is on the overall distribution.
are cited ten or fewer times in the Shepard’s data. Cases receiving one hundred citing references or more comprise only 9.7% of all cited cases. If one included the probably large, but difficult to collect, number of never cited Supreme Court cases, this figure would be even smaller. In a limited dataset collected from the Findlaw website, the top 20% of U.S. Supreme Court cases, in terms of in-degree, garnered 65% of all citing references. Only a small percentage of Supreme Court cases thus exercise virtually all of its authority, at least as measured by citation frequency.

U.S. Courts of Appeal also have power law-like, highly skewed citation frequency distributions, as shown in the following graphs.

---

63. Database on file with author.
64. Database on file with author.
Figure 9. Six representative distributions of citation frequency for U.S. courts of appeal.

Federal district courts also show highly skewed distributions similar to those of the U.S. courts of appeal and the U.S. Supreme Court, as the graphs below illustrate.
Figure 10. Representative citation frequency distributions of federal district courts in the Ninth and Second Circuits.
D. State Courts

State judicial systems, like the federal system, have courts of different levels, including those of original jurisdiction and one or more levels of appellate courts. Interestingly, the citation distributions for the highest state courts are similar to those for the federal appellate courts, as the charts below illustrate.

![Power Law Distribution for New York Court of Appeals](image1)

![Power Law Distribution for California Supreme Court](image2)

![Power Law Distribution for Illinois Supreme Court](image3)

![Power Law Distribution for Idaho Supreme Court](image4)

Figure 11. Highest state appellate court citation distributions for New York, California, Illinois, and Idaho.
In all of the graphs above, the distributions are highly skewed and fit to some degree a power law trend line. However, there is also noticeable curvature in the distributions. The curvature makes the distributions appear concave to the origin. This curvature is also observable, interestingly enough, in the citation frequency distribution of the physics literature, which has been rigorously studied by physicists using network analysis techniques. Indeed, the distributions for the Web of Law and that of the physics literature appear remarkably similar. This suggests a potentially powerful analogy: that the dynamics of legal authority in the Web of Law is similar to the dynamics of scientific authority within a literature such as physics. Part II.F explores this analogy in more detail.

E. The Citation Distribution of Legal Scholarship

The citation degree distribution for legal scholarship published in law reviews and journals has a highly skewed distribution similar to those for cases, as Figure 12 below shows.

Figure 12. Citation distribution for law review articles and other secondary sources appearing in the Shepard’s LAWREV database.
The data graphed above consists of law review articles in the Shepard’s database, which covers about 385,000 law review articles, notes, and comments appearing in 726 U.S. law reviews and journals. The citations counted are those to such sources by either other law review articles or cases.

It is noteworthy first that the distribution for law review articles is visually indistinguishable from those in many jurisdictions for cases. This suggests the dynamics of authority for cases and for law review materials are similar. The degree of the skew is also remarkable. Forty-three percent of articles are not cited at all, and about 79% get ten or fewer citations. At the other end of the distribution, we see that authority in secondary sources, as in cases, is quite concentrated.

F. The Analogy Between the Web of Law and the Physics Literature Network

Figure 13 below shows the citation frequency distribution for a large set of articles in the High Energy Physics literature. The reader should note the roughly power law distribution, with curvature, is strikingly similar to the several representative distributions from various legal jurisdictions displayed above.

FIG. 1: An “atomic” histogram of the citation distribution of the total data set showing the normalized probability, $P(k)$, that a paper has $k$ citations, here plotted versus $k + 1$. The straight lines in the low and high citation regimes have slopes $-1.29$ and $-2.32$, respectively. Note the logarithmic scales.

Figure 13. Citation distribution of a large set of articles in the High Energy Physics literature.\textsuperscript{67}

The discussion of this distribution in Lehmann et al. is worth quoting at some length:

One of the most striking features of this data set is the large number of papers (some 29%) which are uncited, ... In the same vein, 74% of the papers in our network have ten or less citations. In contrast, 6.2% of the papers have 50 citations or more, and only 131 papers ($\approx 0.05\%$) are cited 1000 times or more. The mean number of citations in this sample is 14.6, which is considerably larger than the median of 2.3 citations, implying that a paper with the average number of citations is substantially more cited than the “average” paper. The large factor between mean and median citations suggests that the citation

\textsuperscript{67} Lehmann et al., \textit{supra} note 16, at 026113-2.
distribution has a very long tail with a small fraction of highly cited papers accounting for a significant fraction of all citations. This is indeed the case. Approximately 50% of all citations are generated by the top 4% of the all [sic] papers; the lowest 50% of papers generates only 2% of all citations. The rates of citation production by these two parts of the dataset differ by a factor of approximately 310. These observations regarding citations in SPIRES suggest that the citation distribution follows a power law. As we shall see, this is qualitatively correct.

