The Role of Social Capital in the U.S. Economy: Evidence from Employee-Level Professional Connections from LinkedIn

Shelley Li (University of Southern California)
Frank Nagle (Harvard Business School)
Aner Zhou (University of Southern California)

Abstract

Using a comprehensive dataset on over one billion employee-level professional connections from LinkedIn, we construct a firm-level social capital network and calculate connectivity measures for a sample of more than 11,000 publicly traded U.S. firms from 2004 to 2018. We provide descriptive evidence about firm-level social capital and show that most social capital measures, such as network centrality and the number of various types of connections, are positively and significantly correlated with important firm characteristics such as firm size, net income, R&D expense, and intangible assets. Heat maps of pairwise connections between S&P 100 firms show interesting patterns of connection between the largest firms: higher levels of connectivity in the Finance and Information industries, within the same firm, and amongst firms within the same industry or geography. Industry-level network diagrams show that Information, Finance, and Manufacturing are highly connected with each other and with the rest of the economy, and that Manufacturing is the most central industry (followed by Retail and Information). Our social capital measures are strongly correlated with, yet distinct from, existing measures of economic ties between industries. Using these connectivity measures as another input into a firm's production function, we find positive and significant productivity effects of firm social capital. Our study is the first to use large-scale, individual-level connection data to study firm-level connections, contributing to a more in-depth understanding of firms' strategic positions in the economy and the productivity effect of employee social capital.

Introduction

Connections between firms have long been shown to have important relationships with their economic outcomes (e.g., Podolny, 2001; Uzzi, 1997). In particular, prior studies have examined aspects of firm-level networks based on various types of observed economic ties including supplier-customer relationships at the level of the firm (Atalay et al, 2011; Cohen and Frazzini, 2008) or industry (Antras et al, 2012), patent co-authorship (Breschi and Catalini, 2010), joint ventures (Polidoro et al, 2011), or social and professional connections between high-level firm actors including top executives or board members (Chu and Davis, 2016; Gulati and Westphal, 1999). However, such measures of connectivity only represent the tip of the iceberg in terms of the true depth of connections between firms. In actuality, firms' economic activities are carried out by hundreds or thousands of their employees – whose professional connections could affect the information flow and relational capital upon which business transactions rely. In a world where human capital has steadily become an increasingly important productive input (Ben-Porath, 1967; Black and Lynch, 1996; Romer, 1990), the professional connections of an individual are an extension of their innate human capital, and have important implications for the productivity of the firm who employs them.

In this paper, using a wide-ranging professional connection dataset of the professional social network LinkedIn, we build a firm-level network using employee-level connections for over 11,000 U.S. public firms and generate various network/connection measures for each firm-year from 2004 to 2018. These measures help us gain a deeper understanding of how firms are connected to each other in the modern economy, the types of professional connections they have, which firms are most central in the economic network, and how network measures are correlated with firm characteristics. Further, we estimate the impact of firms' connectivity measures on their total-factor productivity.

Existing studies on firm-level connections face various limitations. For example, input-output tables from governmental sources like the Bureau of Labor and Statistics (BLS) and the Bureau of Economic Analysis (BEA) are at the industry level, not firm specific. Firm-level connections based on board interlock capture a very narrow view of firms' social capital by restricting connections to the board level, and represent an indirect measure of connections by assuming a connection if

two people serve at the same corporate board. Other firm-level connection measures have been built to apply to only a limited set of companies (e.g., co-patenting networks are limited to companies with patents; customer-supplier relationships are limited to a subset of public companies and their major customers). With the rise of online professional social networks like LinkedIn, almost all public companies have employees on these networks who are connected with other professionals. Therefore, our dataset enables us to build a much more comprehensive network of firms based on a much larger set of employees allowing for a more in-depth picture of the connectivity between firms.

Data

LinkedIn is an online professional social network founded in 2003. By the end of 2018, LinkedIn had 624 million users with more than 27 billion connections worldwide. It also has information about 6.9 million companies and 115 million job positions (as reported in user profiles).

To allow for deeper analysis, we chose to focus on employees of publicly traded firms in the United States, for which there is substantial publicly available information. In particular, to generate our sample, we match 19,433 public companies in the CompuStat dataset to companies reported on the LinkedIn platform using the following method. For each CompuStat company, we found the closest LinkedIn company using the *dedupe* package[1], a library for the Python programming language that does a fuzzy match between text items and provides a confidence score from 0 to 1 that represents the accuracy of the match. We improved the matching algorithm by cleaning company names using the python *cleanco* package[2] and incorporating industry classifications as a second variable to match on. 13,971 companies were matched via this algorithmic method, and we manually checked a random subset of these results to ensure accuracy. The remaining 5,462 U.S. public companies and S&P 500 companies with international headquarters were then matched manually. Our analyses focus on three samples of progressively larger firms: the full set of U.S. public companies, the set of S&P 500 companies, and the set of S&P 100 companies.

[1] https://github.com/dedupeio/dedupe
[2] https://github.com/psolin/cleanco

Sample 1 - U.S. public companies. We define a public company as a U.S. public company if its headquarter is in the U.S. We are able to match 11,254 out of 15,438 U.S. public companies between 2004 and 2018. In 2018, 4,683 out of 7,398 U.S. public companies are in our sample with nearly 9 million LinkedIn users employed at those companies, and 1.2 billion connections sent and 1.4 billion connections received by those users.[3] Across the entire sample, these LinkedIn users represent 21% of the total number of company employees (as reported in the public disclosures of the company and captured by Compustat).

Sample 2 – S&P 500 companies. We define a company as an S&P 500 company if it was included in the Standard & Poor's 500 index in a given year, which captures 500 large companies that are representative of the industries in the US economy. Our sample includes 814 out of 826 S&P 500 firms that issued financial statements between 2004 and 2018. In 2018, 505 out of 507 S&P 500 firms are included (the makeup of the S&P 500 can change within a year, hence the total of 507). In 2018, the S&P 500 companies in our sample cover 7 million LinkedIn users with 1 billion connections sent and 1.2 billion connections received. Their LinkedIn users represent 27% of the total number of reported company employees.

Sample 3 – S&P 100 firms. The S&P 100 is a subset of the S&P 500 that includes the largest companies (by market capitalization) of the S&P 500. The S&P 100 represents roughly 60% of the market cap of the S&P 500 and roughly 50% of the market cap of all US public firms. Our sample includes all the 163 S&P 100 firms that issued financial statements between 2004 and 2018. There are 4 million LinkedIn users working for the S&P 100 companies at the end of 2018 with 0.6 billion connections sent and 0.7 billion connections received. Their LinkedIn users represent 28% of the total number of reported company employees.

LinkedIn coverage by industry. Although LinkedIn is the most widely-used professional network in the United States, as alluded to above, not all employees of all companies have a profile on it. Therefore, we consider LinkedIn coverage as the percentage of a firm's employees who are present

[3] On LinkedIn, connections between users are only formed after one user sends a connection request to another user, and that user accepts the request. Here, we report the number of connection requests a user sends and those that they receive separately to allow for a more granular understanding of how the connections are formed.

on LinkedIn and calculated it by dividing the number of profiles on LinkedIn for a given firm (as reported by the employee/LinkedIn user) in that year by the total number of employees reported by the firm (captured in Compustat) for the subset of firms with non-missing Compustat employee data. LinkedIn coverage varies significantly across industries, with Information (NAICS 51), Finance and Insurance (NAICS 52), and Professional, Scientific, and Technical Services (NAICS 54) holding the highest coverage in the U.S. public firm samples. As average company sizes increase (i.e. we move from Sample 1 to Sample 2, and to Sample 3), overall coverage only increases slightly. However, coverage in certain industries increases significantly with size, such as in Mining, Quarrying, and Oil and Gas Extraction (NAICS 21), with its coverage increasing from 33%, to 40% and 54% respectively. Other high-coverage industries also seem to increase their coverage with size (the Information industry's coverage maxes out at 52%, while Real Estate's coverage maxes out at 76%, in the S&P 100 sample).

Using Sample 1 (all public US firms), we counted the total number of connections each employee (of the sample firms) has, removed "extreme" connectors (i.e. those with more than 10,000 connections, a very small proportion of the dataset[4]), and generate a range of firm-level measures based on employee connections, discussed further below. We then use these measures to examine the role social capital plays in firm productivity.

Results

Firm-level connection measures

Using the dataset described above, we build a corporate network based on employee connections. Two firms are connected if there is at least one individual-level LinkedIn connection between these two firms. The strength of the connection between two firms is measured by the natural log of the number of connections between them[5]. Since we aim to measure the "importance" of the firm to the network, we focus on eigenvector centrality as the main network measure in this paper. Eigenvector centrality measures the influence of a node in the network (Newman, 2008). Companies with strong connections to more central firms have higher eigenvector centrality. In

[4] Out of the more than 9 million employees in our sample, only fewer than 3000 were dropped due to "extreme" number of connections.

[5] We use the natural log of the number of connections, rather than the number of connections between firms, because the existence of a power law is usually observed in social networks.

addition to centrality measures, we also generated other firm-level connection measures to capture the extent and type of firm connections. Table 1 shows the descriptive statistics for these firm-level connectivity measures for the year 2018[6]. In our sample of U.S. public firms in 2018, on average, there are 1,887 employees on LinkedIn, with an average of 337 connections per employee (and a total of almost 553,000 connections per firm). Since the diversity of an individual's network has previously been shown to impact economic outcomes (Eagle, Macy, and Claxton, 2010). Therefore, we dig into the type of connections individuals have, and find that, on average, 4% of all connections are internal connections (i.e. connections made with employees currently in the same company), 24% are connections with people who have the same education level, 20% are connections with those who have the same job experience level, 64% are connections with those who have the same job function, 11% are connections with those who have gone to the same school, and 14% are connections with those who used to work in the same company.

----------*Insert Table 1* ----------

Based on the eigenvector centrality measure, we generated a list of the top 100 most central US public companies in the economy (as of 2018). The list is in Table 2. It is important to recognize that the dataset is not the complete set of all individuals, at all companies, and all of their connections, as discussed above. However, it is interesting to note the different types of companies and how central they are to the economy as captured in our data. In particular, none of the firms, even the most central firms, dominate the centrality of the professional social network (the largest eigenvector centrality observed is 0.068). Further, most of the top 20 firms are tech and finance companies, but we see companies from industries including healthcare, consumer goods, retail, and real estate breaking into the top 30. The global clustering coefficient for all US public companies is 0.79, which indicates that the firms are fairly tight knit as the measure goes from zero to one. Examining the global clustering coefficient by industry indicates a small degree of variation across industries, with NAICS 81 (Other service (except public admin)) having the lowest clustering coefficient at 0.70 and NAICS 11 (agriculture, forestry, fishing and hunting) having the highest clustering coefficient at 0.88.

[6] The numbers of observations for different connectivity measures differ as some individuals have missing information in education, seniority level, or job function. Further, the total number of observations is well short of the 11,254 firms matched in Sample 1 as some firms have gone out of business, were acquired, or were no longer public for other reasons as of 2018.

----------*Insert Table 2* ----------

Correlation between connection measures and firm characteristics

In Table 3, we show the correlation between the firm-level connectivity measures and important firm characteristics which were summarized in Table 1. All numbers in **Bold** (*italic*) indicate the correlation is statistically significant at the 1% (5%) level. Eigenvector centrality and total number of firm connections are positively and significantly correlated with all measures of firm size (revenue, assets, number of employees), net income, and intangible assets. Measures representing centrality and total number of firm connections (in general or of a certain type) are positively and significantly correlated with R&D expense. Comparing measures scaled by firm size (assets), we find that the average number of connections is positively and significantly correlated with R&D intensity. Eigenvector centrality is positively and significantly correlated with ROA (return on assets) (at the 5% significance level).

