A Network Analysis of Inventor Collaboration and Diffusiveness on Patents Granted to U.S. Universities

Margaret M. Clements

mclements6@yahoo.com
Indiana University, Department of Education Leadership and Policy Studies,
Wright School of Education, 4228, Bloomington, IN 47405 (USA)

Abstract

By matching individual level data contained in the 2004 United States Patent and Trademark Office (USPTO) PATSIC, CONAME and INVENTOR data files, inventor networks from 44,394 patents granted to 47,556 unique inventors at 326 U.S. universities, and 722 other commercial and international institutions between January 1, 1975 and December 31, 2004 are analyzed. The networks are studied at five year intervals to limit distortions in the data and to lend stability to the structure of the network for comparative purposes. This study ultimately provides a lens through which the Bayh-Dole Act of 1980 is both analyzed and visualized. This paper: 1) describes features of participation in the co-inventor networks of affiliation on patents granted to U.S. universities from 1975-2004; 2) examines the changing dynamics of inventor productivity as measured by patents produced; 3) analyzes the changing dynamics of collaborative strength and collaborative diversity on patents as measured by frequency and uniqueness of collaboration between inventors; 4) interprets the diffusive impact of U.S. university inventors as measured by citation strength, collaboration strength and collaborative diversity.

Introduction

During and after World War II, the federal government of the United States (U.S.) assumed a significant role in funding and sponsoring research at universities and national laboratories (Etzkowitz & Leydesdorff, 1997; Geiger, 1992, 2004). However, before 1980, there was little collaboration between industry and the academy. Indeed, the federal government, industry and colleges and universities--the three sectors most involved in research and development—rarely collaborated (Geiger, 2004; Kerr, 1995). As the United States (U.S.) economy stagnated in the 1970's, collaboration between these three sectors was seen as increasingly pivotal for the growth of a vigorous economy (Etzkowitz & Leydesdorff, 1997; Geiger, 1988, 2004). Since 1980, several legislative efforts have encouraged inter-sector and intra-sector partnerships to promote the transfer and development of technology between industry, the academy and federal laboratories. In essence, these legislative reforms aimed to shift the government's role in stimulating research from the national level to the local level by privatizing the monetary benefits.

The Bayh-Dole Act, in particular, codified the university's shift to the economic market: higher education institutions receiving federal research and development (R&D) funds became routinely eligible to license and patent inventions resulting from publicly funded research. Internationally, many countries are emulating the U.S. system of technology transfer by aligning their policies and university structures with U.S. policies and structures since the passage of Bayh-Dole (Altbach, 2004, 2007; Kilger & Bartenbach, 2002; Mowery & Sampat, 2005; Normile, 1999). Despite vastly different local conditions, the Bayh-Dole Act is seen as a panacea for higher education institutions around the world. Ultimately, the Bayh-Dole Act has played an important role in catapulting the universities into the global marketplace where applied research and technology transfer mechanisms are pivotal to national policies for economic development and competitiveness (Futao, 2006; Leydesdorff & Wagner, 2009).

Proceedings of ISSI 2009, pages 504–515.
Edited by B. Larsen and J. Leta.

Issues in University Patenting

Advocates for the exploitation of university-developed intellectual property argue that the knowledge production process (social or individual) is not harmed by providing legal protection for intellectual products—particularly to those subject to patents. Furthermore, they assert that the knowledge production process is ultimately strengthened by the reduction of conflicts caused by unfair competition. Likewise, they claim that the process of patenting inventions at universities provides an incentive to utilize research produced in laboratories that might otherwise lay dormant. At the university, funds derived from property rights, it is argued, can promote and fund research, help to employ and train graduate students and international scholars, and prevent the attrition of pioneering scientific researchers from the academy who might otherwise be attracted to the private sector. This line of reasoning stresses that legal protection for intellectual property is an indispensable aspect of stimulating and promoting scientific and technological activities. Thus, university participation in patenting activities is critical to productivity in the science and technology sector.