Figure 1 [Figure 13 above] shows a log-log representation of the distribution of citations in the SLAC SPIRES database. The data suggest that this citation distribution is remarkably well described by two power laws. The distribution, \( N(k) \), is approximately proportional to \( (k + 1)^{-1.3} \) for \( 0 \leq k \leq 49 \) and to \( (k + 1)^{-2.3} \) for \( k \geq 49 \).68

Thus, the citation distribution for the high energy physics paper network appears highly skewed similarly to that of the Web of Law. Even more striking is how closely the graph of the high energy physics network visually resembles those from the Web of Law, and how well two power laws describe the physics paper distribution. Of course, the Web of Law data should be subjected to the same analysis, but as a preliminary matter it certainly appears that the legal citation network would also be well described by the two power law model.

Lehmann et al. explain the two power law shape of the high energy physics paper network distribution as follows:

We believe that these different power laws probably reflect differences in the underlying dynamics of citations in the high and low citation regions. That different dynamics rule the two regimes seems clear. The bulk of the papers in the minimally cited part of the distribution are “dead” in the sense that they have not been cited within the last year or more (and will probably never be cited again). Of course, this part of the distribution also contains vigorous young papers of high quality, whose citation count is increasing. However, dead papers vastly outnumber the live population. In the highly cited region, virtually all papers are still alive, with even the oldest of them acquiring new citations regularly. It seems highly likely that citation patterns for such papers are quite different from those of minimally cited papers that are most often cited only by the author and close co-workers.69

Lehmann et al. followed their 2003 paper with another paper pertinent to this topic. In their 2004 paper *Life, Death, and Preferential Attachment*, they write:

That progress in science is driven by a few great contributions becomes disturbingly clear when one considers citation statistics. The vast majority of scientific papers is either completely unnoticed or minimally cited. In high

---


energy physics, 4% of all papers account for 50% of the citations, while 29% of all papers are not cited at all.

... ... [I]t is an empirical fact that the vast majority of nodes in citation networks “die” after a relatively short time and are never cited again. A relatively small population of papers remains alive and continues to accumulate citations many years after publication; this is the main conclusion in [our 2003 paper].

Lehmann et al. develop a statistical model that accounts well for the curved shape of the high energy physics paper network. It is based essentially on the idea that there are two mechanisms at work, one by which most papers die off by not being cited, and a second, cumulative advantage mechanism, by which a relatively few papers survive and accumulate citations over time. The dead or dying papers are in the flatter region (to the upper left) of the distribution curve. Also in that region are newer (but vastly outnumbered) “vigorous” papers destined to become frequently cited papers, but which have not become so yet. In the steeper region of the distribution (to the lower right) are the relatively much more scarce papers that accumulate cites through preferential attachment.

The strong similarity between the Web of Law citation distributions and the high energy physics distributions suggests that scientific authority and legal authority share a similar structure and operate and evolve in similar ways. A few “great contributions” drive progress in both physics and law. A relatively few important decisions exercise the majority of legal influence and authority, and determine the direction of law, just as a few important scientific papers determine the direction and progress of physics (and probably other sciences as well).

71. There is no reason to believe that the high energy physics distribution is atypical of physics papers. For analysis of a much larger database of physics papers, with similar results, see Sidney Redner, *Citation Statistics from More Than a Century of Physical Review* (2004), http://arxiv.org/abs/physics/0407137.
72. That the citation networks of both a rigorous science and the law appear to be structured and evolve in similar ways is suggestive, but of what? One tempting claim would be that because legal authority has the same overall structure as scientific authority, it has a similar degree of intellectual coherence. After all, the intellectually incoherent practice of random citation would produce a Poisson distribution, not those we actually see. However, such a claim might be overreaching. It may be that the two power law distributions one sees are the result of deep sociological facts that affect any human enterprise that produces a citation network, whether or not the underlying
The fact that most judicial decisions, like most physics articles, die, for purposes of “making law,” immediately or soon after being published, is of substantial significance. For example, law professors frequently point out the tension between judges doing justice in the individual case and making the decision that will set the most desirable precedent. Yet this tension would seem much less than probably supposed if the probability is remote that any new case, especially a case decided by a lower court, will ever be cited as precedent, even less if the probability is remote that the decision will ever become a new, widely cited rule of law. In an economic model of judging, judges should presumably discount highly the disutility of a decision that would “set a bad precedent.” The normative implication presumably would be that judges should give more weight to doing justice in the individual case, than might now be commonly assumed, even if that result differed from setting the legally correct precedent. In the contracts law chestnut Williams v. Walker-Thomas Furniture Company, for example, there may be no problem with Mrs. Williams getting her money back on the stereo system she improvidently purchased. The court’s decision probably would not have affected the ability of poor people to get credit, a traditional worry of law-and-economics scholars considering this famous unconscionability case, if its likelihood of being cited in the future was low. The whole idea that every time a judge makes a decision, he tugs a little on the Web of Law, changing its shape, might seem demonstrably false. The fact that the majority of cases are not really important in terms of making law, would also seem to support the practice, common in many jurisdictions, of leaving many cases unpublished.