----------*Insert Table 3* ----------

Pairwise Connections between the largest firms

In Figure 1, we present a heat map representing the 2018 S&P 100 firm-to-firm network. The color of the box represents the number of realized connections initiated from one company (horizontal) going to the other (vertical). Therefore, the colorings are not perfectly symmetric as even if two companies are tightly connected, one of them may be initiating more of the connections than the other. For the boxes where the same company is both on the vertical and horizontal, the color represents the number of internal connections at the company itself. Firms are grouped based on industry (two-digit NAICS), and within each industry, ordered based on size. Each row provides the composition of connections made by the firm. The number of connections has been log-transformed to the scale on the right and darker colors represent larger number of connections. The figure suggests: (1) there are a high number of connections within firms and amongst firms in the same industry. (2) The Finance and Insurance industry, and the Information industry show the highest level of connectivity, which is partially driven by the higher percentage of employees in those industries that are on LinkedIn. (3) There is a geographic element to the connections as well.

For example, companies that are headquartered in Minnesota (e.g., UnitedHealth Group, Target, 3M, and US Bancorp) are well connected to each other, even though they are in different industries. This adds support to prior literature that has discussed the importance of managerial talent to the growth of such geographical clusters (Shaver, 2018). We generated the same heat map for each year going back to 2004 in Figure A1 to Figure A14 (in the Appendix). Looking at heat maps over the years, we see that the connections became overall denser and denser. Even in the very early years (2004, 2005), we can already see a strong diagonal line (indicating connections with those in the same firm) and a denser concentration in the Information industry. These patterns still exist when we use the color of the box to represent the total number of connections between two firms (i.e. without considering who initiated the connection).

----------*Insert Figure 1* ----------

Industry-level network diagrams

Figures 2, 3, and 4 show the industry-level pairwise connections for the S&P 100, S&P 500, and all US public firms at the end of 2018. The darker the line between two industries, the more connections these two industries have to each other and the size of the node represents the centrality of the industry in the overall network. Information, Finance and Insurance, and Manufacturing are highly connected overall and with each other. Manufacturing is the most central industry in this network (which is partially a result of it being the most heavily represented industry in the S&P 100), while Retail Trade and Information are the second and third most central industry, respectively. It is interesting to note that although companies in the Information industry are very well connected in the network of S&P 100 firms (as seen in Figure 1), the industry as a whole is slightly less central compared to traditional industries like Manufacturing and Retail Trade.

We generated the same diagrams for each year going back to 2004 in Figure A15 to Figure A53 (in the Appendix). Similarly to the heat maps, we see that the connections became denser over the years. In the very early years (2004, 2005), Manufacturing, Information, Finance and Insurance, were already the most central industries, and their centralities relative to other industries decrease over time.

----------*Insert Figure 2* ----------

----------*Insert Figure 3* ----------

Comparison with other firm-level network measures

Comparing the firm-level network measures from the sample of all US public firms to the 2018 BEA/BLS input-output tables at the same industry level (which measures sales from one industry to another), we find that the natural log of the number of connections between two two-digit NAICS industries is positively correlated with the natural log of the input-output between the same two industries, with a coefficient of 0.33 and a significance level of 0.00. This indicates that our measure of firm-level connectivity strongly correlates with measures of economic ties between industries, but measures something distinct that is not fully captured by flows of sales between industries.

We further compare our network measures with connection measures calculated with data on boards of directors. In 2018, only 0.2% of connected company pairs based on our sample data (of US public firms) had interlocked board members, suggesting that the firm connectivity measures calculated with board interlock data only apply to a very small portion of U.S. public firms. Among these 0.2% company pairs, the natural log of the number of connections between these two firms is not significantly correlated (at the 5% level) with the natural log of the number of interlocked board members, with a coefficient of -0.03 and a significance level of 0.06. This suggests that our measures based on employee-level professional connections can provide insights not covered by traditional network measures based on top-level actors in firms.

Productivity impact of firm-level connectivity

It has long been known that human capital is a critical input into the productivity of the firm (Romer, 1990). However, our data allows us to more deeply consider the role of employee *social capital* in firm productivity. To examine this relationship, we consider various measures of social capital, as captured in the LinkedIn data metrics discussed above. We use these measures in a total factor productivity (TFP) framework with firm-level employee social capital as an input into the production function in addition to the standard inputs of capital and labor. This is consistent with prior literature that has used TFP to calculate the productivity impact of inputs that are traditionally

difficult to measure (Bloom and Van Reenen, 2002; Brynjolfsson and Hitt, 1996; Nagle, 2019). Such analysis starts with a log-transformed Cobb-Douglas equation and adds the variable of interest into the equation, yielding our primary regression equation:

$$\ln(Y_{it}) = \alpha \ln(K_{it}) + \beta \ln(L_{it}) + \gamma SC_{i,t-1} + \varepsilon_{it} \quad (1)$$

In this equation, Y_{it} is the output of firm *i*, in year *t*, as measured by real sales[7]. K_{it} is the amount of capital invested by firm *i*, in year *t*[8], and L_{it} is the number of employees at the firm[9]. $SC_{i,t-1}$ is the level of social capital as measured by the various measures discussed in Table 1, including eigenvector centrality, total number of connections, and total number of specific types of connections (e.g. internal connections, connections with those who have the same education level or job function). Due to high-levels of correlation between these measures, each is only used one at a time in the analysis. We use the lagged values for these connectivity measures, rather than the current year, to capture the fact that social capital primarily comes from existing ties, rather than newly formed ones. However, results with current year social capital are generally similar. Importantly, although our descriptive statistics above only show firm connectivity at the end of 2018, we use historical data that go as far back as 2004 for the TFP analysis. This allows us to construct a large dataset of firm-years that considers differences in social capital both across firms, and within firms over time (using firm fixed effects). All models include a control for year and NAICS2 industry or firm fixed effects to account for variance in LinkedIn coverage and time trends. Additionally, when the social capital measure of interest is related to a measure of the number of connections at the firm that have a specific characteristic (e.g., the number of employee connections with other people that have the same level of education), a control is included for the total number of employee connections at the firm to help account for the volume of connections. For measures measuring the number of connections, we take the natural log of the measure.

[7] REVT in Compustat deflated by implicit price deflators for gross domestic product (GDP) from the Bureau of Economic Analysis (BEA).

[8] Calculated as the sum of capital stock when the firm first became publicly traded and new investments in each year until year *t*. Capital stock is measured as the firm's net total property, plant, and equipment (PPENT in Compustat) when the firm became public deflated by the GDP deflator for private nonresidential fixed investment from the BEA for that year. The net investment for each year is the difference between in PPENT from *t-1* to *t*, deflated by the appropriate GDP deflator for private nonresidential fixed investment from the BEA.

[9] Results are robust to using labor cost as the labor input measure in the analyses.

Table 4, Column 1 shows coefficients from our baseline regression model which includes industry and year fixed effects. Column 2 uses firm fixed effects[10]. The values shown in the table indicate the coefficients γ (and α, β) from equation 1 for a given social capital measure (i.e. they show coefficients from many regressions, each with a different social capital measure). Coefficients significant at the 1% level are in **bold**. We can see that eigenvalue centrality and various other firm-level connection measures have positive productivity effects. In general, the productivity effects of these connectivity measures become smaller in magnitude when we control for firm fixed effects but remain statistically significant. The positive coefficient on the total number of connections with those who have the same job functions even became statistically significant after including the firm fixed effects. These results show that total amount of employee social capital (as measured by eigenvector centrality, total number of connections, and average number of connections) has positive productivity effects. Controlling for the total amount of employee social capital, we find that the more connections that share similar characteristics (e.g. with those from the same company, the same job function, the same experience level, having gone to the same school, having worked at the same past institution), the higher the productivity effect from these types of social capital. Interpreting the coefficients from the first and second rows of columns 1 yields the following economic impact. An increase of 1 standard deviation in eigenvector centrality (e.g., an increase of 0.015 with a mean of 0.010) leads to a 13% increase in output (sales); an increase of 1 standard deviation in total number of employee connections leads to a 7% increase in output. In comparison, an increase of 1 standard deviation in capital leads to an 85% increase in output.

----------*Insert Table 4* ----------

Table 5 shows the same analysis as Table 4 Column 2 (with firm fixed effects), but includes an interaction of the various measures of social capital with the measure for labor input (number of employees) at the firm (L_{it}). This allows for deeper insights to be drawn by understanding if social capital complements or substitutes for the productivity effect of labor input. For most types of firm connections, there is no significant interaction effect between social capital and labor.

[10] Currently obtained using "demeaning variables" method to run the fixed effects analyses due to analysis limitations.

However, we do see a positive and significant interaction effect between eigenvector centrality and labor (the effect is negative and significant between total number of connections and labor). This suggests that social capital could either complement or substitute the productivity effects of labor input depending on what kind of social capital a firm's employees possess. The difference seems to come from whether the firm is in a structurally important network position (i.e. centrally connected in the professional network), or just has a lot of connections.

----------*Insert Table 5* ----------

Conclusion

In this paper, we use a novel dataset capturing employee-level professional connections between companies to shine new light on a question that economists and sociologists have wrestled with for some time. We show that there is a great deal of connectivity between US public firms, but that this connectivity varies on numerous dimensions. Further, we show that social capital of employees (as measured by connections on LinkedIn), most of whom are generally not involved in high-level strategic decisions, has a measurable impact on firm productivity. Although this dataset provides much more granular insights into the connections between firms than existing data, it is important to note its limitations, particularly that the connections seen on LinkedIn are not a random sample of all people, and over-represent certain industries, white-collar workers, and larger companies (especially given our focus on public companies). Despite these limitations, we believe using employee-level connections as a new measure of firm connectivity allows a more granular analysis than existing measures and will allow scholars to revisit earlier studies of firm networks to explore new avenues through which this connectivity may impact productivity and other social and economic outcomes.

References and Notes

Antràs, P., Chor, D., Fally, T., & Hillberry, R. (2012). Measuring the upstreamness of production and trade flows. *American Economic Review*, *102*(3), 412-16.

Atalay, E., Hortacsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences*, *108*(13), 5199-5202.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of political economy*, *75*(4, Part 1), 352-365.

Black, S. E., & Lynch, L. M. (1996). Human-capital investments and productivity. *The American economic review*, *86*(2), 263-267.

Bloom, N., & Van Reenen, J. (2002). Patents, real options and firm performance. *The Economic Journal*, *112*(478), C97-C116.
Breschi, S., & Catalini, C. (2010). Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. *Research Policy*, *39*(1), 14-26.
Brynjolfsson, E., & Hitt, L. (1996). Paradox lost? Firm-level evidence on the returns to information systems spending. *Management science*, *42*(4), 541-558.
Chu, J. S., & Davis, G. F. (2016). Who killed the inner circle? The decline of the American corporate interlock network. *American Journal of Sociology*, *122*(3), 714-754.
Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*(4), 1977-2011.
Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. *Science*, *328*(5981), 1029-1031.
Gulati, R., & Westphal, J. D. (1999). Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Administrative science quarterly*, *44*(3), 473-506.
Nagle, F. (2019). Open source software and firm productivity. *Management Science*, *65*(3), 1191-1215.
Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics*, *2*(2008), 1-12.
Podolny, J. M. (2001). Networks as the pipes and prisms of the market. *American journal of sociology*, *107*(1), 33-60.
Polidoro Jr, F., Ahuja, G., & Mitchell, W. (2011). When the social structure overshadows competitive incentives: The effects of network embeddedness on joint venture dissolution. *Academy of Management Journal*, *54*(1), 203-223.
Romer, Paul M., 1990. "Human capital and growth: Theory and evidence," Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 32(1), pages 251-286, January.
Shaver, J. M. (2018). *Headquarters Economy: Managers, Mobility, and Migration*. Oxford University Press.
Ugander, J., Karrer, B., Backstrom, L. and Marlow, C., 2011. The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*.
Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative science quarterly*, 35-67.