Critics of intellectual property commercialization at the university point out that proprietary rights accrue more quickly to already advantaged institutions. In the United States, they claim, university commercialization has promoted institutional elitism for universities that are already financially advantaged. Furthermore, they are concerned that the secrecy inherent with proprietary innovation hinders academic freedom and destroys the collegial nature of the university. Especially in the medical field, they believe that the introduction of property rights has caused detrimental change. They complain that "What used to be a scientific community of free and open debate now often seems like a litigious scrum of data-hoarding and suspicion" (Leaf, 2005, p. 152). Likewise, they assert that commercializing technology is expensive. In 2002, North American academic institutions spent over $200 million in litigation. At the same time, over the last six years universities spent $142 million on congressional lobbying efforts (Barton, 2000).

This paper: 1) describes features of participation in the co-inventor networks of affiliation on patents granted to U.S. universities from 1975-2004; 2) examines the changing dynamics of inventor productivity as measured by patents produced; 3) analyzes the changing dynamics of collaborative strength and collaborative diversity on patents as measured by frequency and uniqueness of collaboration between inventors; 4) interprets the diffusive impact of U.S. university inventors as measured by citation strength, collaboration strength and collaborative diversity.

Related Work

This paper expands on the studies of co-author collaboration networks in various scientific journals and scholarly communities (Barabási et al., 2002; Börner, Maru, & Goldstone, 2004; Leydesdorff, 2007; Newman, 2001a, 2001b) . Likewise, the use of citation measures to anlyze the impact of scientific articles influences this study of the impact of patents and particular inventors (Garfield, 1955, 1972, 2004; Leydesdorff & Wagner, 2009). Finally, important developments in data visualization for the analysis of large amounts of data have been achieved that influence this study (Börner, Chen, & Boyack, 2003; Börner, Dall'Asta, Ke, & Vespignani, 2005; Boyack, Klavans, & Börner, 2005; Boyack, Wylie, Davidson, & Johnson, 2000).

Data and Method

This network analysis is performed on the U.S. Patent and Trademark Office's [USPTO] PATSIC, CONAME and INVENTOR data files in addition to the Scholarly Database [SDB] files containing citation information on patents that were granted to U.S. universities between 1975 and 2004. Imprecision in the inventor level data were resolved line by line to correct name misspellings, first and last name inversions and institutional affiliation. This was done by directly contacting the inventor and through publicly available information. This process required thousands of hours to construct a clean and reliable database. Ultimately this network analysis is based on 47,556 unique inventors listed on 44,394 patents issued to U.S. universities involving 1,048 institutions. This paper combines network analyses, patent citation analysis and network visualization to evaluate the effects of the Bayh-Dole Act. Of particular interest is whether or not university patenting has hindered the flow of knowledge or hindered collaborations between inventors. Undirected networks are created from the dataset where each node represents a unique inventor.

Table 1: Summary Statistics for Co-Inventor Networks at Five-year Intervals

Time Interval	75-79	80-84	85-89	90-94	95-99	00-04	75-04
Number of Patents	1674	2269	4118	7474	12978	15881	44394
Mean Patents/Inventor	.06	.09	.18	.35	.69	2.31	2.28
Number of New/Unique Inventors	2008	2488	4740	8295	14055	15970	47556
Mean Collaborations	.08	.14	.29	.65	1.60	2.31	5.08
Mean Collaborative Diversity	.07	.11	.23	.49	1.10	1.63	3.17
Number of Isolated Inventors	385 (19%)	498 (16.8%)	808 (14.3%)	1072 (10.4%)	1224 (6.7%)	1277 (5.4%)	3317 (5.3%)
Number of Cited Patents	1579 (94%)	2149 (95%)	3928 (95%)	7030 (94%)	10843 (83%)	6280 (39%)	31809 (72%)
Total Citations	21470	36509	71876	116117	115659	28144	389775
Mean Citations/cited patent	.578	1.038	2.103	3.296	3.305	.897	7.962
Mean Citations/Inventor	.88	1.57	3.17	5.75	6.59	1.56	19.51

For each time period listed, the statistics are, from top to bottom, total number of patents appearing in the data file issued to inventors at U.S. universities; the mean number of patents per inventor; the number of new and unique inventors entering the network; the mean number of total collaborations on patents per inventor; the mean diversity in collaborations by inventor; the number of isolated inventors with the percent of total inventors listed in parentheses; the number of cited patents with the percent of the total in parentheses; the total citations received during the time period; the mean citations per cited patent per inventor; and the mean citations per inventor.