On the other hand, the story may be more complicated. It may be that many of the cases that die aborning do so because they simply apply the law as it already exists, instead of making new law. Decisions that depart from established doctrine, conversely, might have a better chance of becoming influential. A more sophisticated theory of why judges should be obliged not to make bad precedent, even if it would result in justice in the individual case, may be that while there is only a low intellectual exercise is as coherent as physics. On the other hand, it may be that citation networks organized in this way are a hallmark of intellectually rigorous disciplines.

74. Landes & Posner, supra note 4, at 250–51.
75. 350 F.2d 445 (D.C. Cir. 1965).
76. See id. at 449.
probability that the typical case would become influential, there is a greater probability that a case that departs from established precedent will become an important case, exercising its influence possibly far beyond its original facts and context. Understanding the network dynamics of the Web of Law might not lead us to dismiss our settled intuitions about how law works and judges should behave, so much as realizing that the justifications for these intuitions are much more complicated than we might have assumed.

III. CONSEQUENCES, APPLICATIONS, AND FUTURE RESEARCH

Lawyers, judges, and law professors have long resorted to metaphors of webs, trees, and bramble bushes to evoke the structure of law.78 These metaphors attempt to get at what we can now describe far more precisely as the network structure of law. The Web of Law shares a mathematical structure with many other evolving networks. These structures are being studied intensely by network scientists and are gradually yielding their secrets. The Web of Law has an overall shape and internal structure that can now be profitably studied in new ways. The network perspective holds great promise for deepening our understanding of legal systems and improving the technology we use to access the Web of Law.

The Web of Law grows as judges write opinions which cite cases and other authorities. Legal scholars also add to the network by writing articles and treatises which cite cases and other authorities, and which in turn are (sometimes) cited by cases and other authorities. As the Web of Law grows, a great deal of information gets embedded in citations. Judges cite the cases that they think are the most relevant to the case they are deciding. When two judges deciding different cases cite some of the same authorities, it is also a signal that those cases are relevant to each other. In this way, the millions of decisions regarding what to cite organize the Web of Law into what network scientists call clusters or communities. In other real networks, these clusters form not just structures in link topology, but structures in semantic topology as well. Thus the World Wide Web organizes itself semantically, that is, according to meaning or topic. There is little ground for doubt that the Web of Law does the same. Some search engines, such as Yahoo, impose a

structure from the outside on the World Wide Web. But newer search engines such as Google, which exploit the link-to-semantic topology congruence, have proven more powerful and more popular. Google works by free-riding, as it were, on the information embedded in the linkage decisions made by millions of “webmasters,” the people who create and manage web pages.\textsuperscript{79}

Analogously, our common law system and indeed any common law system will very likely have an organic organization that can be mapped, studied, and probably exploited in a similar way. That common law systems, including ours, spontaneously organize themselves into subject matter or topical clusters, can hardly fail to be of significance to anyone interested in how legal systems evolve and function. More practically, exploiting the information embedded in citation networks has great promise for facilitating lawyers’, judges’, and scholars’ access to cases and other authorities in the growing Web of Law that are the most relevant to their particular projects, just as Google did for users of the World Wide Web.