Table 1 – Descriptive Statistics for Firm-level Measures (2018)

Measures	No. Obs.	Mean	Std. Dev.	Min	25%	50%	75%	Max
Eigenvector centrality	3,787.00	0.011	0.012	0.000	0.001	0.006	0.017	0.068
Number of employees on LinkedIn	4,683.00	1,887.03	9,262.99	0.00	2.00	46.00	548.00	285,561.00
Total number of employee connections	4,683.00	552,915.70	2,968,442.00	0.00	501.50	14,840.00	157,553.50	90,029,070.00
Average number of connections	3,857.00	337.15	273.05	0.00	199.38	287.16	414.63	6,625.00
Total number of employee *internal* connections	4,683.00	54,822.45	357,192.30	0.00	0.00	358.00	9,772.00	12,296,470.00
Total number of employee connections with people who have the same education level	3,724.00	180,920.30	902,822.50	0.00	1,208.75	8,852.00	61,663.50	22,624,620.00
Total number of employee connections with people who have the Same experience level	3,844.00	154,238.30	792,319.90	0.00	813.00	6,437.50	49,841.50	23,601,360.00

Total number of connections with people who have the same job functions	3,844.00	453,888.40	2,281,415.00	0.00	2,973.50	21,432.50	153,808.80	61,984,910.00
Total number of connections with people who have gone to the same school	3,773.00	85,293.82	433,372.60	0.00	459.00	3,753.00	26,300.00	10,427,440.00
Total number of connections with people who have worked at the same institution.	3,857.00	127,806.90	713,203.10	0.00	397.00	4,832.00	36,490.00	20,181,020.00
Revenue (in millions)	4,616.00	3,717.35	16,582.94	-278.11	29.19	301.82	1,751.23	511,729.00
Assets (in millions)	4,640.00	11,712.31	95,805.92	0.00	100.10	802.21	3,856.32	3,418,318.00
Number of employees (in thousands)	4,386.00	9.48	47.53	0.00	0.08	0.70	4.70	2,200.00
ROA	4,410.00	-1.32	20.32	-735.00	-0.09	0.03	0.08	1.75
R&D Expense	2,345.00	179.30	1,132.29	0.00	0.31	11.21	56.00	28,837.00
Net Income (in millions)	4,616.00	300.28	1,815.64	-22,355.00	-9.24	9.06	125.83	59,531.00
R&D/Assets	4,629.00	0.16	2.27	0.00	0.00	0.00	0.05	136.33
Intangible assets (in millions)	4,573.00	1,471.82	8,327.28	0.00	0.00	21.94	376.08	310,197.00

Table 2 Top 100 Firms with Highest Eigenvector Centrality in 2018

#	Stock Ticker	Company Name	Eigenvector Centrality
1	AMZN	AMAZON.COM INC	0.068
2	IBM	INTL BUSINESS MACHINES CORP	0.067
3	MSFT	MICROSOFT CORP	0.066
4	ORCL	ORACLE CORP	0.065
5	GOOG	ALPHABET INC	0.065
6	WFC	WELLS FARGO & CO	0.062
7	AAPL	APPLE INC	0.062
8	JPM	JPMORGAN CHASE & CO	0.061
9	CSCO	CISCO SYSTEMS INC	0.060
10	C	CITIGROUP INC	0.059
11	CRM	SALESFORCE.COM INC	0.059
12	T	AT&T INC	0.058
13	BAC	BANK OF AMERICA CORP	0.058
14	ADP	AUTOMATIC DATA PROCESSING	0.057
15	FB	FACEBOOK INC	0.057
16	CTSH	COGNIZANT TECH SOLUTIONS	0.056
17	HPE	HEWLETT PACKARD ENTERPRISE	0.056
18	MS	MORGAN STANLEY	0.056
19	UNH	UNITEDHEALTH GROUP INC	0.055
20	INTC	INTEL CORP	0.054
21	JNJ	JOHNSON & JOHNSON	0.054

22	COF	CAPITAL ONE FINANCIAL CORP	0.054
23	PEP	PEPSICO INC	0.054
24	HON	HONEYWELL INTERNATIONAL INC	0.053
25	WMT	WALMART INC	0.053
26	GS	GOLDMAN SACHS GROUP INC	0.053
27	TGT	TARGET CORP	0.053
28	IT	GARTNER INC	0.052
29	CBRE	CBRE GROUP INC	0.052
30	GM	GENERAL MOTORS CO	0.052
31	BA	BOEING CO	0.052
32	PFE	PFIZER INC	0.051
33	LMT	LOCKHEED MARTIN CORP	0.051
34	ABT	ABBOTT LABORATORIES	0.050
35	AXP	AMERICAN EXPRESS CO	0.050
36	VZ	VERIZON COMMUNICATIONS INC	0.050
37	TSLA	TESLA INC	0.050
38	NKE	NIKE INC -CL B	0.049
39	DXC	DXC TECHNOLOGY CO	0.049
40	NOC	NORTHROP GRUMMAN CORP	0.049
41	VMW	VMWARE INC -CL A	0.049
42	GE	GENERAL ELECTRIC CO	0.049
43	USB	U S BANCORP	0.049
44	CVS	CVS HEALTH CORP	0.049
45	TMO	THERMO FISHER SCIENTIFIC INC	0.049

46	SYK	STRYKER CORP	0.049
47	PNC	PNC FINANCIAL SVCS GROUP INC	0.049
48	HD	HOME DEPOT INC	0.049
49	ALL	ALLSTATE CORP	0.048
50	ADBE	ADOBE INC	0.048
51	CMCSA	COMCAST CORP	0.048
52	F	FORD MOTOR CO	0.048
53	SBUX	STARBUCKS CORP	0.048
54	HPQ	HP INC	0.048
55	UPS	UNITED PARCEL SERVICE INC	0.048
56	PG	PROCTER & GAMBLE CO	0.048
57	RTN	RAYTHEON CO	0.047
58	MMM	3M CO	0.047
59	BAH	BOOZ ALLEN HAMILTON HLDG CP	0.047
60	KO	COCA-COLA CO	0.046
61	LOW	LOWE'S COS INC	0.046
62	MRK	MERCK & CO	0.046
63	MAR	MARRIOTT INTL INC	0.046
64	KELYA	KELLY SERVICES INC -CL A	0.046
65	DIS	DISNEY (WALT) CO	0.045
66	RHI	ROBERT HALF INTL INC	0.045
67	AIG	AMERICAN INTERNATIONAL GROUP	0.045
68	BSX	BOSTON SCIENTIFIC CORP	0.045
69	XOM	EXXON MOBIL CORP	0.045

70	KFRC	KFORCE INC	0.045
71	SCHW	SCHWAB (CHARLES) CORP	0.045
72	AMP	AMERIPRISE FINANCIAL INC	0.045
73	CAH	CARDINAL HEALTH INC	0.045
74	IQV	IQVIA HOLDINGS INC	0.045
75	V	VISA INC	0.045
76	WDAY	WORKDAY INC	0.044
77	BBY	BEST BUY CO INC	0.044
78	AMGN	AMGEN INC	0.044
79	AAL	AMERICAN AIRLINES GROUP INC	0.044
80	MET	METLIFE INC	0.044
81	PRU	PRUDENTIAL FINANCIAL INC	0.044
82	S	SPRINT CORP	0.044
83	LLY	LILLY (ELI) & CO	0.044
84	BDX	BECTON DICKINSON & CO	0.044
85	M	MACY'S INC	0.044
86	PYPL	PAYPAL HOLDINGS INC	0.044
87	CHTR	CHARTER COMMUNICATIONS INC	0.044
88	CVX	CHEVRON CORP	0.044
89	MCK	MCKESSON CORP	0.044
90	CMCSA2	NBCUNIVERSAL MEDIA LLC	0.044
91	CTL	CENTURYLINK INC	0.044
92	BK	BANK OF NEW YORK MELLON CORP	0.043
93	CERN	CERNER CORP	0.043

94	AFL	AFLAC INC	0.043
95	DAL	DELTA AIR LINES INC	0.043
96	FIS	FIDELITY NATIONAL INFO SVCS	0.043
97	INTU	INTUIT INC	0.043
98	ABBV	ABBVIE INC	0.043
99	JWN	NORDSTROM INC	0.043
100	BMY	BRISTOL-MYERS SQUIBB CO	0.043

Table 3 Correlation between Connectivity (Social Capital) Measures and Firm Characteristics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Eigenvector centrality	**1**																	
2. Number of employees on LinkedIn	**0.594**	**1**																
3. Total number of employee connections	**0.561**	**0.944**	**1**															
4. Average number of connections	**-0.046**	**-0.036**	-0.001	**1**														
5. Total number of employee *internal* connections	**0.491**	**0.904**	**0.967**	-0.002	**1**													
6. Total number of employee connections with people who have the same education level	**0.548**	**0.923**	**0.993**	0.002	**0.948**	**1**												
7. Total number of employee connections with people who have the same experience level	**0.54**	**0.947**	**0.995**	-0.002	**0.961**	**0.988**	**1**											
8. Total number of connections with people who have the same job functions	**0.548**	**0.927**	**0.998**	0.002	**0.973**	**0.992**	**0.99**	**1**										
9. Total number of connections with people who have gone to the same school	**0.544**	**0.893**	**0.963**	-0.003	**0.906**	**0.98**	**0.948**	**0.963**	**1**									

10. Total number of connections with people who have worked at the same institution.	**0.512**	**0.908**	**0.985**	0.004	**0.985**	**0.976**	**0.978**	**0.993**	**0.942**	**1**								
11. Revenue (in millions)	**0.47**	**0.608**	**0.522**	**-0.062**	**0.471**	**0.512**	**0.501**	**0.513**	**0.551**	**0.488**	**1**							
12. Assets (in millions)	**0.305**	**0.375**	**0.326**	**-0.028**	**0.265**	**0.38**	**0.368**	**0.364**	**0.403**	**0.34**	**0.419**	**1**						
13. Number of employees (in thousands)	**0.386**	**0.554**	**0.422**	**-0.073**	**0.372**	**0.404**	**0.409**	**0.405**	**0.415**	**0.377**	**0.742**	**0.214**	**1**					
14. ROA	*0.050*	0.014	0.013	*-0.035*	0.011	0.012	0.012	0.012	0.012	0.011	0.016	0.008	0.014	**1**				
15. R&D Expense	**0.421**	**0.632**	**0.731**	0.007	**0.657**	**0.774**	**0.718**	**0.738**	**0.83**	**0.717**	**0.477**	**0.582**	**0.268**	0.011	**1**			
16. Net Income	**0.418**	**0.605**	**0.589**	*-0.032*	**0.536**	**0.595**	**0.57**	**0.589**	**0.641**	**0.571**	**0.604**	**0.547**	**0.31**	0.012	**0.622**	**1**		
17. R&D/Assets	**-0.077**	-0.012	-0.01	**0.025**	-0.008	-0.016	-0.016	-0.016	-0.016	-0.014	-0.014	-0.008	-0.013	**-0.525**	-0.009	-0.011	**1**	
18. Intangible assets	**0.391**	**0.455**	**0.408**	-0.026	**0.388**	**0.383**	**0.392**	**0.401**	**0.372**	**0.396**	**0.545**	**0.305**	**0.35**	0.012	**0.333**	**0.461**	-0.011	**1**