Discussion of Network Properties

The number of unique inventors in the network grew from 2008 in the 1975-1979 time interval to a total of 47,556 for the entire thirty year period. University scientists have demonstrated an increasing propensity to patent their innovations. As the size of the overall network has increased, so has inventor productivity. The mean number of patents per inventor rose from .06 before 1980 to 2.31 by the end of 2004. The 2000-2004 interval showed a dramatic increase in the productivity of inventors. Because patents represent material and proprietary interests, one would expect inventors to become more guarded about whom they work with in order to maintain a competitive edge or secrecy. It is interesting that increased patenting seems to actually promote both collaboration and collaborative diversity. By the end of 2004, the mean number of inventor collaborators had grown from .08 before 1980 to 2.31. Moreover, before 1980 isolated inventors represented 19% of the network. By the end of 2004, isolated and unconnected inventors only represented 5.3% of the network. Of the total 44,394 patents issued between 1975 and 2004, 25% of the inventors had 5 or more collaborators. Conversely 75% of the inventors had fewer than 5 collaborators. This growing trend toward more frequent collaboration has included more numerous collaborators on patents as well. Although 70% of all patents are assigned to three or fewer inventors, 30% of all university patents include 4 to 20 inventors on a patent.

This analysis points to the observation that total citations to patents have increased over the thirty years studied. Citation strength can be interpreted as the impact of the technologies developed by U.S. university inventors. At the same time, the mean number of citations per patent per inventor has grown despite the fact that the patents issued to these inventors will continue to accrue citations for decades to come. Whereas the mean citations per patent per inventor was .576 before 1980, by the end of 1999, that figure had grown to 3.229. Another indication that scientific importance has not been hampered by university patenting is that a seemingly steady 95% of all university patents accrue citations when truncation—or insufficient aging--is considered. This increasing rate of citation seems to indicate that knowledge is diffusing and that university inventors are contributing substantially to the exchange of ideas through their patenting activities.

Network Measures

In this inventor collaboration network, nodes are the number of inventors and isolates are those inventors who do not collaborate with others on patents. Edges are equivalent to the ties between inventors when they co-patent an invention. From the information about edges and nodes, the remaining analytical features of the graph are computed. The network has evolved from a simple disconnected network to a complex scale-free network over the thirty year period. The average degree increased over time which is an artifact of the growing size and diversity of the inventor network. The power law exponent for the distribution of nodal degrees varied between 2.821 and 3.031. Power law exponents represent the function of preferential attachment in the network (Barabási & Bonabeau, 2003; Barabási, Ravasz, & Vicsek, 2001). Interestingly, it would appear that, despite the inherent competitiveness in the invention and patenting process, there is substantial communication occurring over the inventor network.

Table 2: Inventor Network Analysis

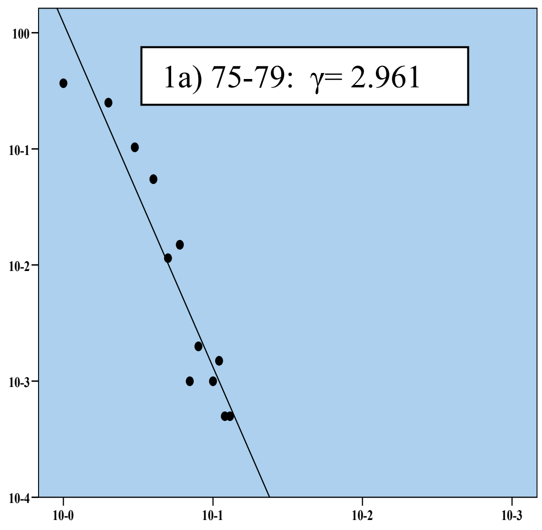
Time Interval	75-79	80-84	85-89	90-94	95-99	00-04	75-04
Network Type	Tree Simple	Tree Simple	Tree Simple	Scale-Free Complex	Scale-Free Complex	Scale-Free Complex	Scale-Free Complex
Patents	1674	2269	4118	7474	12978	15881	44394
Nodes	2008	2954	5662	10273	18108	23501	47556
Isolates	385 19.1%	498 16.8%	808 14.3%	1072 10.4%	1224 6.7%	1277 5.4%	3317 7%
Edges	1611	2524	5467	11724	26151	38761	75464
Average Degree ⟨k⟩	1.604	1.709	1.931	2.282	2.888	3.299	3.174
ASP	1.455	1.825	1.841	3.151	7.209	8.628	13.373
Diameter	5	8	9	15	26	23	44
Density	.0008	.00058	.00034	.00022	.00016	.00014	.00007
Watts-Strogatz Clustering	.885	.903	.894	.885	.891	.891	.844
Connected Components	934	1308	2322	3589	5121	5752	10565
Largest Component	20 1%	37 1.2%	49 .9%	147 1.4%	488 2.7%	948 4.0%	12111 25.5%
Assortativity Pearson's r	.366 .717**	.217 .194	.304 .646*	.238 .640**	.216 .615**	.159 .407*	.013 .059
PL Exponent	2.961	2.807	2.648	3.031	2.858	2.821	2.842
Beta Index	.802	.854	.965	1.141	1.444	1.649	1.587