As discussed in Part II above, the network perspective should also affect the way we view the doctrines of precedent and stare decisis. It probably comes as a surprise to many lawyers, judges, and scholars that the majority of cases have little or no value as precedents, and that the majority of precedential influence is exercised by a relatively small minority of cases. The fact that this pattern emerges equally in the Supreme Court of Alabama and the Supreme Court of the United States indicates that this is a fundamental fact about legal systems organized as ours are. Even the few cases that become legal precedents may also have a natural life span, a term over which they grow in authority, attain a kind of maturity, and then decline into relative obscurity. Scientific authorities do this in scientific citation networks, and legal authorities probably behave in the same way.\textsuperscript{80} The idea that any given case has a low probability of becoming a precedent, and even those that do will be

\textsuperscript{79} I do not mean to claim all scale-free or similar networks organize themselves semantically, because the network might not only be scale-free but also semantics-free. That is, it might just be an artificial network where each node is a mathematical point with no meaning at all, such as airline transportation networks. Nodes in that network do not have any “meaning.” But in networks about which it makes sense to speak of content that can be semantically organized—such as publication citation networks, networks of words, and networks of Web pages—spontaneous semantic organization tends to occur. My claim is simply that the same thing occurs in the Web of Law.

\textsuperscript{80} At this stage, this is a conjecture on my part, based merely on the inspection of the distribution of the data and looking at other models in the literature. However, at this early stage of applying network theory to law, I think it is useful to make plausible conjectures that can be confirmed or falsified by further research.
so only for a time, is certainly different from the standard view commonly taught in law schools and probably taken for granted by most lawyers.

This Part briefly discusses these observations and several others regarding how network theory sheds light on law. This Part is divided into six subparts, each of which discusses some aspect of network theory as applied to law. In subpart A below, I ask whether the legal network is a “small world” and what the answer to that and related questions might tell us about the legal system. In subpart B, I discuss clusters in networks and how this concept applies to the Web of Law. In subpart C, I consider the significance of “hub cases.” In subpart D, I discuss node fitness and how it applies to the legal network. In subpart E, I consider how network theory can be applied to our concepts of legal authority and precedent. In subpart F, I explain how the network structure of the Web of Law might be exploited to improve computerized searching in legal networks. Finally, there is a brief conclusion.

A. Is the Web of Law a “Small World”?  

Measuring the diameter of a network, how “big” or “small” it is, tells us how integrated the network or parts of it are. A basic measure of the diameter of a network is the average number of links one has to cross in order to get from one randomly chosen node in the network to another. Measuring the diameter of the Web of Law, or parts of it, would give us a measure of how well integrated the Web of Law is, and what parts of it are more or less well integrated with the rest.  

In a complex legal system with multiple sovereigns, this would be useful. For example, network analysis could quantitatively measure how autonomous state legal systems are from one another and from the federal system. In law school, most lawyers get a rather standard account of federalism. Each state has its own court system, and the federal system floats above it. The picture gets more complicated as one considers federal courts deciding state law issues in diversity jurisdiction, and that some states, such as California among the Western states, Delaware in corporate law, and New York in insurance law, project legal influence beyond their borders. These complications, however, do not usually cause us to
revise our basic picture of federalism, with its fifty autonomous states and overarching federal system.

It would be interesting, however, to describe empirically how well the actual geography of our federal legal system corresponded to or departed from the law school model. Groups of states might form what are in effect sub-national, but supra-state, legal regimes, at least in certain areas of law. Similarly, one could measure the extent of federal courts’ influence on state courts, and vice versa. One could also measure how well integrated different parts of the federal judiciary were with the federal system as a whole. For example, is the Ninth Circuit really “a law unto itself,” less integrated into the federal judicial system than are other federal circuits?\(^\text{83}\) Sophisticated citation studies using network theory offer an empirical method to improve our understanding of the complex relations between judicial bodies and the legal regimes they create.

Network diameter can change over time, as the network becomes more or less integrated. Barabási et al. found, for example, that in two networks of scientific collaborators, the average separation between nodes was decreasing over time and approaching a limit.\(^\text{84}\) Does the Web of Law, or parts of it, behave in the same way? If the legal profession is becoming increasingly specialized, one might think different regions of the Web of Law would grow more separated from one another. The broader question of whether there are limits of integration or separation towards which common law systems evolve is also worth probing. The idea that integration may have limits is also relevant to federalism. The various state law systems, or the federal and state systems taken as a whole, may be tending toward a certain level of legal integration, but not more than that. If so, it may suggest that in federal systems there are limits to the degree of integration that would naturally occur. Alternatively, integration might be seen to be increasing with no limit. This would suggest that unifying forces in law were so strong that even in a federal system law will tend to integrate itself into one system. Another alternative is disintegration, in which courts, or different areas of law, or both, become increasingly isolated from one another with no limit in sight. Whatever the results of actual analyses reveal, they could be of profound importance to our understanding of legal systems.