Figure 1 Connection Heat Map for S&P 100 Firms

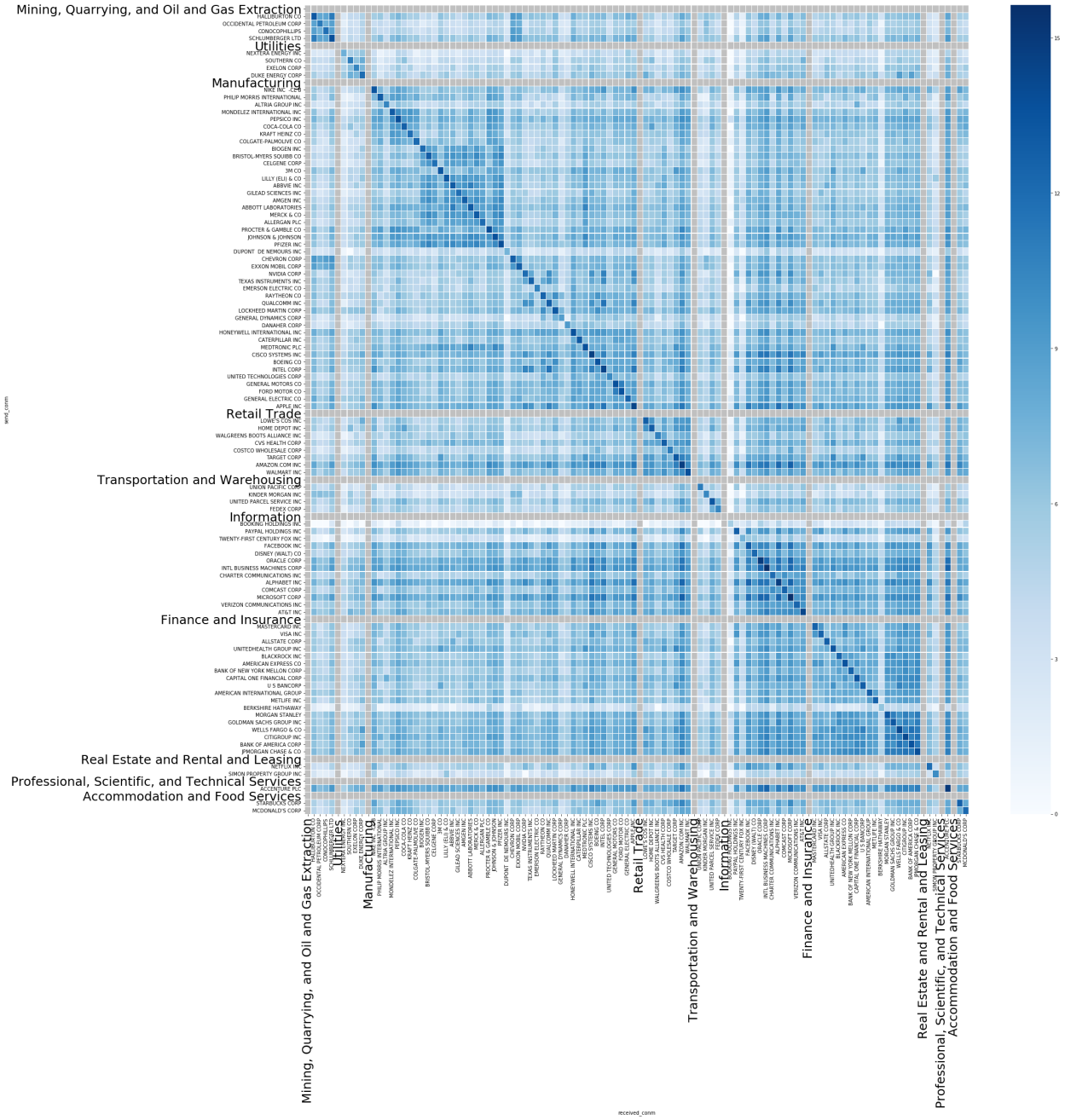
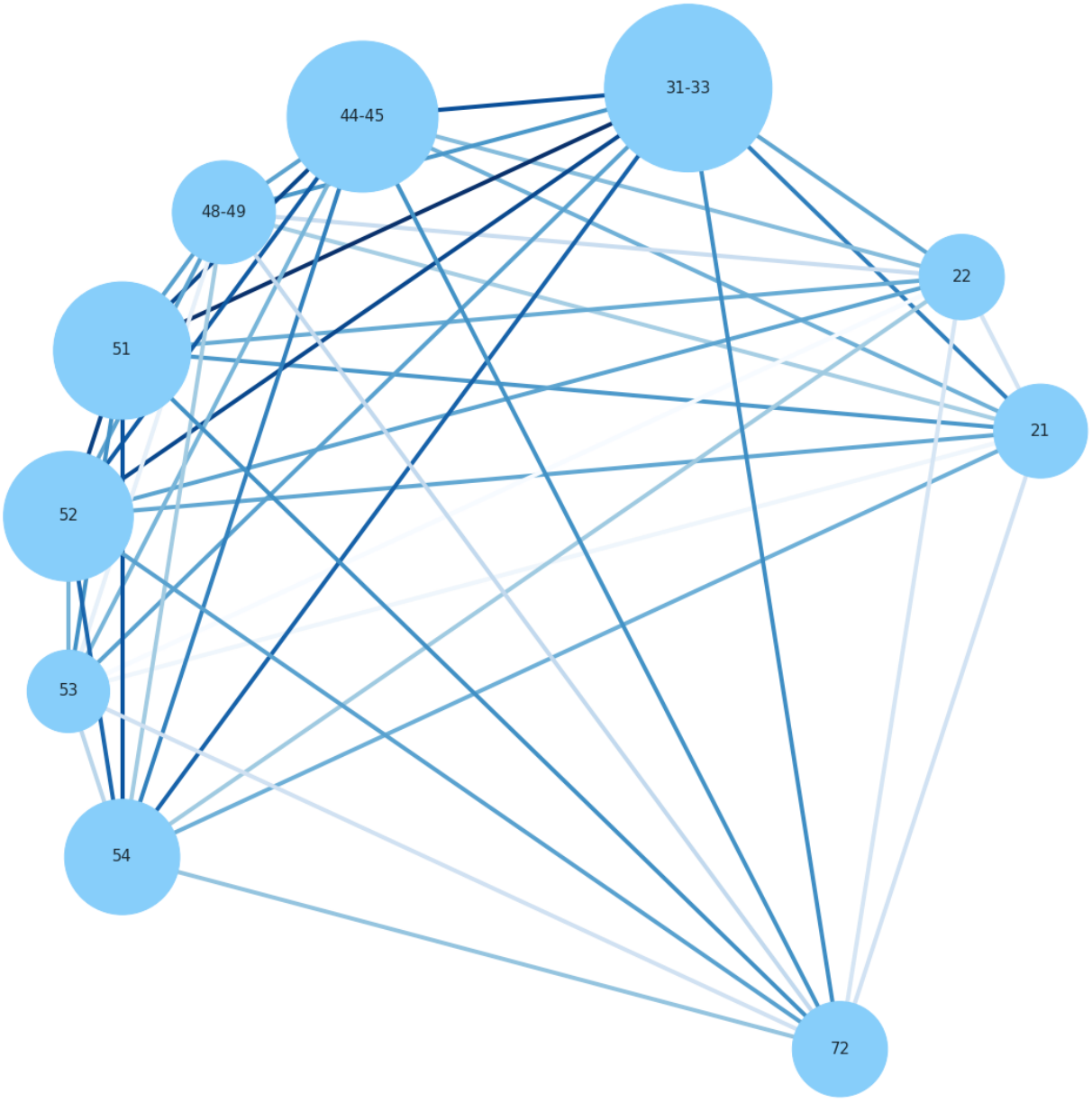


Figure 2 Industry Network Diagram for S&P 100 Firms



2-digit NAICS code	Industry Name
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing

42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
92	Public Administration
99	Missing

Figure 3 Industry Network Diagram for S&P 500 Firms

An explanation for the 2-digit industry codes is provided on the previous page (after **Figure 2**).

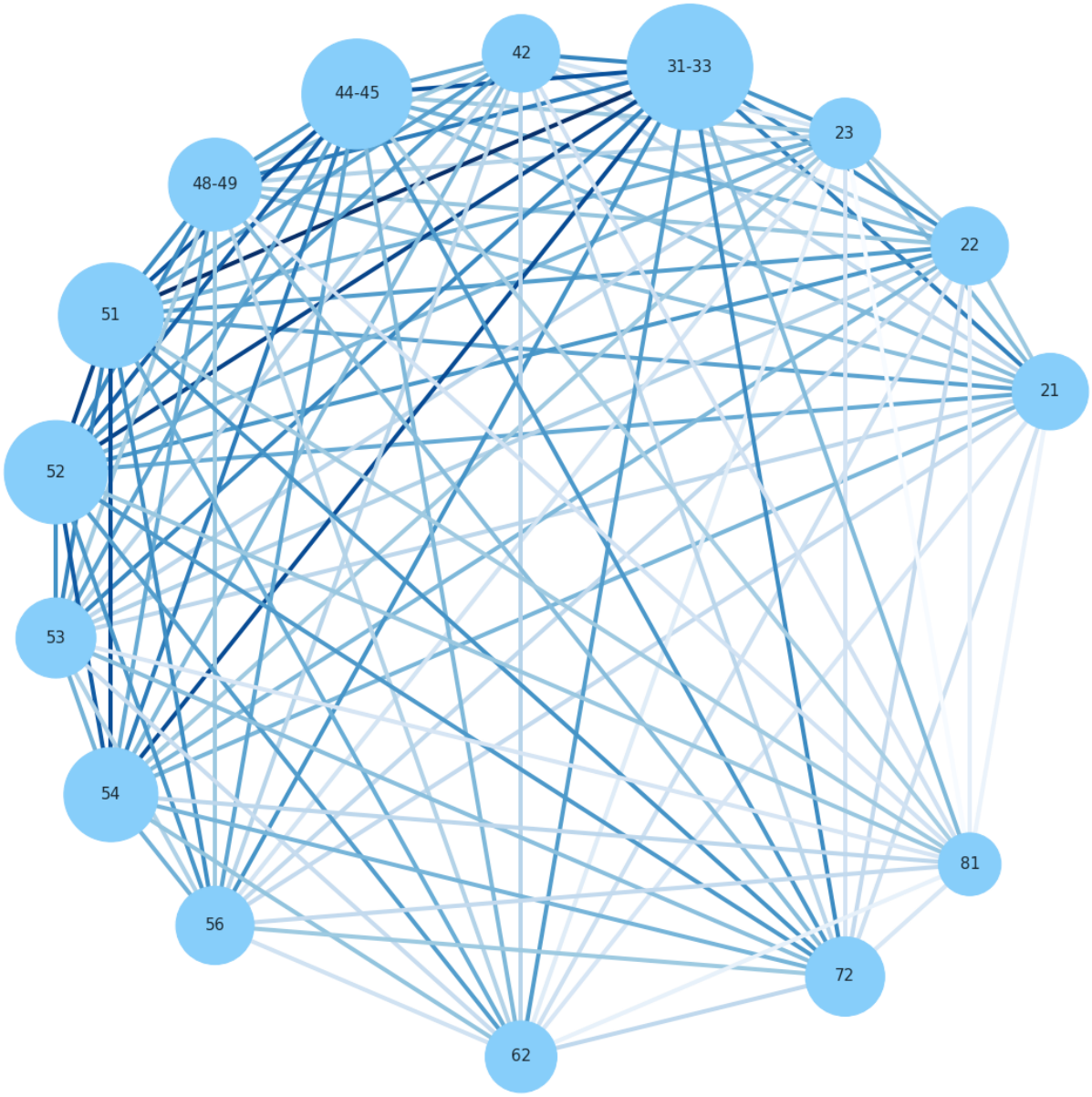


Figure 4 Industry Network Diagram for US Public Firms

An explanation for the 2-digit industry codes is provided on the previous page (after **Figure 2**).