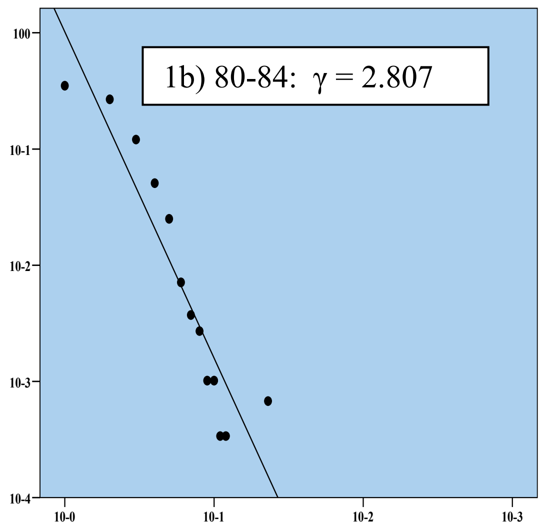
**significant at the .01 level for two-tailed test *significant at the .05 level for two-tailed test

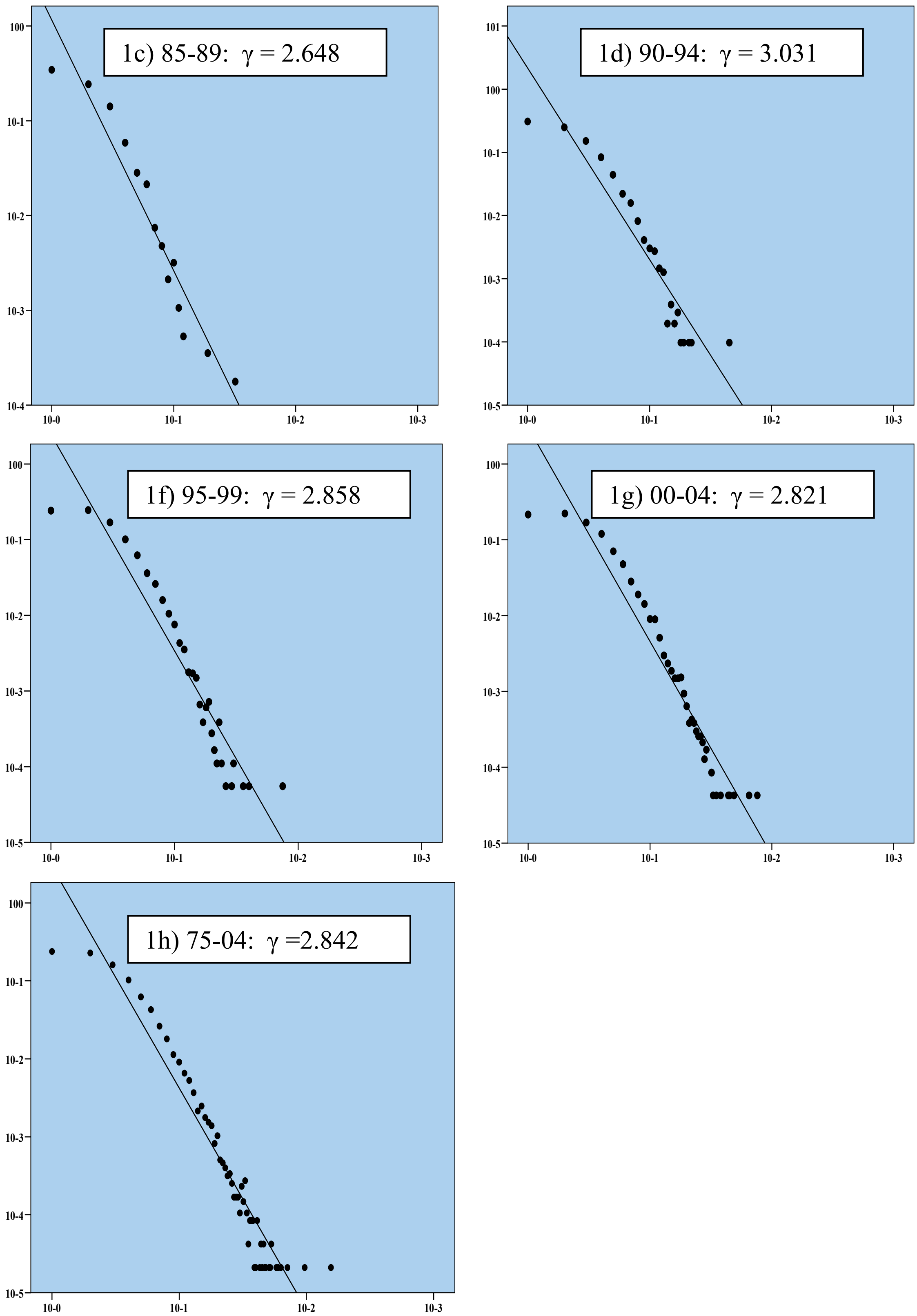
The diameter of the network is decreasing indicating better connections between inventors in the graph. The high clustering coefficient reflects the probability that the nodes connected to one inventor are more likely to connect with each other. The number of connected components has grown from 934 to 5,752 connected components. These represent separate communities of inventors. The largest component now represents 25% of the network. This indicates that inventors are collaborating with each other and that there is substantial communication between inventors. The network is assortative with highly connected inventors generally collaborating with other high degree inventors. The inventor network reaches complexity in the 1990-1994 interval as the beta index arrives at 1.141. Table 2, above, provides detailed analyses of the structural dynamics of the inventor collaboration network.