B. Legal Clusters

Real scale-free and similar highly skewed networks tend to share topological traits. The Web of Law is such a network, and is thus likely to possess these mathematical properties as well. These properties are a potentially rich source of general facts about common law systems generally and ours in particular. A good example is that real scale-free networks tend to be organized in clusters. The Web of Law thus probably consists of clusters of cases which are relatively tightly linked within themselves, but more sparsely linked to each other, analogous to the structure of the World Wide Web. The preliminary visualization of U.S. Supreme Court cases in Figure 8 strongly suggests this is the case. Furthermore, these clusters probably correlate highly with underlying legal semantics. That is, cases in the same legal cluster are likely to be related to each other in terms of meaning and subject matter, as are communities of Web sites.85

Indeed, tightly linked legal cases ought to be, one might think, even more closely related semantically than similarly linked Web pages. Courts cite the most relevant cases they can. These citations are produced by persons who are intimately familiar with the case at hand and with the relevant law. Judges also have strong motivations to make relevant links, as citations are part of what persuades higher courts and other potential critics that their decisions are correct. The link topology to semantic topology congruence, therefore, would seem likely to be even tighter in the legal network than it is in the Web, where this phenomenon was first described.86

This implies an important conclusion for general jurisprudence. If link topology maps well onto semantic topology in the Web of Law, then analysis of clustering in the legal network should give us a rare, objective picture of the natural organization of law. Instead of some topology imposed from the outside, by legal scholars or judges with particular points of view, clustering analysis may reveal a mode of organization that is naturalistic, that is, an organization that is found in the legal system, rather than imposed upon it. This organization is likely

85. David Gibson et al., *Inferring Web Communities from Link Topology* (1998), http://citeseer.ist.psu.edu/ (type the article title into the central field, click “Search Documents,” and select the article title).

86. See id. at 2.
to yield insights about law, and it may or may not conform to conventional organizations of law.

C. Hub Cases

Scale-free and similar networks are as integrated as they are because of the relatively few nodes that have many links. In the Web of Law, a small minority of cases get the large majority of citations. As noted above, a mere one thousand cases, roughly 0.025% of the total, get about 80% of the total number of citations. One might say that American case law, state and federal, is really only about a thousand important cases. These may be thought of as the hubs of the Web of Law.

More research needs to be done to characterize these cases. Some tentative generalizations are possible, however. Powerful courts, such as the U.S. Supreme Court, the federal circuit courts, and the state supreme courts, probably decide most of the hub cases. Hub cases seem also to include many procedural decisions, perhaps because procedural jurisprudence applies to many different types of substantive controversy, and because law has a tendency to force a wide variety of factual situations into a relatively few procedural forms.87 Similarly, constitutional cases might prove disproportionately to be hubs, because the principles they articulate apply to many cases in otherwise diverse legal areas.

It is probably the case that particular judges are responsible for more than their share of hub cases.88 Future research should include studies of citation distributions by particular judges and courts. It seems highly likely that these distributions will resemble those of cases in the Web of Law. A similar conjecture may be made about the citation distribution of legal scholarship.89 It is probable that a relatively few law review articles garner a large majority of the total number of citations. Depressing as it may be for legal scholars, the converse also seems probable: The large majority of law review articles quickly and irreversibly become completely obscure or “dead,” and like the majority of physics articles, are never or rarely cited. The biological metaphor is more to insects than humans: many are born, but only a few survive.

Network analysis would also reveal which cases are important to scholars and which articles important to judges. The cases law professors cite may be quite different from those cited by judges, and the articles

87. Professor Farber observes that many of the most frequently cited cases in his data are procedural cases. See Farber, supra note 61, at 865, 870. I suspect many of the hub cases are procedural cases.

88. See Montgomery N. Kosma, supra note 4; Landes et al., supra note 4.

cited by judges may be different from those most cited by law professors. Thus, for example, I would conjecture that *Roe v. Wade* \(^90\) has been cited many more times by law professors than by judges. The legal scholarship network and the network of case law and statutory authority probably interpenetrate, but are probably also to some extent independent spheres.

One must be cautious in interpreting what it means that a particular case is a hub. The most frequently cited case in American case law, with about 72,000 citations to its credit, is the *Liberty Lobby* \(^91\) case, a pillar of the U.S. Supreme Court’s summary judgment jurisprudence. It is, without doubt, an important case. Presumably it is cited so often because the federal courts, and state courts following federal courts’ lead, handle many motions for summary judgment, and very often when they do, they cite *Liberty Lobby*. It does not follow, however, that it is therefore the “most important” or “most authoritative” case in American law. On the other hand, it would be hard to deny that *Liberty Lobby* provides law that is at the core of what federal courts do. Would a set of hub cases, defined in some appropriate way, do a good job of capturing the fundamental doctrines of American law? Perhaps the better question would be, do the so-called “fundamental doctrines” of American law do a good job of capturing the hub cases?