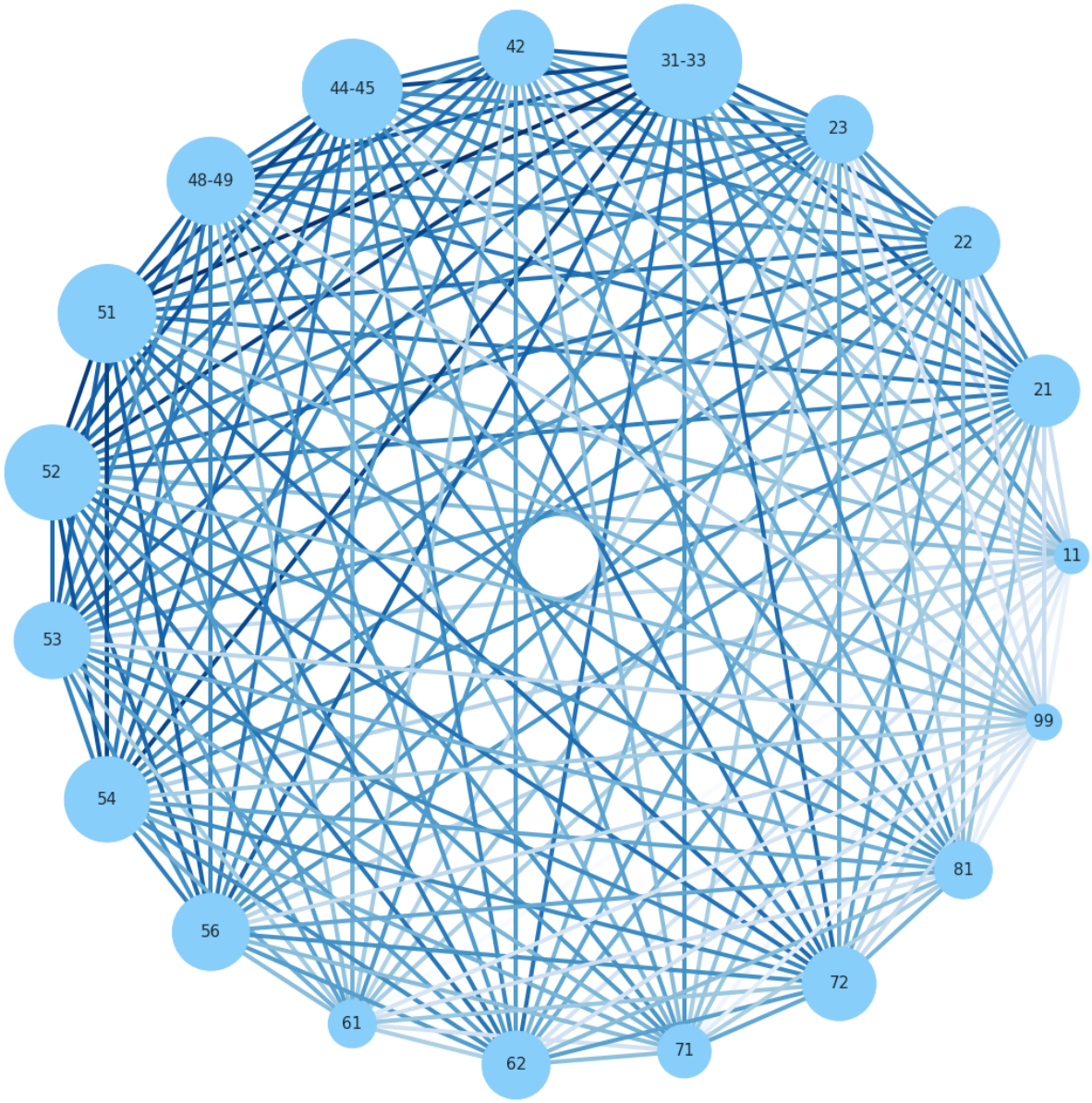


Table 4 Firm Social Capital and Productivity

	1				2			
Social Capital Measures	$SC_{i,t-1}$	$\ln(K_{it})$	$\ln(L_{it})$	R^2	$SC_{i,t-1}$	$\ln(K_{it})$	$\ln(L_{it})$	R^2
Eigenvector centrality	**8.484**	**0.274**	**0.731**	**0.877**	**2.759**	**0.182**	**0.608**	**0.353**
Total number of employee connections	**0.015**	**0.263**	**0.772**	**0.871**	**0.010**	**0.189**	**0.579**	**0.341**
Average number of connections	**0.055**	**0.277**	**0.762**	**0.875**	**0.011**	**0.187**	**0.607**	**0.350**
Total number of employee internal connections	**0.026**	**0.262**	**0.770**	**0.871**	**0.008**	**0.188**	**0.578**	**0.342**
Total number of employee connections with people who have the same education level	**0.085**	**0.270**	**0.738**	**0.877**	**0.015**	**0.183**	**0.613**	**0.357**
Total number of employee connections with people who have the same experience level	**0.098**	**0.273**	**0.737**	**0.876**	**0.011**	**0.185**	**0.606**	**0.352**
Total number of connections with people who have the same job functions	0.042	**0.273**	**0.741**	**0.876**	**0.009**	**0.185**	**0.606**	**0.352**
Total number of connections with people who have gone to the same school	**0.048**	**0.269**	**0.739**	**0.877**	**0.012**	**0.182**	**0.614**	**0.357**
Total number of connections with people who have worked at the same institution.	**0.044**	**0.272**	**0.740**	**0.876**	**0.012**	**0.185**	**0.603**	**0.351**
Industry Fixed Effects	Yes				No			
Year Fixed Effects	Yes				Yes			
Firm Fixed Effects	No				Yes			

Table 5 The Interactive Effect of Social Capital and Labor on Firm Productivity

	1	2	3	4	5
Social Capital Measures	$SC_{i,t-1}$	$SC_{i,t-1} * L_{it}$	$\ln(K_{it})$	$\ln(L_{it})$	R^2
Eigenvector centrality	**2.925**	**2.086**	**0.182**	**0.608**	**0.353**
Total number of employee connections	**0.010**	**-0.005**	**0.189**	**0.578**	**0.342**
Average number of connections	**0.011**	-0.005	**0.187**	**0.606**	**0.351**
Total number of employee internal connections	**0.078**	-0.048	**0.188**	**0.578**	**0.342**
Total number of employee connections with people who have the same education level	**0.015**	-0.004	**0.182**	**0.614**	**0.357**
Total number of employee connections with people who have the same experience level	**0.011**	-0.003	**0.185**	**0.607**	**0.353**
Total number of connections with people who have the same job functions	**0.009**	-0.003	**0.185**	**0.607**	**0.353**
Total number of connections with people who have gone to the same school	**0.012**	-0.001	**0.181**	**0.615**	**0.357**
Total number of connections with people who have worked at the same institution	**0.012**	-0.005	**0.184**	**0.603**	**0.352**
Industry Fixed Effects	No				
Year Fixed Effects	Yes				
Firm Fixed Effects	Yes				

