LEGAL NETWORKS: THE PROMISES AND CHALLENGES OF LEGAL NETWORK ANALYSIS

Ryan Whalen[*]

2016 MICH. ST. L. REV. 539

INTRODUCTION

The adoption of a "network perspective" has been one of the most influential recent developments in the biological, physical, and social sciences. Legal scholars too have begun to take a network approach, posing legal questions as network problems and exploring legal data from a network perspective. This Article reviews the existing literature on legal network analysis, suggests directions for

---

* Law & Science Fellow Northwestern University Pritzker School of Law, Lecturer Dalhousie University School of Information Management. JD – Northwestern Law, PhD – Northwestern University.

future work, and notes some important caveats for scholars hoping to adopt network analytic methods.

The first Part of this Article provides a brief network analysis background. It first describes the traditional data forms that networks take—nodes or vertices joined by links or edges—and how those structures can contain different types of information, such as weighted links. This background section subsequently goes on to discuss the history of network analysis, introducing some early work on networks and describing its increasing importance in recent years as network data and awareness increase, and computational power enables more and more sophisticated analysis methods.

The next Part briefly touches on the different ways that scholars have applied network analytic techniques to questions of interest to legal scholars. Legal networks take many forms, including precedent and patent citation networks, statutory and regulatory networks, and social and organizational networks.

Subsequently, this Article presents some methodological considerations for future legal network analyses. I argue that scholars should strive to retain more detailed and nuanced data, to analyze those data in more sophisticated ways, and to embrace the explanatory power of network analyses by drawing on multiple levels of theory about their formation and structure. Finally, I take my own advice and provide a brief empirical example of how more nuanced network analytic techniques can provide insight that more simple methods do not.

I. NETWORK ANALYSIS BACKGROUND

While the world has grown increasingly complex and interrelated, network analysis has become an important tool for twenty-first century science.[1] As scholars have begun to apply network methods to reveal insights about complex social, economic, and technological systems, a new network-oriented field of study has begun to emerge.[2] In recent years, legal scholars have begun to take notice of this network approach and apply network analytic techniques to questions about the law, legal systems, and the social systems that the law seeks to regulate.[3] However, despite some

---

1. *See* DAVID EASLEY & JON KLEINBERG, NETWORKS, CROWDS, AND MARKETS: REASONING ABOUT A HIGHLY CONNECTED WORLD, at xi (2010).

2. *Id.*

3. *See, e.g.*, James H. Fowler et al., *Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court*, 15 POL.

studies applying network methods to legal data, there are still comparatively few legal network studies. This is likely the result of a number of factors, including the traditional training model for legal academics, and the publishing model for legal research. This Section will briefly introduce networks and explore the history of network analysis in a variety of fields. By demonstrating the diverse ways in which scholars in other fields have developed network techniques and applied them to scientific and social scientific questions, I hope this overview will help demonstrate the promising utility that network techniques have for the study of the law.

A. What is a Network?

Network analysis is suited to many different substantive domains and forms of data. Anywhere there are relationships, there is also the potential to examine those relations as a network. Nodes and links are the basic constituent parts of any network, where relationships between nodes are represented with links.[4] These nodes can represent many types of data, including people,[5] organizations,[6]

ANALYSIS 324 (2007); James H. Fowler & Sangick Jeon, *The Authority of Supreme Court Precedent*, 30 SOC. NETWORKS 16 (2008); Daniel Martin Katz & M. J. Bommarito II, *Measuring the Complexity of the Law: The United States Code*, 22 ARTIFICIAL INTELLIGENCE & L. 337 (2014); Daniel Martin Katz et al., *Reproduction of Hierarchy? A Social Network Analysis of the American Law Professoriate*, 61 J. LEGAL EDUC. 76 (2009); Thomas A. Smith, *The Web of Law*, 44 SAN DIEGO L. REV. 309 (2007); Brian Uzzi & Ryon Lancaster, *Embeddedness and Price Formation in the Corporate Law Market*, 69 AM. SOC. REV. 319 (2004); *How Network Science Is Changing Our Understanding of Law*, MIT TECH. REV. (Jan. 29, 2015), http://www.technologyreview.com/view/534576/how-network-science-is-changing-our-understanding-of-law/ [https://perma.cc/U54V-EE5J]; Ryan Whalen, *Modeling Annual Supreme Court Influence: The Role of Citation Practices and Judicial Tenure in Determining Precedent Network Growth*, *in* COMPLEX NETWORKS 169 (Ronaldo Menezes, Alexandre Evsukoff & Marta C. Gonzalez eds., 2013) [hereinafter Whalen, *Modeling Annual Supreme Court Influence: The Role of Citation Practices and Judicial Tenure in Determining Precedent Network Growth*]; Ryan Whalen, *Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Supreme Court Precedent*, *in* NETWORK ANALYSIS IN LAW 1 (Radboud Winkels, Nicola Lettieri & Sebastiano Faro eds., 2014) [hereinafter Whalen, *Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Supreme Court Precedent*].

4. MARK NEWMAN, NETWORKS: AN INTRODUCTION 109 (2010).

5. *See, e.g.*, Gueorgi Kossinets & Duncan J. Watts, *Empirical Analysis of an Evolving Social Network*, 311 SCIENCE 88 (2006).

6. *See, e.g.*, Herminia Ibarra, Martin Kilduff & Wenpin Tsai, *Zooming In and Out: Connecting Individuals and Collectivities at the Frontiers of Organizational Network Research*, 16 ORG. SCI. 359 (2005).

or judicial opinions.[7] The links between nodes can represent any form of relationship, whether it be a friendship relationship between people,[8] the fact that two corporations share board members,[9] or a citation from one judicial opinion to another.[10]

There are a variety of types of networks, each of which has potential applications in legal network studies. One of the most basic distinctions between types of networks is whether the relationships (or links) are directed or undirected.[11] A directed network represents relationships that are not necessarily mutual. For example, most communication network studies represent the data in a directed format because messages can be both sent and received. I may send you a message (suggesting a tie from me to you), but you may never send me a message in return. In legal network studies, these directed networks come up most frequently in analyses of legal citation networks. Citation networks have an additional constraint in that they are generally acyclic. This arises because of the temporal nature of citations. Citations only go "backwards" in time, so there are no "loops" in the network structure.[12]

Undirected networks are used to represent relationships that are mutual. In these networks, when two nodes are connected to one another the relationship between them does not exist "from" one node and "to" another, but simply exists between the two actors. For example, acquaintance networks are often considered undirected. If two individuals know one another they are said to have an acquaintance tie. There is no directionality in this relationship, so this data is represented in an undirected form.

In addition to directionality, networks can be used to represent multiple types of nodes and multiple types of relationships. When there are two types of nodes in a network it is referred to as a bipartite network (or bipartite graph).[13] These are used frequently for affiliation networks, where individuals (one type of node) are

---

7. *See, e.g.*, Fowler & Jeon, *supra* note 3.

8. *See, e.g.*, Evelien Zeggelink, *Evolving Friendship Networks: An Individual-Oriented Approach Implementing Similarity*, 17 SOC. NETWORKS 83 (1995).