Power Law Exponent

The emergence of scale free properties in evolving social structures can be evidenced through the plotting of the equally binned degree distributions on a double logarithmic scale. Using curve estimation techniques, lines are fit to the distribution. The resulting slopes are the power law exponents and can be derived for five year intervals. Power law relationships occur when the typical value around which individual measurements are taken are not centered or not distributed normally. Figure 1 a-g, below convey these slopes and the power law exponent is expressed by γ. The emergence of scale free properties arise in this network during the 1990-94 timeframe (figure 1d) as the tail of the degree distribution becomes heavy and the power law exponent is equal to 3.031. As more inventors enter the network, the tail of the distribution becomes even heavier and the power law exponent decreases to 2.821 in the final time period (figure 1f) and is 2.842 (figure 1g), overall. It should be noted that inventors with no connections (isolates) are not represented in these distributions because they have no degree measure. In the inventor network, 70% of the inventors have collaborated with at least one other inventor three or fewer times. Of that 70%, most inventors (38% of the total network) have collaborated with another inventor only once. At the other extreme, a few inventors are especially active collaborators as twenty-two, or .01% of the inventors have collaborated more than 39 times.







Figures 1a-g: Co-inventor Degree Distributions on Log-Log Scale with Corresponding Power Law Exponents for Each Time Step and for the Network Overall