Appendix

Figure A1 Connection heat map for S&P 100 firms in 2017

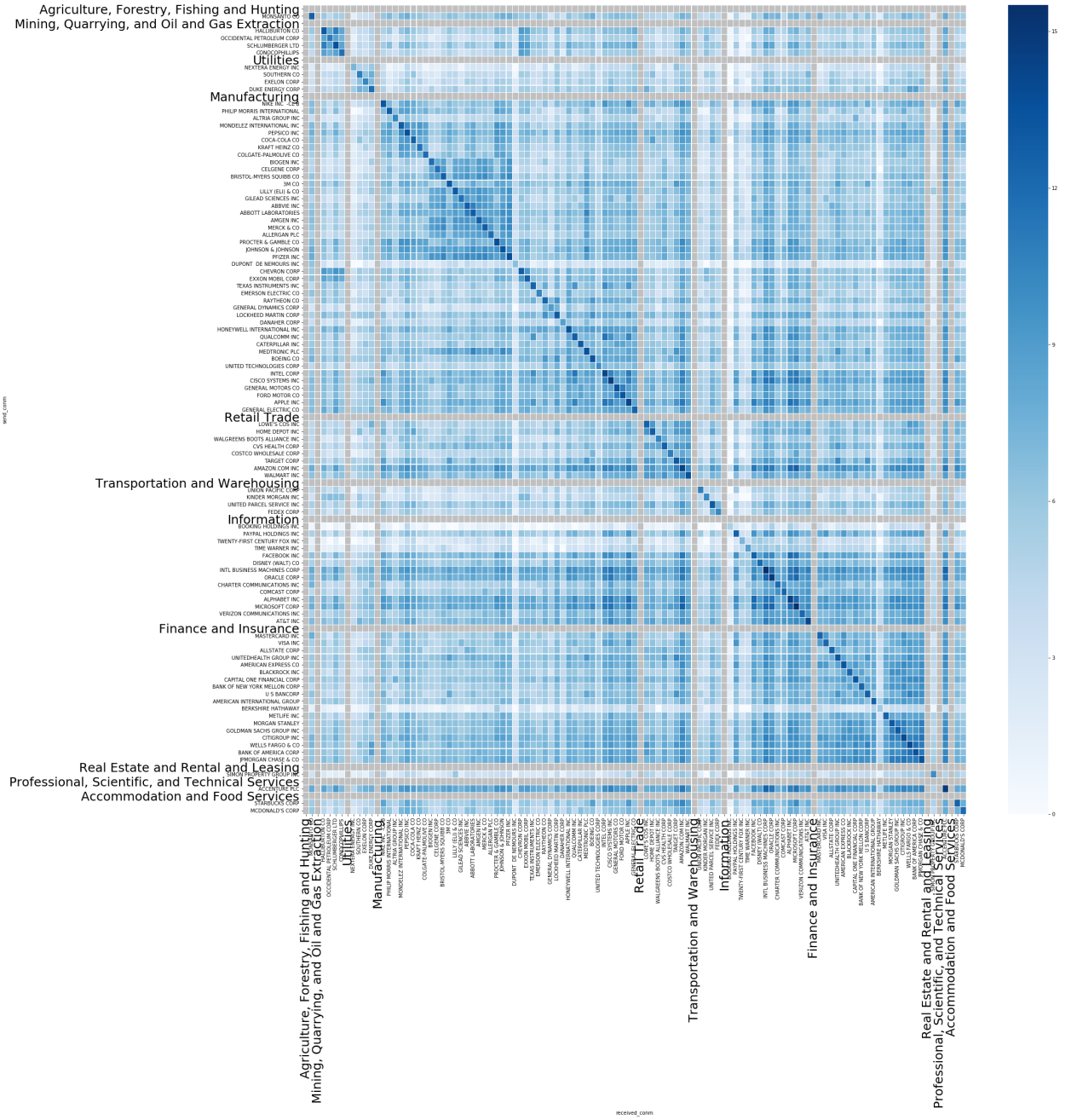


Figure A2 Connection heat map for S&P 100 firms in 2016

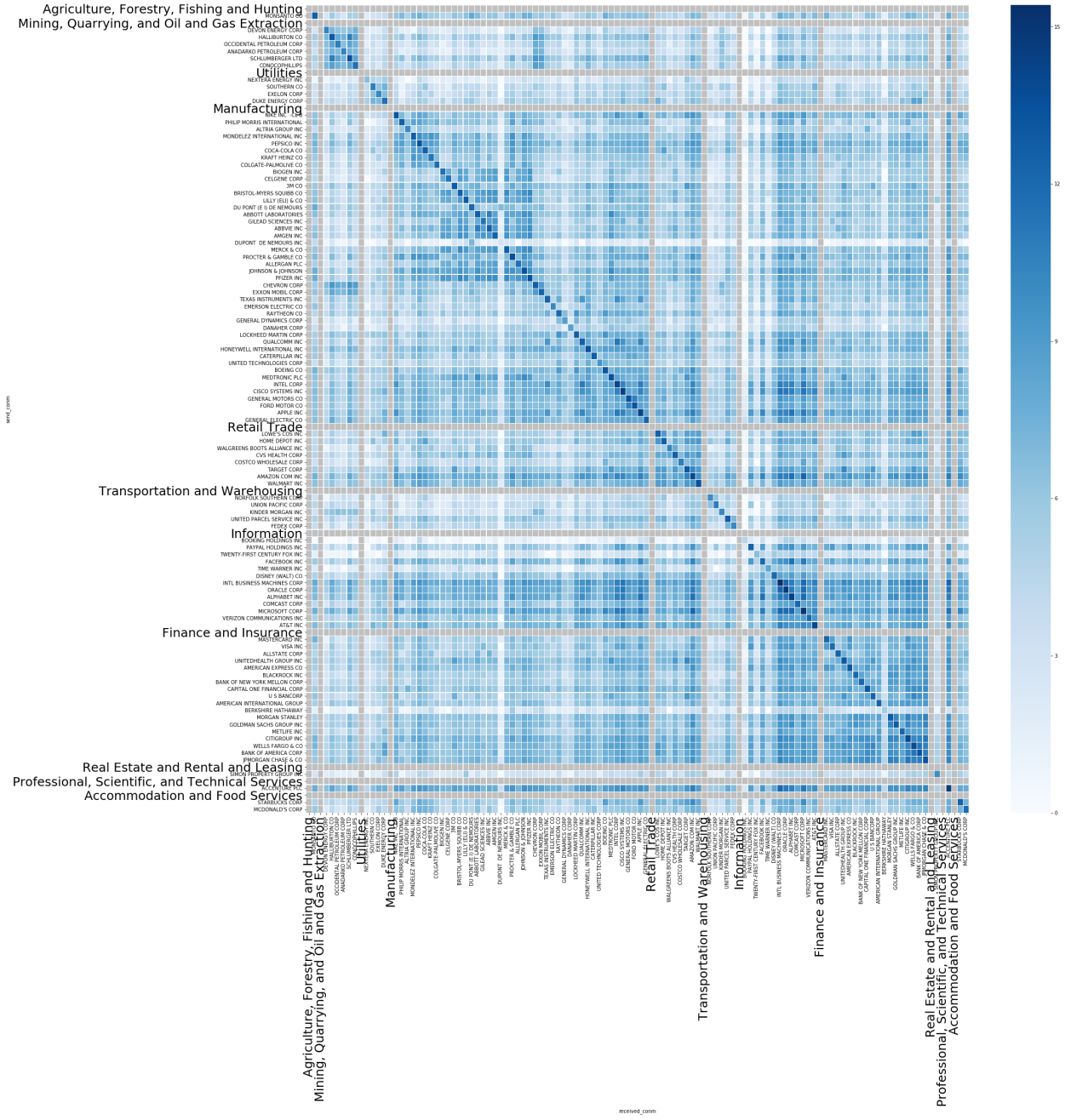


Figure A3 Connection heat map for S&P 100 firms in 2015

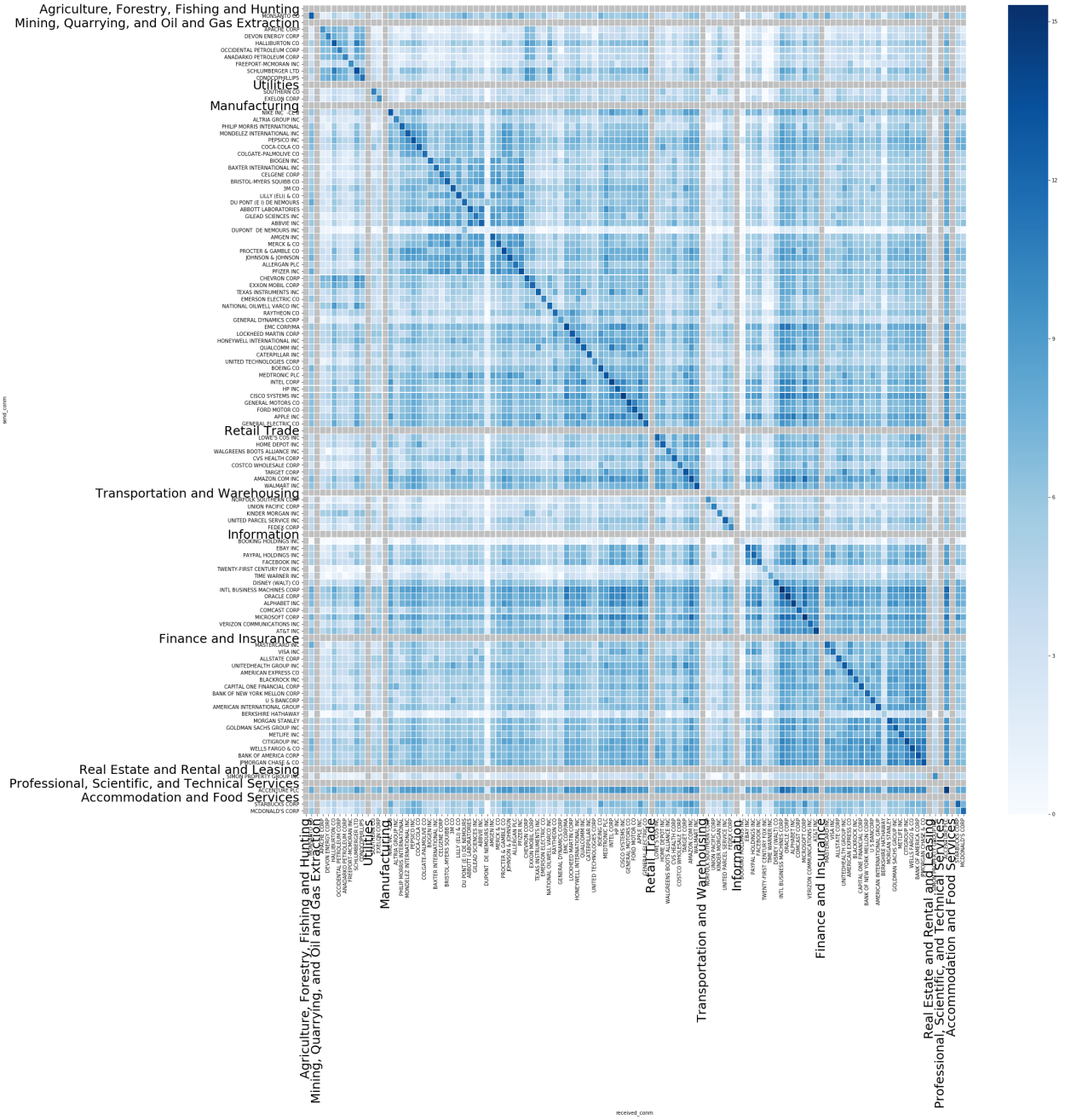


Figure A4 Connection heat map for S&P 100 firms in 2014

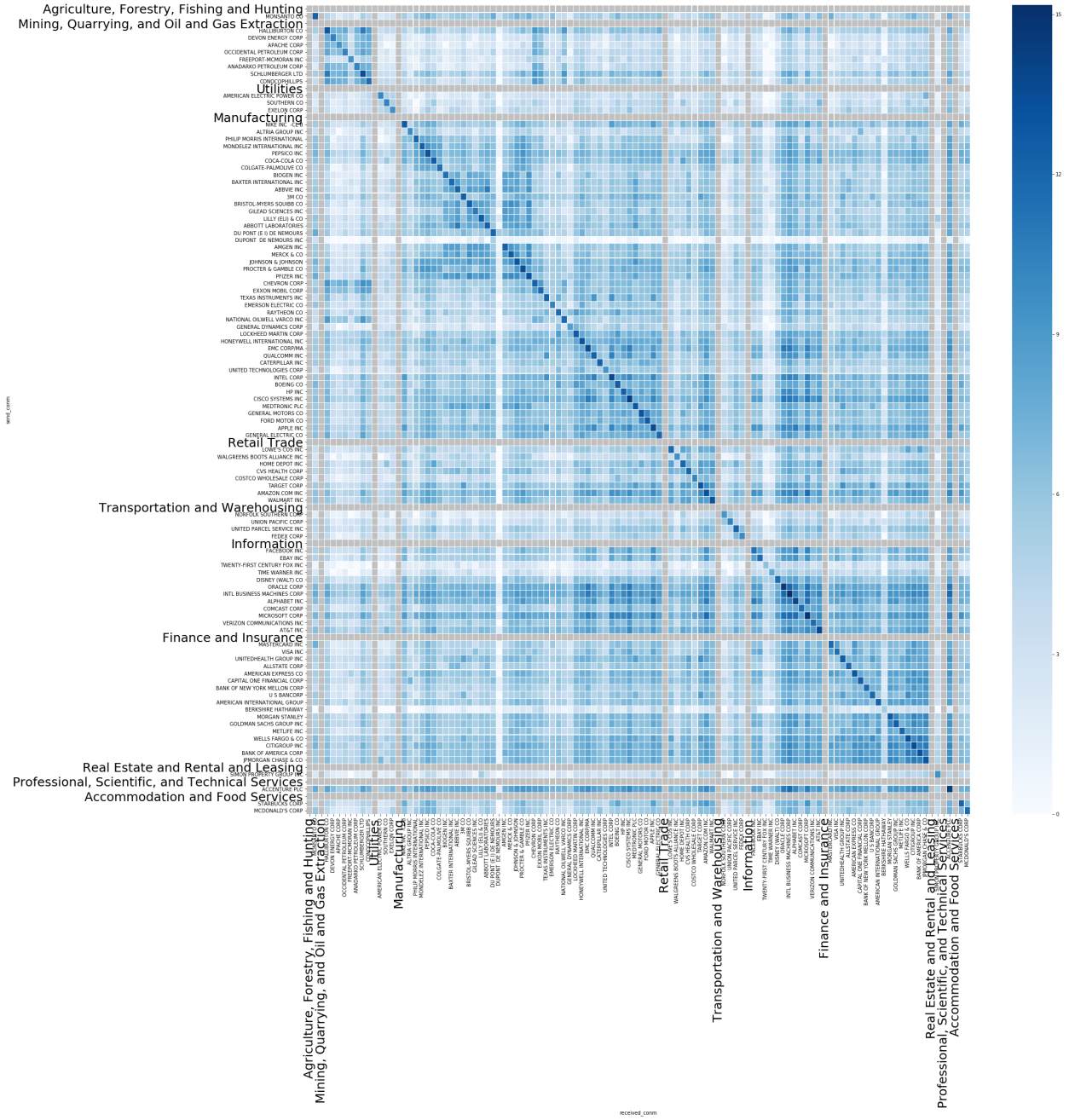