While it may be hazardous to conclude that one actively cited case is more important or authoritative than another, it certainly seems plausible to distinguish between cases that are very actively cited, modestly cited, rarely cited, and never cited. A citation is, after all, the invocation by a court of the authority of a previous decision by some court or other juridical body, such as an administrative agency. A decision that is utterly ignored cannot be said to be influential or authoritative. However, many factors go into making a case frequently cited. *Gregg v. Georgia* \(^92\) is a very frequently cited U.S. Supreme Court case, yet it may be so because it is invoked so frequently in relatively routine matters before the Court. Part of the problem here is that network analysis itself generates distinctions that are not captured by relatively crude concepts such as “importance” and “authoritativeness.” One case may generate a lot of routine law, and be important in one respect. Another case may generate a lot of scholarship, but not be cited by many judges. Yet

---

90. 410 U.S. 113 (1973).
another case may settle some important issue, such as the border between Texas and Oklahoma, but not be relevant to many other issues. Cases that are important in legal evolution, moreover, may not be the most important socially or economically. Finally, some cases may become so much a part of the fabric of the law that they are no longer frequently cited. Arguably, every time a federal court reviews legislation for constitutionality, it invokes the authority of *Marbury v. Madison*.93 Yet to cite *Marbury* in a routine case today reviewing a statute for constitutionality would seem somewhat pedantic or cranky.94 Frequency of citation does not capture the influence of important cases that have been absorbed into the fabric of law. However, it is also true that very few cases, as a percentage of the total, fall into this rarified category. For every *Marbury*, there were thousands of cases of which no one has heard, or ever will.

Given what we now know about citation patterns, an important jurisprudential question suggests itself. What is the relationship between the hub cases and what we call “American law”? It is tempting to speculate that nearly all of the rules and principles that we say are embodied in cases, in our common law system, are in fact embodied in a very small subset of the total set of cases. American common law appears to derive from a small kernel of cases. If a mere one thousand cases get 80% of all citations, and a mere 2% of U.S. Supreme Court cases get 96% of cites to Court cases, then it seems natural to conclude that it is in those elite cases that the law is to be found. The rest of the cases, those that cite, but are never or rarely cited, would seem to be merely applying the law found in the kernel.

D. Fitness in the Web of Law

Nodes in scale-free networks fit a power law curve, in terms of “in-degree.” In the log-log format, the nodes with the most links, the hub nodes, are those to the southeast, while the much more numerous nodes with few or no links, are to the northwest. Measuring the in-degree of a network, such as the Web of Law, is like taking a snap shot of it at a particular time. The nature of scale-free and similar real networks, however, is to evolve as new nodes and new links get added. Can we picture dynamically what this evolution of the Web of Law looks like? The following fable gives an idea of what the evolution of a network like the Web of Law, which has nodes of varying fitness, might look like.

93. 5 U.S. (1 Cranch) 137 (1803).
Let us imagine that when a court issues an opinion, it produces a copy of it, which weighs a few ounces. If that decision gets cited, a small helium balloon, like children get at parties, is attached to it. Balloons are our metaphor for links, which in the Web of Law are citations by other cases. Cases start out with no citations. As of mid-2004, about 400,000 cases, some old, and some new, were in this state of complete obscurity. In our fable, these 400,000 booklets are earthbound. Another 773,000 cases have only one citation. They float barely above the ground. Other cases have been cited more often. As cases get more citations, they get more balloons, and rise higher off the ground. The altitude of each case is determined by how many balloons are attached to it. This presents a curious spectacle. Many cases, about a tenth of the total, just sit on the ground. Another fifth are barely off the ground. Most of the rest are at different altitudes, but the vast majority are still close to the ground. A relatively few cases have many, hundreds, or thousands of balloons attached to them. They soar high above the rest. If we were to film this evolving spectacle, we would see that some cases worked their way to a high altitude by virtue of having been around longer than most others. These cases get balloons tied to them relatively slowly, but surely, and so rise in the ranks, some making it into the highest rank. Other cases would get a few balloons attached to them, but then stall fairly close to the ground. Still others would rise quite quickly, getting balloons attached to them at a rapid pace, and rise anywhere from somewhat to much faster than other cases that were making more stately upward progress. But most cases just stay close to the ground.