Visualizing the Evolution of the Co-Inventor Network

Visualizations can reveal the overall context and content of a scientific or organizational domain, allow for easy access to multiple levels of data and facility investigation and hypothesis formation (Börner, Sanyal, & Vespignani, 2007; Chen, 2004; Hopcroft, Khan, Kulis, & Selman, 2004; Shneiderman, 2002; Tufte, 2003). Here it is used as an analytical component for the organization of scientific inventors in the context of the U.S. system of higher education. The inventor network, due to the highly specialized nature of many scientific endeavors, remains sparsely connected. The graphs represented in figures 2 through 9 construct a symbol of the interactions between inventors on patents issued to U.S. universities over thirty years. In order to reveal the dynamic evolutionary processes of the network, the graphs are presented in seven panels. Six of the panels relate to the network at each of 5 five year intervals. The last panel presents the comprehensive thirty year network. This graph utilizes a force directed ordination algorithm developed at Sandia National Laboratories that locates similar objects together through the VxOrd program (Börner et al., 2003; Boyack et al., 2005; Boyack et al., 2000) . After calculating the similarity between inventors from the number of times they collaborated on a patent together, the edge values from the connected nodes in the network were fed into VxOrd which was used to determine the coordinates for each institution on the x and y axis. VxOrd then recursively generated cluster assignments through the k-means clustering algorithm. The nodes were positioned on the graph such that the most similar nodes according to edge strength appear closer together. Thus, the closer the nodes are to one another, the more connected they are.

Modelling the diffusion of knowledge by considering aspects of collaboration and technological impact, this visualization provides an opportunity to peer into the invisible college of academic inventors. Because diffusion of knowledge is a focus of this study, those inventors who did not collaborate with others are not represented in this visualization. In addition, to further focus the visualization, those inventors who collaborated with five or more different inventors are included in the following network visualizations. Finally, of all the inventors presented in the graph, fifty were selected as the most diffusive overall by calculating a diffusiveness score. This diffusiveness score was based on the number of different collaborators an inventor worked with and the number of times citations were made to the patents on which he/she was listed as an inventor.

Each circle in the visualization represents a node, or an inventor. The size of the node is dependent on the overall impact of the inventor based on a normalized value for citation strength. Here, the impact is determined by the number of citations to patents that have received citations for each inventor (not all patents receive citations). Then, the color of the nodes represents the inventor’s strength in terms of collaborative diversity. As the color of the node approaches red, it means that the inventor has collaborated with a more diverse array of inventors. Note that some inventors collaborate frequently with the same inventor. In this case, the diffusive impact of such an inventor would be discounted by his/her collaborative diversity score. International inventors are sized small, colored gray and are featured with some transparency so as to not interfere with the ability to see the underlying U.S. inventors.

Because he has been identified as the most “diffusive” inventor in terms of his collaborative intensity, collaborative diversity and patent impact, it is interesting to follow Robert S. Langer’s evolution. He appears for the first time in the 1980-1984 time period as a dark pink node that turns red by the 1985-1989 time period. By 1990-1994, the time period depicted in figure 5, Robert Langer had attracted several different inventors who worked with him on patents—many of whom were international. International collaborations became more noticeable in figure 6, the 1995-1999 time period. It should be noted that inter-disciplinary collaboration grew as well. By the 2004, seventeen separate disciplines are included in the largest connected component. This demonstrates that interdisciplinary cooperation is an important factor in knowledge diffusion and innovation. Notably, it is possible to see the growing complexity of the network which evolves into a scale-free, self-organizing network by 1990-1994, depicted in figure 5, below.

Figure 2: The 50 Most Diffusive Inventors 1975-1979

Figure 3: The 50 Most Diffusive Inventors 1980-1984

Figure 4: The 50 Most Diffusive Inventors 1985-1989

Figure 5: The 50 Most Diffusive Inventors 1985-1989

Figure 6: The 50 Most Diffusive Inventors 1995-1999

Figure 7: The 50 Most Diffusive Inventors 2000-2004

Conclusions and Future Work

The undirected co-inventor networks of university patenting have evolved into scale-free, non random networks of affiliation adhering to models of preferential attachment (Barabási, 2003; Barabási, Ravasz, & Vicsek, 2001; Colizza, Flammini, Serrano, & Vespignani, 2006). These networks appear to be self-selective and governed by affinity. It is interesting to note that as more inventors enter the network, a correlation between collaboration and diffusion (as measured through citation strength) arises. While the overall number of collaborations is increasing, collaborative diversity is also growing. Furthermore, there is growing international participation on patents awarded to U.S. universities.

Visualizing the inventor networks indicates that despite the proprietary nature of patenting, collaboration between inventors has grown. Not only do inventors collaborate more frequently, they collaborate with an increasingly diverse number of inventors. Citation strength of patents does not appear to have been harmed and it would appear as though a great deal of knowledge diffusion is occurring through the inventor networks.