Figure A5 Connection heat map for S&P 100 firms in 2013

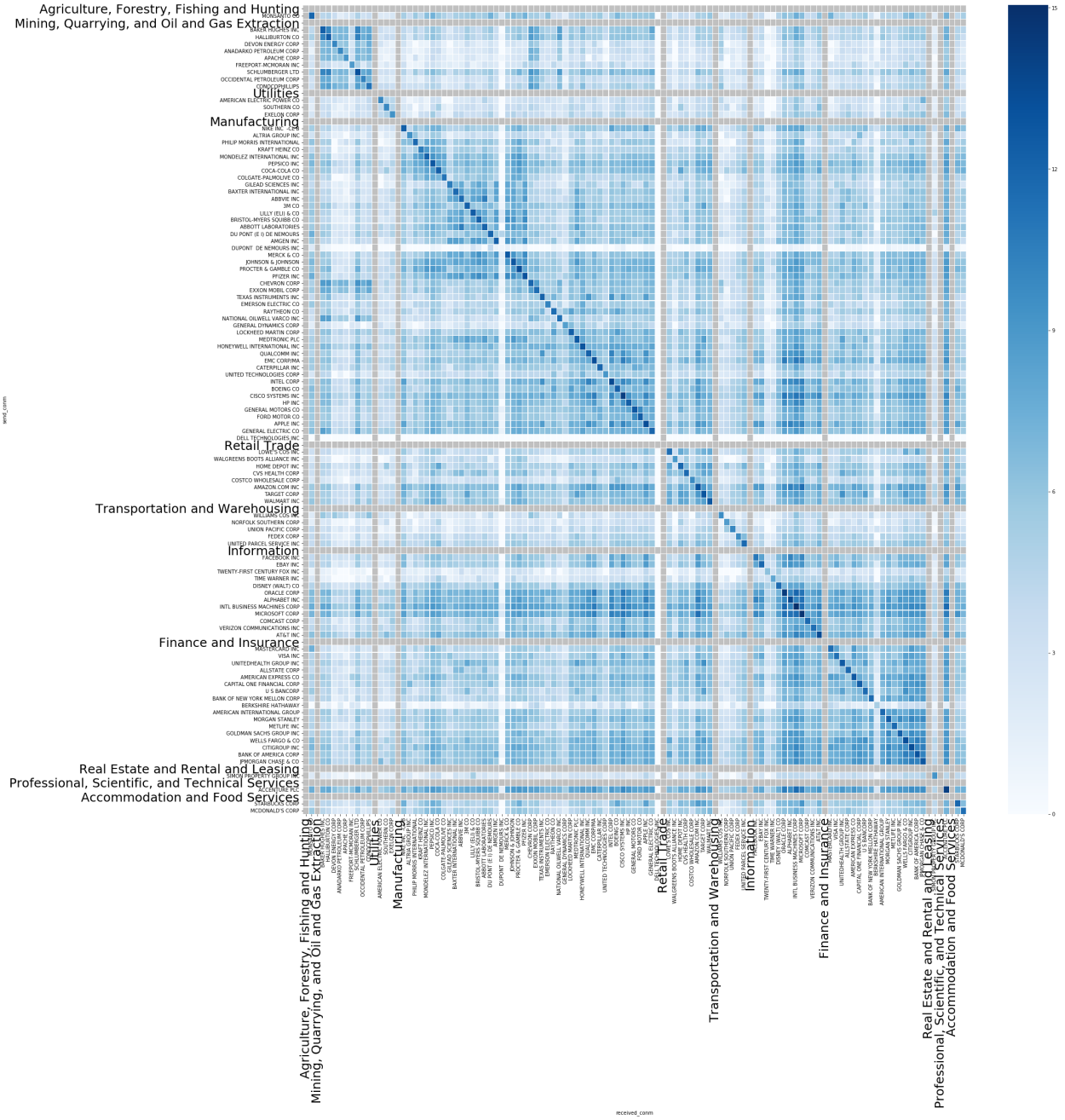


Figure A6 Connection heat map for S&P 100 firms in 2012

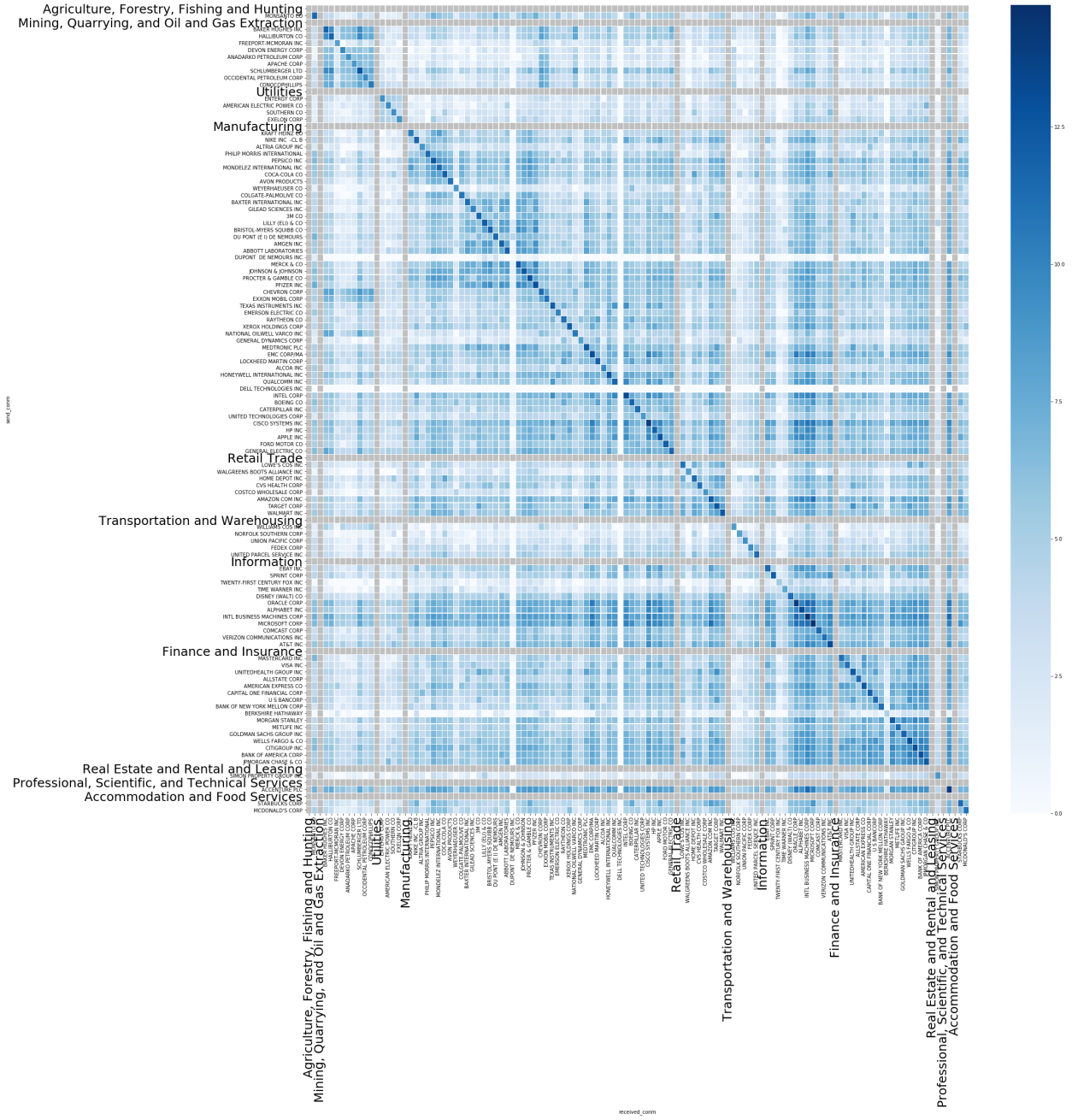


Figure A7 Connection heat map for S&P 100 firms in 2011

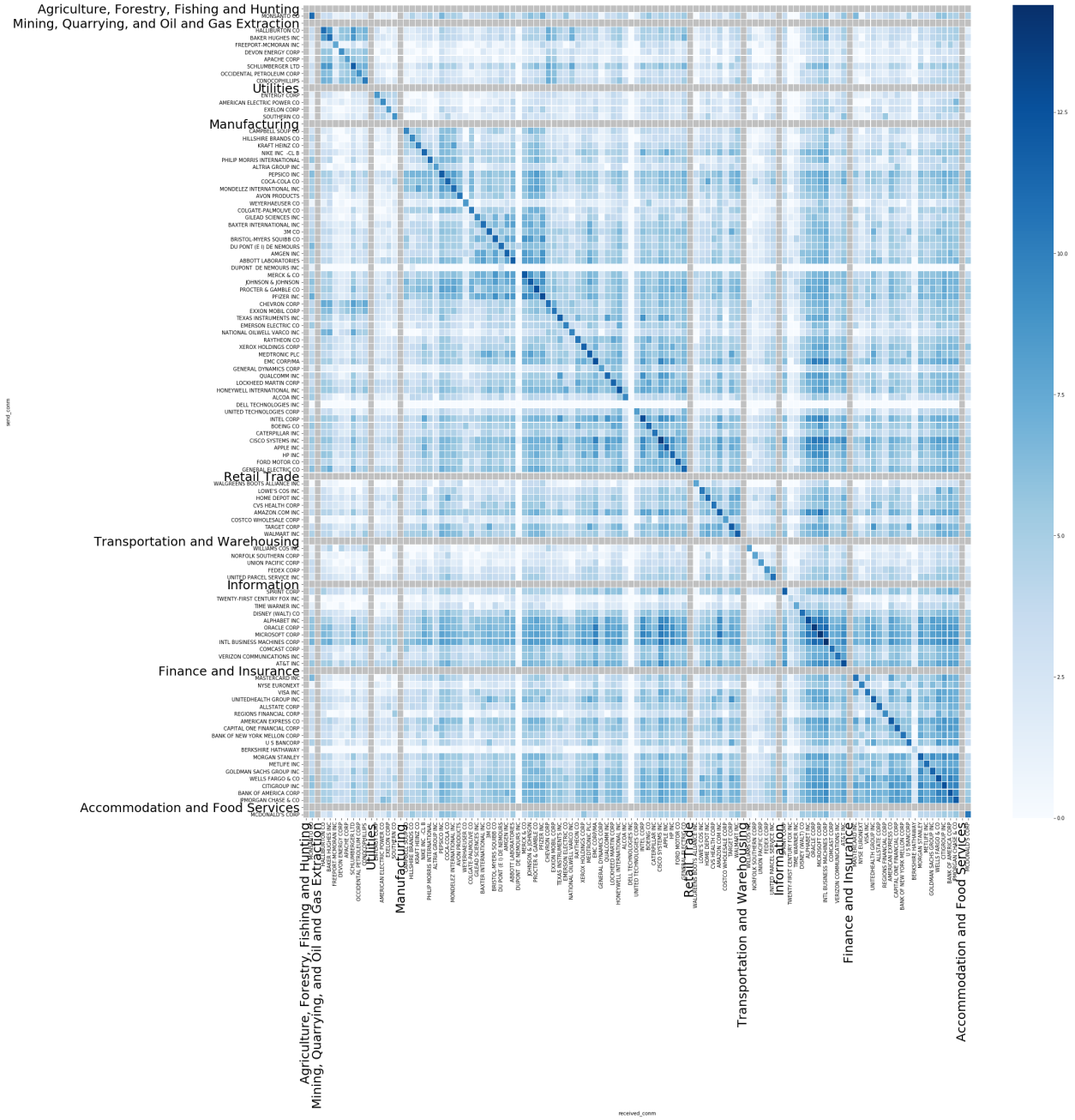


Figure A8 Connection heat map for S&P 100 firms in 2010