In this informal model, balloons get attached to cases partly as a matter of preferential attachment, that is, as a function of how many balloons a case already has attached to it. We can imagine that judges looking for cases to cite look up and see those that are high above the rest. Those attention-grabbing cases are most likely to get still more balloons attached to them. But this is not the only dynamic at work. Cases also get additional balloons attached to them depending on how “fit” they are. With the BA preferential attachment model, we can calculate the probability that a given case will get an additional citation when new links are added to the network. We can also determine which cases are getting new citations at a rate faster than that predicted by the preferential attachment model. How much greater this rate is, is a measure of the fitness of the case. Fitness is that portion of a node’s ability to garner links that is not explained by its already having a certain
number of links. When a case garners citations beyond what the preferential attachment model would predict, we say the case is relatively fit. How fit is a matter of how many links it gets beyond what preferential attachment predicts.  

Measuring fitness of cases (and other authorities) in the Web of Law would be useful. It should be possible to determine whether authorities were frequently cited because they were so well established or because they were particularly fit. Some cases might be cited more than others because they are better reasoned, more persuasive, more accurately interpret statutory language or legislative intent, or better serve political goals. Measuring fitness should also allow us to identify cases that were not yet authoritative, but were heading in that direction, and conversely, cases that were currently authoritative, but seemed to be in the process of losing their authority. The fitness of other sources, such as law review articles, also could be measured. Law scholars would probably like to know how fit their scholarship was, compared to that of other scholars. One can imagine a search engine that provided the option of searching for “emerging authorities” by calculating a fitness index for cases within given relevance parameters (produced by a text-based search, for example) and highlighting cases of high fitness, or ranking them in fitness order.

E. Improving Legal Research Technology with Network Science

If the Web of Law has a structure similar to the World Wide Web, it follows that searching for relevant authorities in the legal network could be improved by drawing on techniques that work well on the World Wide Web. Searching in networks is now a highly developed and

95. This process is similar to measuring fitness in evolutionary biology. See, e.g., Post & Eisen, supra note 4, at 569 n.37 (citing examples).

96. In recent years, law and economics scholars have used statistical methods to evaluate judicial performance. See, e.g., Stephen Choi & Mitu Gulati, A Tournament of Judges?, 92 CAL. L. REV. 299, 305–09 (2004). Could measures of the fitness of citations improve these methods? Probably, but it would be a complex undertaking. Measuring citation fitness would correct somewhat for bias in favor of older, more famous judges. Some scholars have attempted to do this respecting legal scholarship. See Fred R. Shapiro, The Most-Cited Legal Scholars, 29 J. LEGAL STUD. 409 (2000); see also Landes & Posner, supra note 89. It would not correct for increased citations judges get for having dockets in more active areas. To really measure fitness, one should presumably measure it relative to other cases in similar areas. Some jurisdictions are also much busier than others. Unless credit is to be giving for working harder (and perhaps it should be) citation rates should be measured relative to the level of activity in that particular area, both jurisdictionally and by subject matter. Finding the average fitness of cases may correct for the advantage busier judges have, but it would penalize judges who must shoulder a lot of routine work, but still manage to produce heavily cited opinions.
technical field. By offering a superior search algorithm, Google has become a multi-billion dollar public company. Yet, the search technology that works well on the World Wide Web has not yet been applied, at least not on any large scale, to legal networks.

Google exploits the information imbedded in web links to rank the relevance of search results. Google uses its patented PageRank algorithm to rank search results produced by a more conventional text-based system. Roughly speaking, the Google search engine first produces a set of results based on the occurrence of terms on a web page. (The most used legal search tools work similarly.) Then PageRank scores those web pages for relevance on the basis of how many web pages link to them. This score is calculated, however, so that being linked to by a web page that itself has many links increases a web page’s relevance score by more than being linked to by a webpage with only a few links. So if a web page is linked to by Yahoo or CNN.com, it will have a higher relevance score than a page linked to by, say, some obscure blog.

In theory, something similar could be done with legal network searching. Legal search results could be ranked by citation frequency, and in fact, one small company, fastcase.com, does just this. The key to Google’s powerful results, however, comes not from counting all links equally, but from scoring them according to how authoritative (or heavily linked) the linking site is. In the legal network, a similar approach could be implemented in various ways. The most straightforward would be to score citations depending on how many cites the citing cases themselves had. (The process is recursive, but in practice, going back a few levels in a citation network is usually all that is necessary to achieve greatly improved results.) Because law is hierarchical, scoring algorithms based on the level of the courts citing a case might also prove powerful. In this approach, a federal circuit court case that was cited approvingly by the U.S. Supreme Court would be scored more highly for relevance than would a similar case that was cited approvingly by another federal circuit court. The dimension of time is also different in the World Wide

Web than in the Web of Law, which is a citation network, not an electronic network. It seems likely legal search algorithms would be more powerful if relevance scores based on links were discounted for the freshness of links. So, a case that was cited by a federal circuit court five times in the last two years might rank more highly, all other things equal, than a case cited ten times by the same court, but ten years ago.