9. *See, e.g.*, John A. Sonquist & Thomas Koenig, *Interlocking Directorates in the Top U.S. Corporations: A Graph Theory Approach*, 5 INSURGENT SOCIOLOGIST 196 (1975).

10. *See, e.g.*, Stephen M. Marx, *Citation Networks in the Law*, 10 JURIMETRICS J. 121 (1970).

11. NEWMAN, *supra* note 4, at 114-15.

12. *Id.* at 118.

13. *Id.* at 123-27.

affiliated with another type of node—for instance, shoppers linked to the products they purchase on an online shopping site.[14] Bipartite networks can be converted to unipartite projections, in which nodes will be closely related to one another if they share connections to many of the same nodes in the other mode of the affiliation network (e.g., if they purchased many of the same products).[15]

When they represent more than one type of relationship, networks are said to be "multiplex."[16] For instance, instead of simply graphing acquaintances between individuals, I might be interested in tracking a variety of relationships between lawyers. A lawyer social network might contain links that represent having worked in the same office, having graduated from the same law school, having clerked for the same judge, having advocated on the same case, etc. This sort of multiplexity retains more of the information about the underlying relationships of interest, but also complicates the analysis.

One of the major strengths of a network analytic approach is that it allows scholars to approach their content from a very powerful multilevel perspective.[17] While traditional approaches allow for relatively small-scale, yet in-depth, analyses—think traditional doctrinal analysis of a subset of cases, showing the evolution of the law in a given area—network approaches allow a system-level perspective that can demonstrate both global properties and further our understanding of constituent parts by demonstrating how components of the network are related to one another.

B. Network Analysis History

The birth of network science can be traced back to 1736, when Leonhard Euler presented the first mathematical proof utilizing

---

14. *See, e.g.*, Zan Huang, Daniel D. Zeng & Hsinchun Chen, *Analyzing Consumer-Product Graphs: Empirical Findings and Applications in Recommender Systems*, 53 MGMT. SCI. 1146 (2007).

15. NEWMAN, *supra* note 4, at 124.

16. *See, e.g.*, Michael Szell, Renaud Lambiotte & Stefan Thurner, *Multirelational Organization of Large-Scale Social Networks in an Online World*, 107 PROC. NAT'L ACAD. SCI. 13636 (2010) (examining a multiplex network containing friendship and enmity ties).

17. *See* Noshir S. Contractor, Stanley Wasserman & Katherine Faust, *Testing Multitheoretical, Multilevel Hypotheses About Organizational Networks: An Analytic Framework and Empirical Example*, 31 ACAD. MGMT. REV. 681, 681 (2006) (demonstrating the application of a multilevel multitheoretic model of network analysis).

graph theory.[18] Euler's proof solved the Seven Bridges of Königsberg problem, in which individuals were challenged to prove whether or not it was possible to navigate the city of Königsberg's seven bridges joining two sides of a river and two mid-river islands in such a manner that each bridge was crossed once and only once.[19] By abstracting the problem from its geographic constraints and instead representing the landmasses as nodes and bridges as links in a network, Euler demonstrated that in order to cross each bridge once and only once, every node but two must have an even number of links (entry and exit points in this case).[20] As none of Königsberg's landmasses had an even number of connections, Euler proved that the challenge could not be successfully completed.[21] This represents one of first graph theoretic proofs and marks the beginning of network science.[22]

1. *Social Network Analysis*

While graph theory remained a relatively obscure area of mathematics, by the twentieth century network approaches to social scientific questions became much more popular. In the 1930s, Jacob Moreno began using what he called "sociograms" to help explain social behaviors.[23] These diagrams graphed human relationships to give the analyst some sense of their underlying structure.[24] A product of the Vienna school of psychoanalysts, Moreno's major contribution was in realizing that in order to understand humans and their behavior it was necessary not only to focus on the inner workings of their minds, but also the structure of relationships that they are embedded within.[25]

---

18. Throughout this Article I will interchangeably refer to networks as both "graphs," the term more commonly used in mathematics, and "networks," the term more commonly used in the social sciences.

19. NEWMAN, *supra* note 4, at 140-41.

20. *Id.*

21. *Id.*

22. *Id.*

23. *See* JACOB L. MORENO, WHO SHALL SURVIVE?: FOUNDATIONS OF SOCIOMETRY, GROUP PSYCHOTHERAPY, AND SOCIODRAMA 95-96 (3d ed. 1978).

24. *Id.*

25. *Id.* at 5-6.

Following Moreno's contributions, Milgram's small world experiment remains one of the more famous social network studies.[26] Milgram's work aimed to better understand the structure of relationships in the United States, and whether or not the common folk wisdom that "it is a small world" was empirically supported.[27] In order to do so, he selected individuals in Wichita, Kansas and Omaha, Nebraska and asked them to forward letters toward a target person located in Boston.[28] However, they were only to forward the letters to individuals that they know, those individuals would then in turn forward the letters on to individuals they know.[29] At each step the intent was to get the message "closer" to its ultimate target.[30] Upon completing the experiment Milgram revealed that there were on average approximately six degrees of separation between each two, effectively random, social actors.[31] These results were confirmed in subsequent research[32] and still hold in today's large online social networks.[33]

In the 1970s through 1980s, social network approaches again enjoyed substantial attention as scholars began to focus not just on describing the structure of social relations, but also on developing network-based theories to explain social phenomena. Much of the focus during this era of social network research was on the relationship between social structure and information diffusion. Mark Granovetter famously demonstrated the "strength of weak ties" in exposing individuals to fresh information that was not previously available to them.[34] In a similar vein, Ron Burt's work focused on demonstrating the importance of network position, showing that

---

26. *See* Stanley Milgram, *The Small-World Problem*, 1 PSYCHOL. TODAY 61 (1967); Jeffrey Travers & Stanley Milgram, *An Experimental Study of the Small World Problem*, 32 SOCIOMETRY 425 (1969).

27. Milgram, *supra* note 26, at 62.

28. *Id.* at 64.

29. *Id.*

30. *Id.*

31. *Id.* at 65.

32. Travers & Milgram, *supra* note 26, at 431-32.

33. Johan Ugander et al., The Anatomy of the Facebook Social Graph 1-2 (Nov. 18, 2011) (unpublished manuscript), http://arxiv.org/pdf/1111.4503v1.pdf [https://perma.cc/JDH6-7X6C].

34. Mark S. Granovetter, *The Strength of Weak Ties*, 78 AM. J. SOC. 1360, 1360 (1973).

occupying "structural holes" within the network leads to distinct information advantages.[35]

2. *Large-Scale Network Analysis*

The end of the twentieth century brought a number of developments that proved to be important for network science. Increased computational power, the popularization of the Internet, and the discovery of a number of important network models all contributed to network science's increasing capabilities and popularity. Increasing computational power has been essential to the development of network science. Because of the memory and computational resources required by network analysis, those done before the 1990s were limited in the scope they could encompass and in the types of analyses they could apply. This began to change in the 1980s and especially by the late 1990s and the 2000s, when more powerful and more affordable computers became widely available.