Future studies will demonstrate more information about individual as well as institutional level data such as funding, nationality and evolutionary properties of the inventor network. Likewise, comparing the evolutionary processes in the U.S. network of academic innovators to those in other countries will provide another window of the international scope and dynamics of academic patenting.

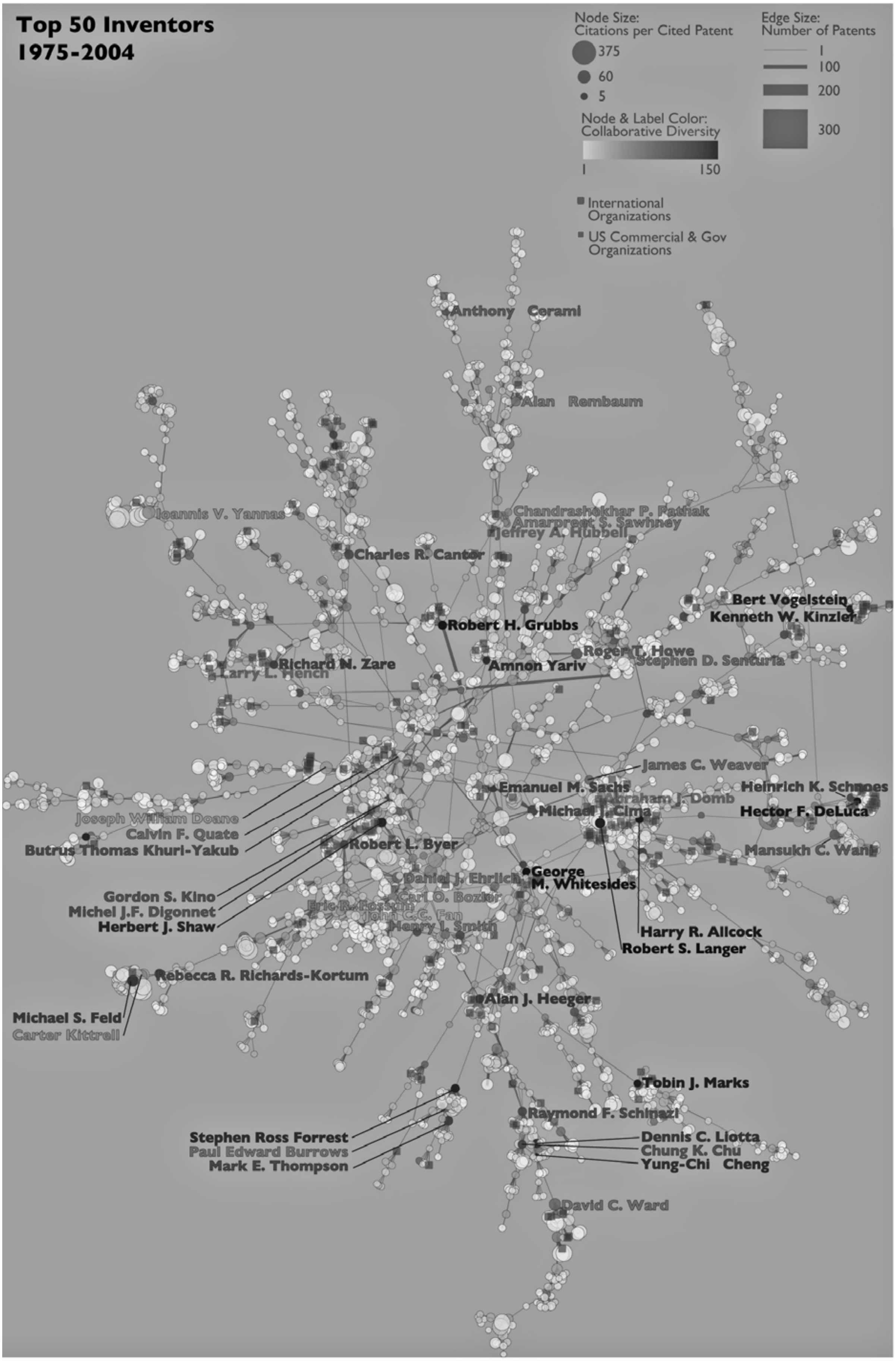


Figure 8: The 50 Most Diffusive Inventors 1975-2004

Acknowledgments

This work has been influenced by Dr. Katy Börner's initiatives at Indiana University to map the world of Science and Dr. Kevin Boyack's work at "Maps of Science." I would like to thank Bruce Herr and Elisha Hardy who helped with the visualizations. Finally, I would like to acknowledge the network workbench tool for network analysis: http://nwb.slis.indiana.edu.