A search engine that ranked cases (or other authorities) in order of fitness would also be useful in some applications. Legal scholars and legal practitioners, for example, often want to know what are the cases, articles, or other authorities that are attracting the most attention lately. Recent cases may not have had time to accumulate many citations, and yet might still be attracting a relatively large number of citations for their age. These “emerging authorities” could be identified by calculating the fitness of cases returned in Boolean or other text-based search. Doing this would require calculating an expected citation weight for cases of a particular type and age, using a preferential attachment model, and then comparing that weight to the weight of cases actually produced by the text-based search. Fitness searches of the area with which a lawyer or scholar wanted to remain abreast would show what were the “hottest” of relatively recent cases and articles.

Another feature of real networks that can be exploited by search algorithms is clustering. It seems likely that the next generation of web search technology will exploit the natural tendency of real networks to cluster. Cluster analysis may well be applicable to the Web of Law as well. To suggest how this would work, consider how one performs a thorough job of legal research prior to the existence of the Internet. Using an index based system such as West’s, one finds one or more cases that are as close to on point as possible. One reads those cases, and then the cases they cite. If one is fortunate, this leads to a case or cases that are directly on point or at least close to the issue one is researching. Then one reads outward, as it were, to the cases cited by those cases, and the cases cited by those cases, and so on. One establishes the boundaries of the little universe of cases that bear on one’s issue. When one does this sort of research, one is exploring the Web of Law from the inside, like someone mapping a maze by walking around inside it.

Cluster analysis can be thought of as looking down at the maze from above, and using algorithms to identify the clumps of relatively tightly connected cases. If one began with one or two cases that one knew were relevant, the algorithm would give you the cluster that those cases were part of. An algorithm of this sort could save a great deal of time for lawyers or scholars who wanted to quickly ascertain the relevant set of cases for a given issue.
IV. CONCLUSION

American case law and other legal authorities are organized in a certain way, as a web or network, with its nodes connected by the links of citations. This network can be considered as a mathematical object whose topology can be analyzed using the tools pioneered by network scientists who wanted to explore the structure of the World Wide Web and other real networks. The Web of Law has a structure very similar to that of other real networks, such as the Web and the network of scientific papers. The Web of Law is a scale-free network, or something quite similar. Analogously to the Web, it has a relatively few hub cases that have many citations while the vast majority of cases have very few. The distribution of citation frequency approximates a power law distribution, as is common with real scale-free networks, and resembles closely the citation distribution in the network of physics papers. The most immediately striking feature of this distribution is the extraordinary concentration of precedential authority in a relatively few cases. This pattern of highly skewed distribution is evident at every level of our legal system. It is a fundamental feature of our common law system.

Many promising hypotheses can be generated by considering the law as a scale-free network. State and federal systems can be examined empirically to measure how well integrated each is with itself, and with each other, and how this is changing over time. Legal authorities can be measured to determine whether their authority is emerging or declining. Institutional bodies, such as courts, can be examined in the same way. Clusters of cases, which will reveal the semantic topology of law, can be mapped to determine whether traditional legal categories are accurate. These methods can be operationalized in computer programs to improve the efficiency of searching electronic legal databases. The dynamics of authority in law generally can be studied much more rigorously. How nodes age may profoundly affect overall network structure and therefore affect the shape of the Web of Law. Network theory hints at complex but analyzable interactions between the legal doctrines of precedent and the systems of federalism and common law.

Because law grows and because it has doctrines of authority, it creates a network of a certain shape. Our legal system has features that cause it to spontaneously organize itself. This is the product of laws arising from the underlying mathematics of networks, laws that govern networks of computers, proteins, and firms as inexorably as they govern networks of cases.
Understandably enough, network scientists have been far more interested in the network of scientific publications than they have been in the Web of Law. Part of this Article’s purpose is to advocate collaboration between legal scholars and network scientists to explore what may be the oldest, largest, and best documented citation network ever created. The Web of Law is probably the largest citation network in existence, and stretches back some two centuries. Legal databases are huge, well documented, and readily accessible. They present a perfect opportunity for the application of network science. This research would produce new knowledge of general jurisprudence that has simply been impossible until now, when we have the necessary advances in network science, fast computers, and a complete record of the legal network in electronic form, waiting to be explored.