Along with these more powerful computers came the popularization of the Internet. The Internet contributed to the increasing popularity of network science in a number of ways. By networking individuals and ideas together, the Internet led people to view the world from a network perspective.[36] This created an atmosphere in which scientists, students, grant funding bodies, and the population, more generally, were receptive to thinking about networks.[37] In addition to the increasing popularity of a network perspective, the Internet led to the development of a number of important network methods and metrics. Perhaps most famously, the competitive advantage that led to Google's dominance in the search domain is based upon network theory.[38]

As the Internet increased network awareness and provided incentive for the development of new network measures, it also provided data with which social scientists could begin to study very

---

35. Ronald S. Burt, *Structural Holes and Good Ideas*, 110 AM. J. SOC. 349, 349 (2004).

36. *See, e.g.*, YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM 1-2 (2006).

37. *Id.*

38. *See* LAWRENCE PAGE ET AL., THE PAGERANK CITATION RANKING: BRINGING ORDER TO THE WEB (1998), http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf [https://perma.cc/4UGW-E8MW] (detailing the PageRank algorithm that bases a webpage's importance as a function of its position within the hyperlink network).

large social networks. Much of the early large-scale network analysis examined the Internet's hyperlink network as scientists attempted to describe and understand large-scale networks.[39] These and related studies led to a variety of network models that described the formation of networks that reflected those observed in reality. Mathematical network models, such as the preferential attachment model,[40] and the small world model,[41] helped network scientists understand the processes that underlie network formation and also eased network simulation and thus the computational study of networks.

C. Legal Network Analysis

While legal scholars have not adopted network analytic methods to the same extent as those in other fields of study, there have been a variety of legal network studies, and the trend appears to be increasing. This Section will briefly review the legal scholarship applying network analyses before the following Part explores more fully the future potential for legal network studies and the challenges faced by legal scholars working in the field.

Researchers have used network analytic techniques in a variety of contexts relevant to legal scholars. These include the analysis of legal social networks,[42] examining statutes and regulatory codes as networks,[43] studying the networks of criminals and terrorists,[44] and studying the structure created by case law citations.[45] As case law citation network analyses are the most common and perhaps the most accessible of these, we will begin the review by looking at the history of this body of scholarship.

---

39. *See, e.g.*, Réka Albert, Hawoong Jeong & Albert-László Barabási, *Internet: Diameter of the World-Wide Web*, 401 NATURE 130 (1999).

40. *See* Albert-László Barabási & Réka Albert, *Emergence of Scaling in Random Networks*, 286 SCIENCE 509, 511 (1999).

41. *See* Duncan J. Watts & Steven H. Strogatz, *Collective Dynamics of 'Small-World' Networks*, 393 NATURE 440, 440 (1998).

42. *See, e.g.*, JOHN P. HEINZ & EDWARD O. LAUMANN, CHICAGO LAWYERS: THE SOCIAL STRUCTURE OF THE BAR 7 (1982).

43. *See, e.g.*, Romain Boulet, Pierre Mazzega & Danièle Bourcier, *A Network Approach to the French System of Legal Codes—Part I: Analysis of a Dense Network*, 19 ARTIFICIAL INTELLIGENCE & L. 333 (2011).

44. *See, e.g.*, Jialun Qin et al., *Analyzing Terrorist Networks: A Case Study of the Global Salafi Jihad Network*, *in* 3495 LECTURE NOTES IN COMPUTER SCIENCE: INTELLIGENCE & SECURITY INFORMATICS 287 (Paul Kantor et al. eds., 2005).

45. *See, e.g.*, Fowler & Jeon, *supra* note 3.

1. *Legal Citation Network Studies*

Legal citation studies can be traced back to at least the 1950s.[46] Early work in this vein paid limited attention to the network structure created by case law citations, but focused instead on how citations accrued over time, and how judges decided which cases to cite.[47] John Henry Merryman wrote one of the earliest empirical analyses of judicial citation behavior when he examined the precedent citation behavior of the California Supreme Court.[48] He followed up on this with a 1977 piece that added a longitudinal perspective on his citation analysis.[49] By the early 1970s some were calling for more focus on judicial citations and emphasizing how they can provide important information that can assist in the legal information retrieval process.[50]

This body of work was extended by Landes and Posner as they turned their focus to an empirical examination of judicial citations and attempted to theorize citation behavior.[51] They argued that "to the extent that judicial citation practices exhibit regularities explicable within a systematic analytical framework, a statistical analysis of precedent should reveal them."[52] Their law and economics inspired analysis conceives of judicial opinions as "legal capital" and examines the rate at which precedent related to different legal areas ages. Posner and Landes liken this aging to capital depreciation, with different depreciation rates for different types of precedent.[53]

Posner subsequently called for more citation studies in the law, and pointed to the increasingly sophisticated work being done in bibliometrics as a potential inspiration for legal scholars.[54] This work built upon earlier judicial citation work, noting some challenges

---

46. *See, e.g.*, John Henry Merryman, *The Authority of Authority: What the California Supreme Court Cited in 1950*, 6 STAN. L. REV. 613 (1954).

47. *See, e.g.*, *id.* at 614.

48. *See id.*

49. *See generally* John Henry Merryman, *Toward a Theory of Citations: An Empirical Study of the Citation Practice of the California Supreme Court in 1950, 1960, and 1970*, 50 S. CAL. L. REV. 381 (1977).

50. *See, e.g.*, Marx, *supra* note 10.

51. *See* William M. Landes & Richard A. Posner, *Legal Precedent: A Theoretical and Empirical Analysis*, 19 J.L. & ECON. 249, 251 (1976).

52. *Id.* at 252.

53. *Id.* at 267-69.

54. Richard A. Posner, *An Economic Analysis of the Use of Citations in the Law*, 2 AM. L. & ECON. REV. 381, 383 n.3 (2000).

faced by the field. In particular, the issue of heterogeneity in citation motivation makes many accepted citation analysis methods challenging in legal contexts.[55] The varied meanings that legal citations have—e.g., to cite authority, to question a previous rulings reasoning, to cite for procedural reasons, etc.—complicate citation analyses. This early body of legal citation analysis focused more on counting citations, and little on the structural aspects of judicial citation networks.

It took until the twenty-first century, with increased access to data and computational capacity, for larger scale legal citation network analyses to begin to appear. The most well-known of these studies was published in 2007–2008 as two articles, one in the journal *Social Networks*[56] and one in *Political Analysis*.[57] These studies focused on the United States Supreme Court precedent network, exploring ways to measure precedent centrality. Fowler et al. applied the "Hubs and Authorities" algorithm to identify particularly important Supreme Court cases, and demonstrated that network analysis provided a meaningful way to identify important cases.[58] This work has been extended to some non-American contexts, for instance as when Winkels et al. applied a similar analysis to cases from the Dutch Supreme Court.[59]

As Fowler et al. approached legal citation analyses from a social scientific or bibliometric standpoint, legal scholars too began paying more attention to the potential that precedent citation networks had to shed light on the legal system. Analyzing the degree distribution of all federal and state cases in the Lexis-Nexis database, Smith demonstrated that judicial opinions follow the familiar scale-free distribution seen in many other network contexts.[60] In previous work, I have extended research done describing the Supreme Court citation network by taking a complex systems approach and examining citation network development,[61] and by focusing in on

55. *Id.* at 387.