References

Altbach, P. G. (2004). Higher education crosses borders. *Change, 36*(2), 18-24.

Altbach, P. G. (2007). Peripheries and centres: Research universities in developing countries. *Higher Education Management & Policy, 19*(2), 111-134.

Barabási, A.-L., & Bonabeau, E. (2003). Scale-free networks. *Scientific American*, 50-59.

Barabási, A.-L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications, 311*(3-4), 590-614.

Barabási, A.-L., Ravasz, E., & Vicsek, T. (2001). Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications, 299*(3-4), 559-564.

Barton, J. H. (2000). Reforming the patent system. *Science, 287*(5460), 1933-1934.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*.

Börner, K., Dall'Asta, L., Ke, W., & Vespignani, A. (2005). Studying the emerging global brain: Analyzing and visualizing the impact of co-authorship teams. *Complexity, 10*(4), 57-67.

Börner, K., Maru, J. T., & Goldstone, R. L. (2004). The simultaneous evolution of author and paper networks. *Proceedings of the National Academy of Sciences, 101*(suppl. 1), 5266-5273.

Börner, K., Sanyal, S., & Vespignani, A. (2007). Network science. *Review of information science and technology, 41*, 537-607.

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*(3), 351-374.

Boyack, K. W., Wylie, B. N., Davidson, G. S., & Johnson, D. K. (2000). *Analysis of patent databases using VxInsight*™. Paper presented at the 9th International Conference on Information and Knowledge Management.

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences, 101*(suppl. 1), 5303-5310.

Etzkowitz, H., & Leydesdorff, L. (Eds.). (1997). *Universities and the global knowledge economy: A triple helix of university-industry-government relations*. London, England: Pinter.

Futao, H. (2006). Incorporation and university governance: A comparative perspective from China and Japan. *Higher Education Management & Policy, 18*(2), 1-15.

Garfield, E. (1955). Citation indexes for science. *Science, 122*(3159), 108-111.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science, 178*(4060), 471-479.

Garfield, E. (2004). The unintended and unanticipated consequences of Robert K. Merton. *Social Studies of Science, 34*(6), 845-853.

Geiger, R. L. (1988). Milking the sacred cow: Research and the quest for useful knowledge in the American university since 1920. *Science, Technology, & Human Values, 13*(3/4), 332-348.

Geiger, R. L. (1992). Science, universities, and national defense, 1945-1970. *Osiris, 7*, 26-48.

Geiger, R. L. (2004). *Research and relevant knowledge: American research universities since World War II* (2nd ed.). New Brunswick, N.J.: Transaction Publishers.

Hopcroft, J., Khan, O., Kulis, B., & Selman, B. (2004). Tracking evolving communities in large linked networks. *Proceedings of the National Academy of Sciences, 101*(suppl. 1), 5249-5253.

Kerr, C. (1995). *The uses of the university*. Cambridge, MA: Harvard University Press.

Kilger, C., & Bartenbach, K. (2002). New rules for German professors. *Science, 298*(5596), 1173-1175.

Leaf, C. (2005). The law of unintended consequences. *Fortune, 152*.

Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *J. Am. Soc. Inf. Sci. Technol., 58*(9), 1303-1319.

Leydesdorff, L., & Wagner, C. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics, 78*(1), 23-36.
Mowery, D. C., & Sampat, B. N. (2005). The Bayh-Dole Act of 1980 and university-industry technology transfer: A model for other OECD governments? *The Journal of Technology Transfer, 30*(2), 115-127.
Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E, 64*(1), 016131.
Newman, M. E. J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, 64*(1), 016132.
Normile, D. (1999). Japanese faculty show signs of catching start-up fever. *Science, 286*(5439), 397-398.
NWB Team. (2006). Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, http://nwb.slis.indiana.edu.
Shneiderman, B. (2002). *Leonardo's laptop: Human needs and the new computing technologies*. Cambridge, Mass.: MIT Press.
Tufte, E. R. (2003). *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, Conn.: Graphics Press.