56. Fowler & Jeon, *supra* note 3.

57. Fowler et al., *supra* note 3.

58. Fowler & Jeon, *supra* note 3, at 20; Fowler et al., *supra* note 3, at 343.

59. *See, e.g.*, Radboud Winkels, Jelle de Ruyter & Henryk Kroese, *Determining Authority of Dutch Case Law*, 235 LEGAL KNOWLEDGE & INFO. SYS. 103 (2011).

60. Smith, *supra* note 3, at 327.

61. Whalen, *Modeling Annual Supreme Court Influence: The Role of Citation Practices and Judicial Tenure in Determining Precedent Network Growth*, *supra* note 3.

distinctions between different types of cases.[62] Others have argued that judicial citations should be used as a performance metric to evaluate the work that judges do.[63]

Along with the large body of work studying precedent citation networks, there is a thriving body of research focusing on patent citation networks. This body of work can be traced back to the late 1940s when the United States Patent and Trademark Office (USPTO) began receiving requests to begin tracking prior art citations between patents.[64] Scholars hoped that these citations would assist in patent searching by "furnish[ing] a network of paths which cut across the major highways marked by the (inevitably artificial) boundaries of Classification."[65] This plan of course came to fruition, and now patent applicants are legally required to cite relevant prior art that they are aware of.

Patent citations have been used as a proxy measure for an invention's value,[66] as measures of knowledge flow,[67] and to better

---

62. Whalen, *Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Supreme Court Precedent*, *supra* note 3 (demonstrating that cases which go on to be overruled have different levels of centrality over time than similar matched pairs).

63. Peter A. Hook, *Evaluating the Work of Judges*, *in* BEYOND BIBLIOMETRICS: HARNESSING MULTIDIMENSIONAL INDICATORS OF SCHOLARLY IMPACT 345 (Blaise Cronin & Cassidy R. Sugimoto eds., 2014).

64. Arthur H. Seidel, *Citation System for Patent Office*, 31 J. PAT. OFF. SOC'Y 554, 554 (1949); Harry C. Hart, *Re: "Citation System for Patent Office,"* 31 J. PAT. OFF. SOC'Y 714, 714 (1949).

65. Hart, *supra* note 64, at 714.

66. *See, e.g.*, M.B. Albert et al., *Direct Validation of Citation Counts as Indicators of Industrially Important Patents*, 20 RES. POL'Y 251 (1991); Mark P. Carpenter, Francis Narin & Patricia Woolf, *Citation Rates to Technologically Important Patents*, 3 WORLD PAT. INFO. 160 (1981); Dietmar Harhoff et al., *Citation Frequency and the Value of Patented Inventions*, 81 REV. ECON. & STAT. 511 (1999); Mark Schankerman & Ariel Pakes, *Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period*, 96 ECON. J. 1052 (1986); Manuel Trajtenberg, *A Penny for Your Quotes: Patent Citations and the Value of Innovations*, 21 RAND J. ECON. 172 (1990).

67. *See, e.g.*, Paul Almeida & Bruce Kogut, *Localization of Knowledge and the Mobility of Engineers in Regional Networks*, 45 MGMT. SCI. 905 (1999); Benjamin Gomes-Casseres, John Hagedoorn & Adam B. Jaffe, *Do Alliances Promote Knowledge Flows?*, 80 J. FIN. ECON. 5 (2006); Adam B. Jaffe, Manuel Trajtenberg & Rebecca Henderson, *Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations*, 108 Q.J. ECON. 577 (1993); Carlos Rosell & Ajay Agrawal, *Have University Knowledge Flows Narrowed?: Evidence from Patent Data*, 38 RES. POL'Y 1 (2009).

understand the creative process.[68] By demonstrating how inventions are related to their predecessors, patent citation analysis allows us to analyze the products of the innovation system as a network of related knowledge.

2. *Statutory and Regulatory Network Studies*

Statutory and regulatory network analyses are conceptually similar to citation network analyses. Many statutes and regulations contain references to other sections or parts of the code. These references are in some ways similar to a citation. At times, they draw on information contained in the cited section (e.g., as when linking to a definition), and at other times they point the reader to another section of the code, relevant in some way to the origin section. Mazzega et al. demonstrate a network approach to statutory citation analysis by analyzing the majority of the French legal codes and showing that there is a highly skewed citation distribution, with a "rich club" of nodes (codes in this case) both sending and receiving the majority of the citations.[69] This work was later extended to show that this rich club consists of codes largely related to important social issues such as the regulation of territories and natural resources.[70]

Bommarito and Katz applied a similar analysis to the United States Code.[71] In addition to showing connections between different components of the Code, Bommarito and Katz apply their analysis over time showing that as revisions are made, and sections are added or removed, the Code grows in structure, interdependence, and language.[72] They follow this work up by focusing on complexity measures, comparing the complexity of various sections of the U.S.

---

68. *See, e.g.*, Kristina Dahlin, Margaret Taylor & Mark Fichman, *Today's Edisons or Weekend Hobbyists: Technical Merit and Success of Inventions by Independent Inventors*, 33 RES. POL'Y 1167 (2004); Lee Fleming & Olav Sorenson, *Technology as a Complex Adaptive System: Evidence from Patent Data*, 30 RES. POL'Y 1019 (2001); Lee Fleming & Olav Sorenson, *Science as a Map in Technological Search*, 25 STRATEGIC MGMT. J. 909 (2004) [hereinafter Fleming & Sorenson, *Science as a Map in Technological Search*].

69. *See* Pierre Mazzega, Danièle Bourcier & Romain Boulet, *The Network of French Legal Codes*, 12TH INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L., 2009, at 236.

70. Boulet, Mazzega & Bourcier, *supra* note 43, at 333-34.

71. *See* Michael J. Bommarito II & Daniel M. Katz, *A Mathematical Approach to the Study of the United States Code*, 389 PHYSICA A 4195, 4195 (2010).

72. *Id.*

Code, and showing that Titles 42, 26, and 5 are the most complex titles of the U.S. Code.[73]

3. *Legal Social Networks*

The studies discussed above, featuring precedent citation networks and regulatory networks, are examples of information network analysis. These types of studies use network analysis to help us understand the structure and dynamics of an information system. Social networks provide another context within which network analysis can shed light on topics of interest to legal scholars. Heinz and Laumann presented one of the first examples of a legal network analysis when they examined the "social structure of the bar" in their work on Chicago lawyers.[74] This work was later extended in an analysis of the structure of policymaker relations in Washington.[75] They find that Washington elites are situated on the periphery of networks with "hollow cores."[76]

As social network analysis became more common throughout academia, a number of scholars focused their analyses on legal social networks. Network approaches have been used to explore how lawyers balance intrafirm pressures to cooperate and compete,[77] how firm members spread the costs of protecting resources across their social networks,[78] how collegiality manifests in social networks within firms,[79] how large firms create international networks of lawyers and legal knowledge by posting lawyers in various offices,[80]

---

73. Katz & Bommarito II, *supra* note 3, at 368.

74. HEINZ & LAUMANN, *supra* note 42.

75. JOHN P. HEINZ ET AL., THE HOLLOW CORE: PRIVATE INTERESTS IN NATIONAL POLICY MAKING 377 (1993).

76. *Id.*

77. *See, e.g.*, Emmanuel Lazega & Marijtje van Duijn, *Position in Formal Structure, Personal Characteristics and Choices of Advisors in a Law Firm: A Logistic Regression Model for Dyadic Network Data*, 19 SOC. NETWORKS 375, 376 (1997).

78. *See, e.g.*, Emmanuel Lazega & David Krackhardt, *Spreading and Shifting Costs of Lateral Control Among Peers: A Structural Analysis at the Individual Level*, 34 QUALITY & QUANTITY 153, 154 (2000).

79. *See, e.g.*, EMMANUEL LAZEGA, THE COLLEGIAL PHENOMENON: THE SOCIAL MECHANISMS OF COOPERATION AMONG PEERS IN A CORPORATE LAW PARTNERSHIP (2001).

80. *See, e.g.*, Jonathan V. Beaverstock, *'Managing across Borders': Knowledge Management and Expatriation in Professional Service Legal Firms*, 4 J. ECON. GEOGRAPHY 157 (2004).

and how firm network structure predicts information flow and service pricing.[81]

4. *Criminology Network Studies*

Criminologists have long been using network analysis in ways relevant to many legal scholars.[82] This work considers social structure as an important lens through which to understand and/or attempt to prevent criminal behavior.[83] Much of the recent work in the area of criminal network analysis has focused on developing ways to identify criminals,[84] as well as ways to target particularly important actors in criminal or terrorist networks so as to impede the network's ability to function.[85]

This expansive body of work demonstrates how network analysis can provide insight into not only the law itself, or the social network of legal practitioners, but also into the social phenomena that the law concerns itself with. By doing so, criminological network analyses offer the potential to help legal scholars explore the implications of policies currently in place, or proposed policy alternatives. In this instance, network analytic techniques are essentially another tool in the "law and" toolbox. Much as the law and economics literature has adopted the power of economics to explain and predict the social phenomena that the law regulates, law and criminological networks offers the potential to help legal

---

81. *See, e.g.*, Uzzi & Lancaster, *supra* note 3, at 331-33.

82. *See* Barry Leighton, *The Community Concept in Criminology: Toward a Social Network Approach*, 25 J. RES. CRIME & DELINQ. 351, 351-52 (1988).

83. *See* Nicholas C. Athey & Martin Bouchard, *The BALCO Scandal: The Social Structure of a Steroid Distribution Network*, 14 GLOBAL CRIME 216, 220 (2013); Rebecca Nash, Martin Bouchard & Aili Malm, *Investing in People: The Role of Social Networks in the Diffusion of a Large-Scale Fraud*, 35 SOC. NETWORKS 686, 687 (2013).

84. *See, e.g.*, Evan C. McCuish, Martin Bouchard & Raymond R. Corrado, *The Search for Suitable Homicide Co-Offenders Among Gang Members*, 31 J. CONTEMP. CRIM. JUST. 319 (2015); Bryce G. Westlake & Martin Bouchard, *Criminal Careers in Cyberspace: Examining Website Failure Within Child Exploitation Networks*, JUST. Q. ONLINE (May 29, 2015), http://tandfonline.com/loi/rjqy20#.Vwf7SvkrKCg [https://perma.cc/SM7G-WCVW].

85. *See, e.g.*, Jonathan David Farley, *Breaking Al Qaeda Cells: A Mathematical Analysis of Counterterrorism Operations (A Guide for Risk Assessment and Decision Making)*, 26 STUD. CONFLICT & TERRORISM 399 (2003); Carlo Morselli, Cynthia Giguère & Katia Petit, *The Efficiency/Security Trade-off in Criminal Networks*, 29 SOC. NETWORKS 143 (2007); Qin et al., *supra* note 44.

scholars understand crime by analyzing the attendant social structures.

II. EMPOWERING LEGAL NETWORK RESEARCH

The previous Part introduced network analysis generally and discussed some of the ways that it has been applied in legal studies. This Part will focus more closely on considerations scholars should take into account when engaging in legal network analysis. While much high quality research has been done in the area of legal network studies, scholars in the future should work to adopt more sophisticated analysis methods, while also striving to ensure that the data they use is as rich and detailed as possible.

A. Improving Legal Network Data

One of the most important considerations for future legal network studies will be to ensure that the data they use remains as rich and detailed as possible. This can be done by resisting the urge to convert nuanced relationships into binary network links, unless it is absolutely necessary to do so. For instance, consider the vast majority of legal citation analyses that have been done. These studies almost universally represent a citation as a binary construct; it either exists or does not.[86] While this allows for efficient analysis, it discards much of the contextual information between the citing and cited documents. In the case of precedent citations, we know that judges include these in their opinions for many different reasons. They are used to provide authority for a rule interpretation, distinguish a case from a similar precedent, note the existence of binding precedent, or for many other reasons.[87] We are perhaps even more aware of this in the law than in other disciplines that perform citation analyses. While there is some appreciation for the varied meanings of citations in other disciplines,[88] legal publishers have

---

86. *See, e.g.*, Fowler & Jeon, *supra* note 3, at 26; Winkels, de Ruyter & Kroese, *supra* note 59; Whalen, *Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Precedent*, *supra* note 3, at 5.

87. Posner, *supra* note 54, at 383-85.

88. *See, e.g.*, Lutz Bornmann & Hans-Dieter, *What Do Citation Counts Measure?: A Review of Studies on Citing Behavior*, 64 J. DOCUMENTATION 45 (2008); Christian Catalini, Nicola Lacetera & Alexander Oettl, *The Incidence and Role of Negative Citations in Science*, 112 PROC. NAT'L ACAD. SCI. 13823 (2015); Martin A. Safer & Rong Tang, *The Psychology of Referencing in Psychology Journal Articles*, 4 PERSP. ON PSYCHOL. SCI. 51 (2009); Xiaojun Wan & Fang Liu,

more than 150 years of experience in working to document the various types of citations that judges make.[89]

In order to avoid losing all of the nuances that citations and other legal network datasets contain, scholars should strive to use weighted or valued networks where possible. To do so means that the network, in addition to representing the nodes and links of interest, contains information of the type or strength of relationship between the nodes.[90] This information can take a number of forms, depending on the data in question. In many cases it makes sense to weight relationships based on how strong they are. So, if we were examining the network of lawyers that argue together at the Supreme Court, it would make sense to weight this network so we do not lose the distinction between those who often argue together and those who have only argued one case together.[91] In this case, we would weight the relationship between lawyers by the number of cases they had worked on together, so those having worked together only once would have a relatively weak relationship, whereas those who had worked together many times would be strongly tied to one another. Doing so allows for a better appreciation of the system, for instance by shedding light on the link–weight distribution, which will demonstrate whether a network is dominated by a small group of strong ties or whether tie strength is more evenly distributed across the network.

Similarly, while it might not always make sense to precisely weight a network, ties can be used to represent different types of relationships. This would be appropriate in the context of judicial citations described above. A citation intended to invoke the authority of a precedent is not necessarily stronger than a citation used to

---

Are All Literature Citations Equally Important?: Automatic Citation Strength Estimation and Its Applications, 65 J. ASS'N FOR INFO. SCI. & TECH. 1929 (2014); Ping Zhou & Loet Leydesdorff, *Fractional Counting of Citations in Research Evaluation: A Cross- and Interdisciplinary Assessment of the Tsinghua University in Beijing*, 5 J. INFORMETRICS 360 (2011).

89. Legal citators like *Shepard's* citations have a long history of classifying judicial citations based on their meaning. *See* Patti Ogden, *"Mastering the Lawless Science of Our Law": A Story of Legal Citation Indexes*, 85 L. LIBR. J. 1, 1-2 (1993); *see also* James F. Spriggs, II & Thomas G. Hansford, *Measuring Legal Change: The Reliability and Validity of* Shepard's *Citations*, 53 POL. RES. Q. 327, 328-29 (2000).

90. NEWMAN, *supra* note 4, at 112-14.

91. For an example of a Supreme Court advocates' network, see *The Supreme Court Advocacy Network*, RYAN WHALEN, http://ryanwhalen.com/2015/03/07/a-supreme-court-advocacy-network/ [https://perma.cc/XZP6-FA6L] (last visited Apr. 6, 2016).

factually distinguish a case, but they are of qualitatively different types. To the extent that legal citation network studies can account for these different citation types, they will strengthen their ability to accurately describe and understand the legal system.

While categorizing links by their weight or their type can help more accurately represent the relationships in question, oftentimes it can also help to represent multiple types of relationships, as in a "multiplex" network. Multiplexity allows for multiple types of links to exist between the nodes in a network. There are many legal contexts in which a multiplex approach would help scholars more accurately understand the structures they analyze. For instance, a multiplex approach could be used to represent relationships between law firms that share links in terms of lateral hiring, clients, law schools they recruit from, cities they have offices in, etc. Taking this sort of multiplex approach would allow for a much more nuanced perspective on the network of law firms, allowing researchers to better understand how they are, or are not, related to one another.

B. Improving Legal Network Analysis Techniques

In addition to striving to work with network data that preserves more of the information present in the legal system, legal network scholars should also attempt to draw on more of the cutting-edge network analytic techniques being developed and applied in the very active network science discipline. This includes both more usage of advanced network analytic techniques, and also more rigorous use of the scientific method, building from theory to hypothesis testing.

1. *Network Statistical Models*

As network analyses have become more common, methodologists have developed increasingly powerful network analytic techniques. Legal network studies, like many other studies of networks, have tended to remain descriptive in nature and have focused overwhelmingly on identifying particularly central—or peripheral—entities within the networks they study.[92] While this approach has provided great insight, it is unable to reveal whether the patterns of centrality described are statistically significant, and does little to help us understand *why* a network may have its observed structure.

92. *See supra* note 88 and accompanying text.

In order to overcome this shortcoming of descriptive network analytic approaches, scholars have developed a collection of statistical modeling techniques, specifically designed to analyze networks. These developments are important because many traditional inferential techniques cannot be easily applied to network data. The most popular statistical techniques rely on the assumption that the data in question satisfies the IID assumption.[93] However, network data violates this assumption. Centrality measures for instance are by definition not independent. One actor's centrality in a network depends on another actor's centrality. Methodologists have developed a number of network analytic techniques to help address this issue.[94] These often rely on simulating null models that researchers can use to compare their observations against, in order to determine the degree of "statistical significance" they can attribute to their observations. One relatively straightforward approach is to use quadratic assignment procedure (QAP) modeling, which permutes the observed network a number of times in order to build a probability distribution for the observed variables.[95] This allows researchers to compute standard errors and thus estimate statistical significance.

Exponential Random Graph Models (ERGMs) are conceptually similar, but much more powerful than a QAP analysis. ERGMs are a family of statistical models developed for network analysis.[96] They rely on bayesian inference, and Markov Chain Monte Carlo (MCMC) simulation to explain network structure as a function of potentially related statistical parameters.[97] This allows researchers to determine the independent effect size (i.e., parameter estimates) for a variety of network structures, and node or link-level attributes that might contribute to the observed network's structure.[98] Applying these sorts of advanced analytic techniques will help legal network analyses move beyond description towards explanation, which will represent a significant step in the progress of empirical legal studies.

---

93. IID stands for independent and identically distributed, and is a data feature assumed by many statistical methods. *See* David A. Freedman, *Statistical Models and Shoe Leather*, 21 SOC. METHODOLOGY 291, 305 (1991).

94. *See, e.g.*, David Krackhardt, *QAP Partialling as a Test of Spuriousness*, 9 SOC. NETWORKS 171, 171-72 (1987).

95. *Id.* at 172-75.

96. *See* David R. Hunter et al., *Ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks*, 24 J. STAT. SOFTWARE 1, 2 (2008).

97. *See id.* at 19.

98. *See id.* at 4-6.

2. *Network Dynamics*

In addition to improved statistical modeling of static networks, legal network analytics could benefit from increased focus on network dynamics. Much of the legal data that lends itself to network analysis is dynamic in nature. Instead of focusing on the network as a static object at one point in time, studies can improve their ability to explain and understand by focusing more on how the network evolves over time.

At the most basic level, dynamic network analysis can simply describe network evolution over time: How have global network properties like degree distribution changed over time? How has the network size and density changed? Do these changes relate to any theoretically relevant phenomena that might explain what is driving network evolution? Describing network changes in this way can provide significantly more insight than simply focusing on the network at a static point in time.

Along with describing dynamic networks, there are a number of statistical techniques capable of explaining network evolution. Both the relational event model (REM)[99] and the stochastic actor oriented model (SAOM)[100] are capable of analyzing dynamic network data and statistically modeling the processes that underlie the network's formation.

3. *Network Theory*

One of the strengths of a network approach is that it allows researchers to simultaneously examine both global structural properties and local traits of the system's constituent parts. This perspective provides the potential for powerful insights into the examined system. However, in order to effectively model systems such as these, researchers need to focus on multiple levels of analysis and likely need to apply multiple theoretical frameworks.[101] Increased use of multitheoretical, multilevel (MTML) approaches will help legal network analyses transition from the largely descriptive work that has been done in the past to work that advances

---

99. *See* Carter T. Butts, *A Relational Event Framework for Social Action*, 38 SOC. METHODOLOGY 155, 159 (2008).

100. *See* Tom A.B. Snijders, *Stochastic Actor-Oriented Models for Network Change*, 21 J. MATHEMATICAL SOC. 149, 149 (1996).

101. *See* Contractor, Wasserman & Faust, *supra* note 17, at 681.

social theory and increases our understanding of not only *how* legal networks are structured, but *why* they are structured as they are.

There are many contexts within which an MTML approach would strengthen legal network analyses. For example, consider the relatively large body of work examining precedent citation networks. This is perhaps the most active legal research area applying network analytic techniques.[102] However, these works tend toward a system-level approach, attempting to understand the citation network as a whole.[103] There is a great deal of research conducted by judicial behaviorists that could supplement precedent citation analyses.[104] In conjunction with one another, the actor-level research on judicial behavior could greatly supplement the system-level work on citation networks. By focusing not just on centrality and structure, but also on why judges choose to cite the cases they do, a combined multilevel approach offers distinct—and perhaps better—insight into why we observe the precedent network structures that we do.

III. AN EMPIRICAL DEMONSTRATION

The above has discussed the basics of network analysis, its background, and some ways that legal network analyses can become more powerful and insightful in their application. This final Part will briefly demonstrate the potential for more rigorous network analysis by examining the patent prior art citation network. In doing so, I will take my own advice and use more nuanced data that does a better job of capturing the reality of the relationships between patents, and I will also engage in some simple dynamic network description, demonstrating how doing so can provide a great increase in explanatory power. This will demonstrate how striving to represent more nuanced relationships within the analyzed network offers better insight into the system in question.

---

102. *See* Marx, *supra* note 10.

103. *See id.* at 121-22.

104. *See, e.g.*, Jilda M. Aliotta, *Combining Judges' Attributes and Case Characteristics: An Alternative Approach to Explaining Supreme Court Decisionmaking*, 71 JUDICATURE 277 (1988); Andrew D. Martin, Kevin M. Quinn & Lee Epstein, *The Median Justice on the United States Supreme Court*, 83 N.C. L. REV. 1275 (2005); Anthony Niblett, *Do Judges Cherry Pick Precedents to Justify Extra-Legal Decisions?: A Statistical Examination*, 70 MD. L. REV. 234 (2010); Rorie Spill Solberg & Stefanie A. Lindquist, *Activism, Ideology, and Federalism: Judicial Behavior in Constitutional Challenges Before the Rehnquist Court, 1986–2000*, 3 J. EMPIRICAL LEGAL STUD. 237 (2006).

A. Patent Citation Networks Generally

In recent years, scholars, pundits, and politicians alike have bemoaned the state of the patent system.[105] Some argue that the way innovation occurs has changed, and that the USPTO is ill-equipped to cope with these changes.[106] However, finding empirical evidence of a qualitative shift in innovation styles is difficult to do. One could take a network approach to this and attempt to measure changes in the patent citation network as a way to demonstrate how innovation has changed.

This approach relies on the citations as markers of the relationship between knowledge. As inventors apply for patents, they are obliged to cite relevant prior art that they are aware of.[107] Similarly, as examiners assess applications, they have the option of adding citations to prior art that they feel is sufficiently related to the claimed invention.[108] As these citations demonstrate the relationship between new knowledge and antecedent knowledge, they provide insight into the innovation process. Innovation relies fundamentally on recombining preexisting knowledge.[109] Thus, if the way innovation occurs has truly altered, we would expect to see these changes reflected in the way that new knowledge is related to antecedent knowledge, and therefore the patterns of prior art citations.

---

105. James Bessen & Michael J. Meurer, *A Third of the Economy Is at Stake—and Patent Trolls Are to Blame*, WASH. POST (Nov. 18, 2015), https://www.washingtonpost.com/news/in-theory/wp/2015/11/18/patent-trolls-are-costing-us-billions-they-must-be-stopped/ [https://perma.cc/M392-62WZ]; Scott Wolinsky, *An Inside Look At The Patent Examination Process*, VOLPE & KOENIG (Mar. 14, 2010), http://vklaw.com/publication/an-inside-look-at-the-patent-examination-process-updated-from-january-2007 [https://perma.cc/2X7C-D5KF] (describing the current situation as a patent system in crisis).

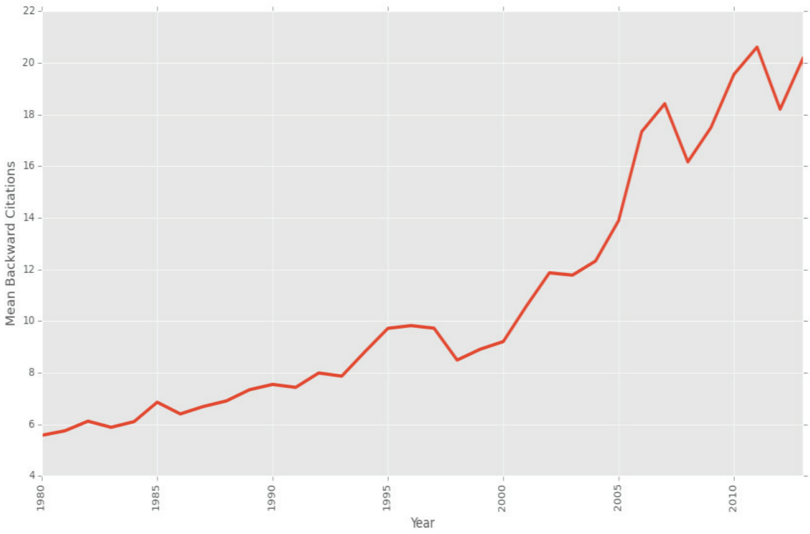
106. *See, e.g.*, COMM. ON INTELLECTUAL PROP. RIGHTS IN THE KNOWLEDGE-BASED ECON. BD. ON SCI., TECH., AND ECON. POLICY & GLOB. AFFAIRS DIV., NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., A PATENT SYSTEM FOR THE 21ST CENTURY 1 (Stephen A. Merrill, Richard C. Levin & Mark B. Myers eds., 2004) ("[B]oth economic and legal changes are putting new strains on the [patent] system.").

107. *See* 37 C.F.R. § 1.56 (2015).

108. *See* MPEP § 2141 (9th ed. Rev. 7, Nov. 2015).

109. *See* JOSEPH A. SCHUMPETER, BUSINESS CYCLES: A THEORETICAL, HISTORICAL AND STATISTICAL ANALYSIS OF THE CAPITALIST PROCESS 14 (1939); ABBOTT PAYSON USHER, A HISTORY OF MECHANICAL INVENTIONS 1-2 (1929); *see also* D. Charles Galunic & Simon Rodan, *Resource Recombinations in the Firm: Knowledge Structures and the Potential for Schumpeterian Innovation*, 19 STRATEGIC MGMT. J. 1193 (1998).

If we were to take a straightforward approach and simply plot the mean out degree (number of backwards citations) of patents over time, we see a stark increase over the last few decades.



This results in a denser patent citation network and suggests that perhaps the way innovation occurs has been changing in recent years.

However, it is difficult to conclude that the innovation environment has substantially changed simply because there are more patent prior art citations. The increase in prior art citations could be due to changes in the way the examination process occurs, or in the way that applicants prepare their applications. While both of these might lead to increased patent citations, neither necessarily demonstrates that the way we innovate has changed. However, by improving the fidelity of our data we can improve our analysis and gain significantly more insight into the way that the innovation environment has changed.

Most patent citation studies treat the relationships between patents as binary. They either exist, or do not. For many studies this has been a necessary concession. When analyzing thousands, or even millions of citations, categorizing them to better capture the nature of the relationship between the citing and cited references is simply intractable. However, advances in computational power and natural language processing techniques enable us to automatically assess the relationship between patents based on their text.

As discussed above, patent prior art citations provide a record of knowledge recombination. This fact can help provide meaningful weighting for the links in a patent citation network. By taking into account the relationship between the knowledge combined, we can meaningfully weight patent prior art citations. In order to understand why the relationship between combined knowledge is a meaningful metric with which to weight patent citations, it is helpful to consider the structure of knowledge and its relationship to creative recombination. One way to do so is to consider the entirety of knowledge as a *knowledge space* through which innovators search in order to find novel and useful recombinations.[110] As innovators search through the knowledge space, they can choose to recombine ideas that are located close to one another (e.g., two methods for insulating a coffee cup) or alternately, they could choose to combine ideas that are very distant from one another (e.g., a method for insulating a coffee cup, and a method for insulating an orbital launch system rocket).

A wide variety of research suggests that the types of ideas that innovators combine are an important predictor of their eventual success.[111] Evidence suggests that mixtures of ideas that are typically combined, along with the addition of some atypically combined ideas, lead to the greatest probability of high success.[112] Similarly, research that brings together ideas for the first time or that haven't been frequently combined in the past tends to be more successful than research that relies on commonly combined ideas or knowledge.[113] This body of research suggests that the knowledge space "distance" between ideas is a vital aspect of the relationship between combined ideas.[114]

Taking into account the knowledge space distance between citing and cited inventions provides useful information that we can use to weight links in the patent citation network. We can do this by comparing the text of each patent and assessing how similar they are to one another. The results below use Latent Semantic Analysis—an

---

110. *See* Fleming & Sorenson, *Science as a Map in Technological Search*, *supra* note 68, at 909.

111. *See, e.g.*, Brian Uzzi et al., *Atypical Combinations and Scientific Impact*, 342 SCIENCE 468 (2013).

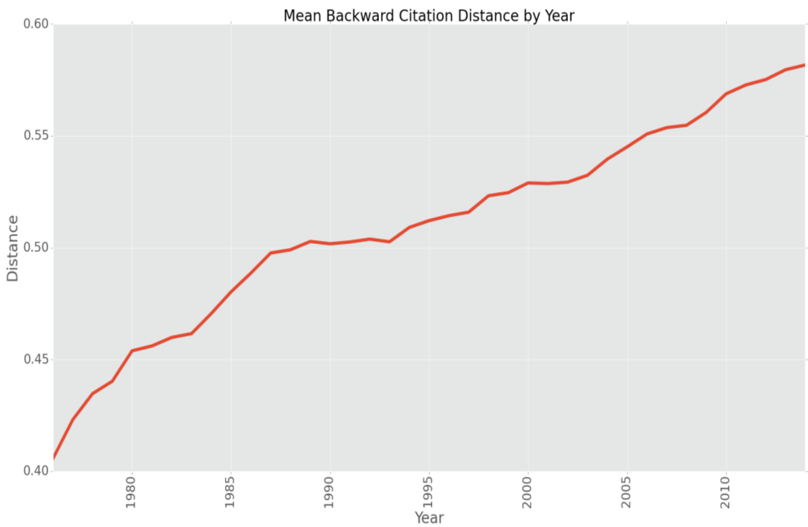
112. *Id.*

113. *See* Jacob G. Foster, Andrey Rzhetsky & James A. Evans, *Tradition and Innovation in Scientists' Research Strategies*, 80 AM. SOC. REV. 875, 895 (2015).

114. *Id.*

established method from the field of information retrieval—to reduce each text to a common set of dimensions.[115] We then compute the cosine distance between each citing/cited pair and use that distance to weight the link between those patents. This weight represents the distance between the content in the citing patent and the content in the cited patent, allowing us to see which patents reach far across the knowledge space to draw on distant knowledge and which primarily draw on proximate knowledge.

Doing so demonstrates that, in addition to the increasing frequency of citations that we observed above, there has been a qualitative shift in the type of citations that are made. Over the past three decades, inventions have become not only more likely to cite prior art, but also more-and-more likely to cite to *distant* inventions.



This suggests that, in addition to the increasing connections that we observed above, the connections in the patent network have steadily linked together more-and-more disparate pieces of knowledge. Not only are there more connections within the patent

115. *See* Scott C. Deerwester et al., *Indexing by Latent Semantic Analysis*, 41 J. AM. SOC'Y FOR INFO. SCI. 391, 391 (1990); Thomas K. Landauer, Peter W. Foltz & Darrell Laham, *An Introduction to Latent Semantic Analysis*, 25 DISCOURSE PROC. 259, 259-60 (1998). The results presented here converts the patent texts to 500 dimension LSA vectors and compares those texts to one another by measuring their distance from one another in multidimensional space.

prior art network, but those connections are qualitatively changing over time.

The heightened complexity of the innovation system, where inventors draw from increasingly distant knowledge, suggests that there has been not only a quantitative change in the way innovation occurs, but also a qualitative change in the way ideas are combined. The twenty-first century has brought not just an increase in innovative activity that the patent office must cope with, but an alteration in the way innovation occurs that the patent office must respond to. This sort of qualitative change in the nature of innovation requires policymakers to rethink the current model that is very much designed to promote and regulate twentieth century innovation. Perhaps one of the greatest challenges that policymakers face is how to effectively assess these increasingly interdisciplinary inventions that draw on knowledge highly distant from their own areas.

Traditionally, the USPTO's examination processes have focused on categorizing the invention claimed into a discrete technological area, and then assigning the assessment to an art-unit that then assigns the application to an individual patent examiner with expertise in that area.[116] In the twentieth century innovation context, this individual-expertise model was appropriate. When inventions tended to draw more on local knowledge, an expert in that area was well-suited to determine whether or not it was sufficiently novel, useful, and nonobvious to merit patentability.[117] However, as inventors increasingly draw on distant areas of knowledge, it becomes less-and-less likely that an individual subject-area expert will be equipped to assess their inventions.

This observation—that the patent network is not only becoming increasingly connected, but is also featuring increasingly distant connections—would not have been possible using a traditional binary-link network analysis. A simpler analysis would have been able to demonstrate the existence of a denser network, but would be unable to show so clearly that the nature of innovation is changing. These changes have important implications for the way the innovation system is structured and the way the USPTO is organized. This Article is not intended to go in-depth into these substantive implications, but rather to show that, by employing somewhat more sophisticated network methods—in this case an analysis that

---

116. *See* MPEP § 903.08(a)-(b) (9th ed. Rev. 7, Nov. 2015).
117. *See* 35 U.S.C. § 101 (2012).

includes weighted links—legal network analysis becomes a more powerful and insightful tool.

As the world increases its adoption of powerful network analytic techniques, legal scholars should look to their peers in industry and in other academic disciplines for approaches that can help improve understanding of the legal system and the society it regulates. By keeping abreast of these developments, and striving to apply rigorous methods in their analyses, legal scholars will be better able to harness the power of network analytics.

CONCLUSION

This Article is intended to serve two purposes. First, I briefly reviewed the literature on legal networks demonstrating that legal network analyses show promise in a variety of applications, including the study of legal information networks, legal social networks, legal organizational networks, and criminal networks. It is our position that network analysis will offer an important set of tools to these and other areas of interest to legal scholars in the future.

Subsequently, I argue that the potential of network analysis is most likely to be fulfilled if scholars strive to apply sophisticated network analytic techniques. The brief demonstration of weighting a patent prior art citation network by semantic similarity shows this in action. I demonstrate that a weighted prior art citation network leads to qualitatively different insight than its binary counterpart. In the future I hope that legal scholars will continue to improve the rigor of their analyses by retaining nuanced data, applying cutting-edge analytic techniques, and theorizing legal networks.

Michigan State
Law Review