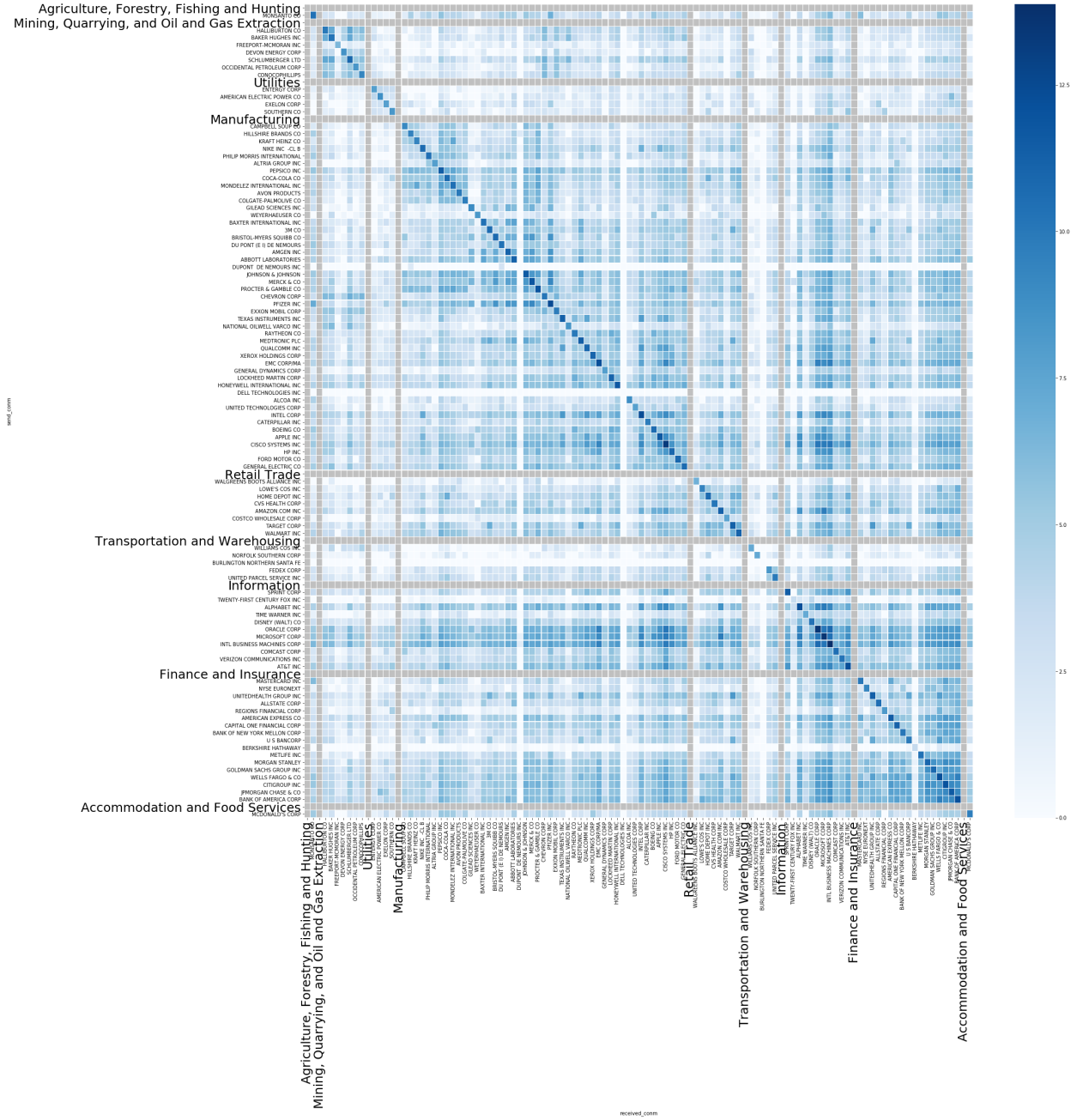


Figure A9 Connection heat map for S&P 100 firms in 2009

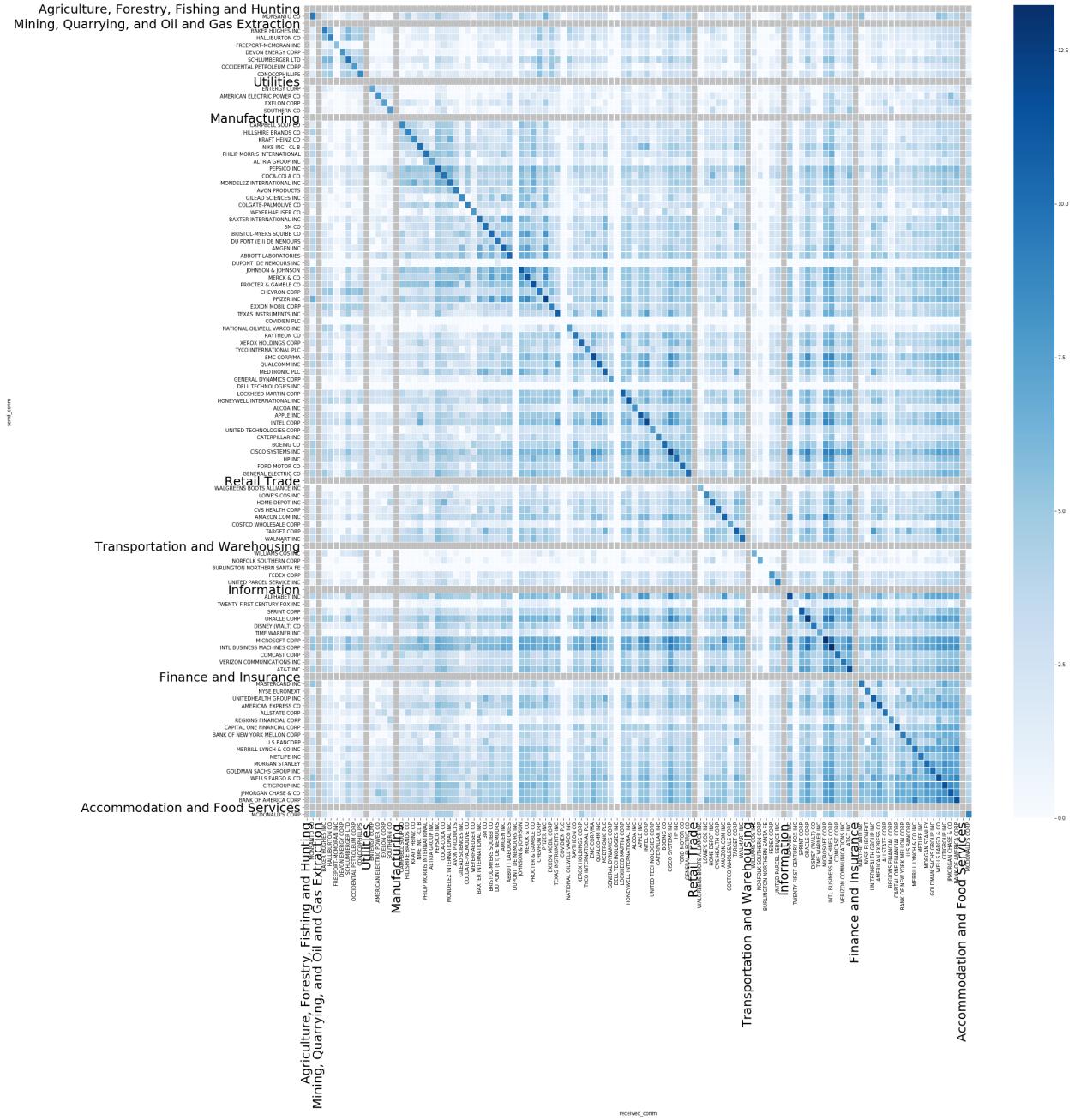


Figure A10 Connection heat map for S&P 100 firms in 2008

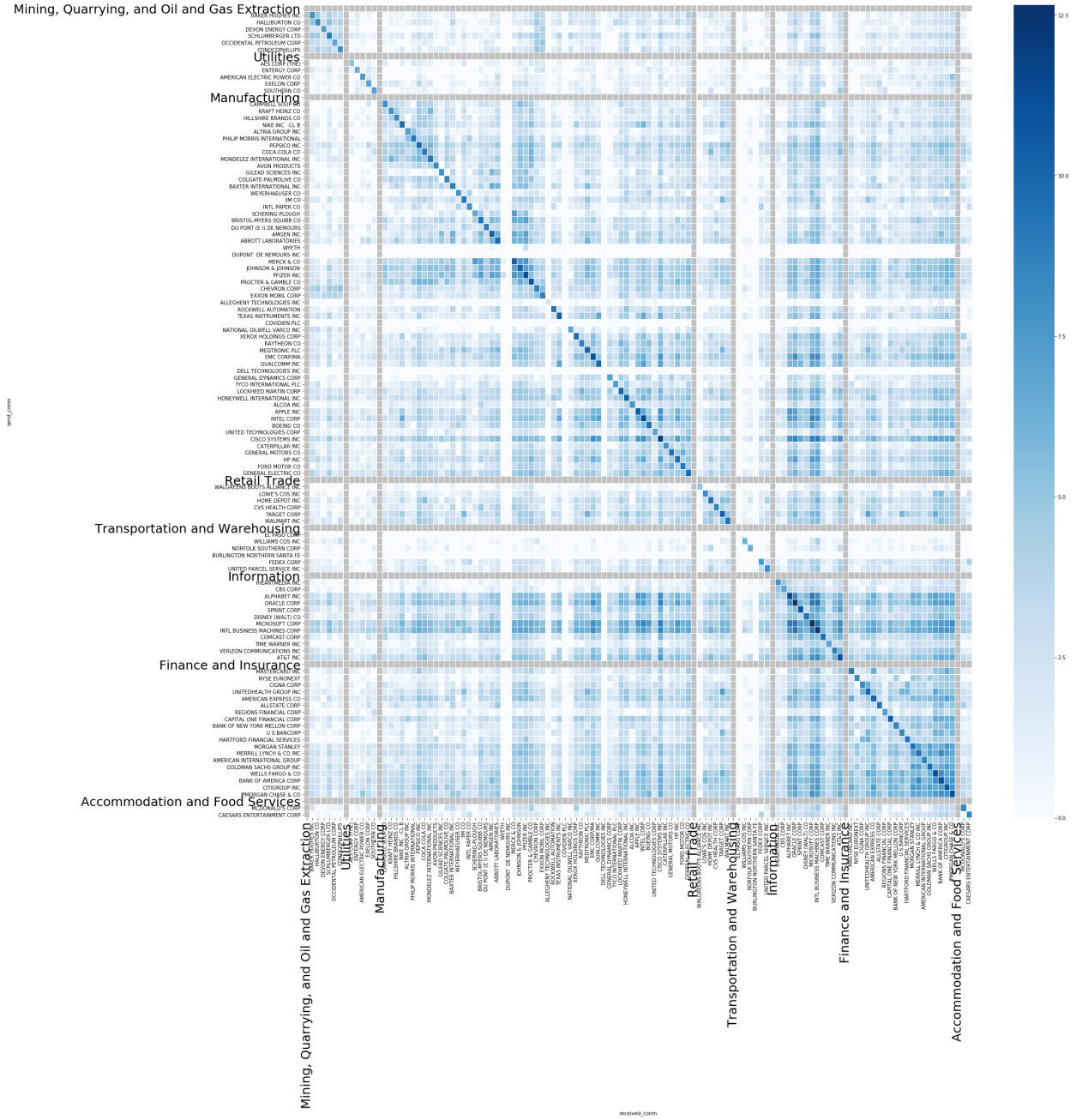


Figure A11 Connection heat map for S&P 100 firms in 2007

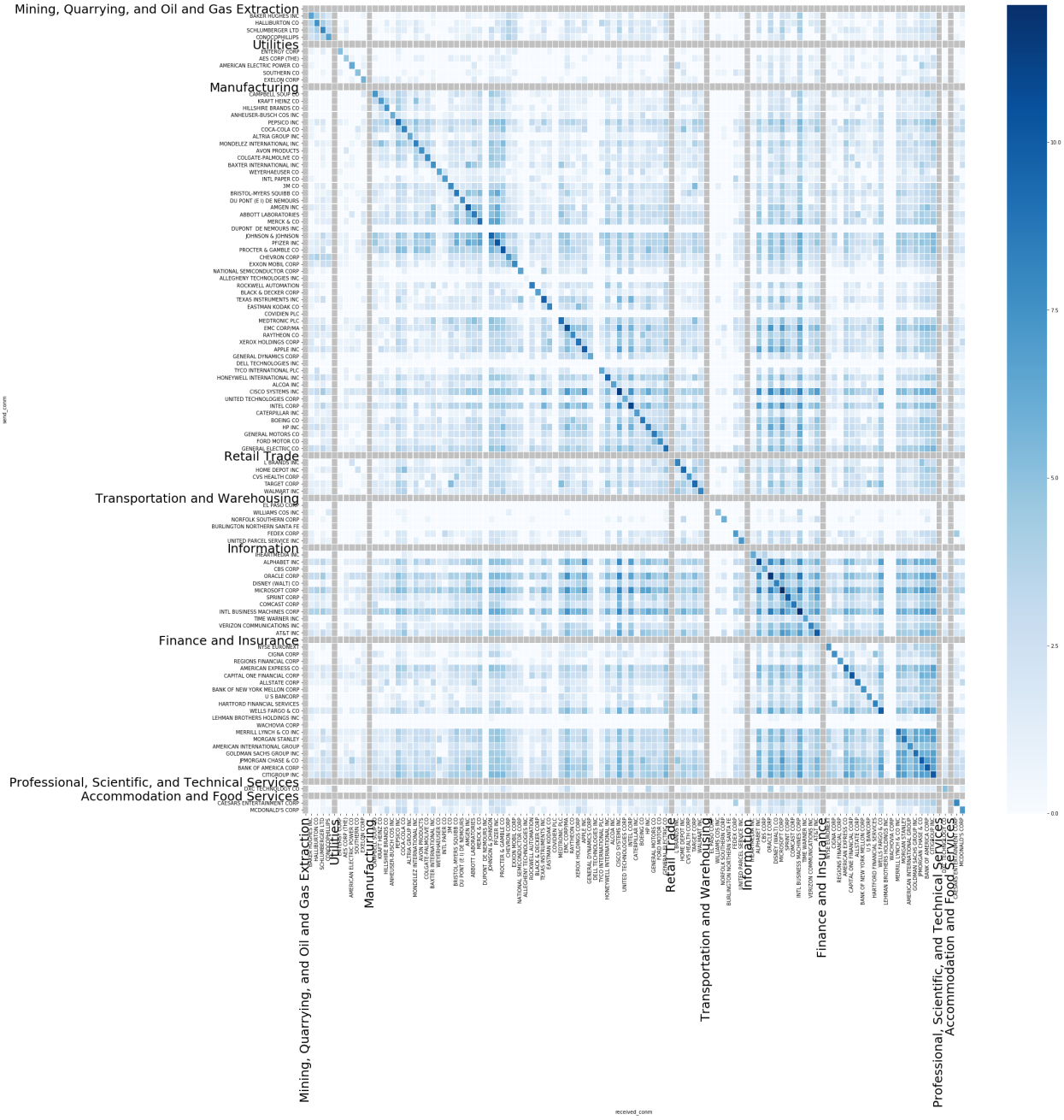


Figure A12 Connection heat map for S&P 100 firms in 2006

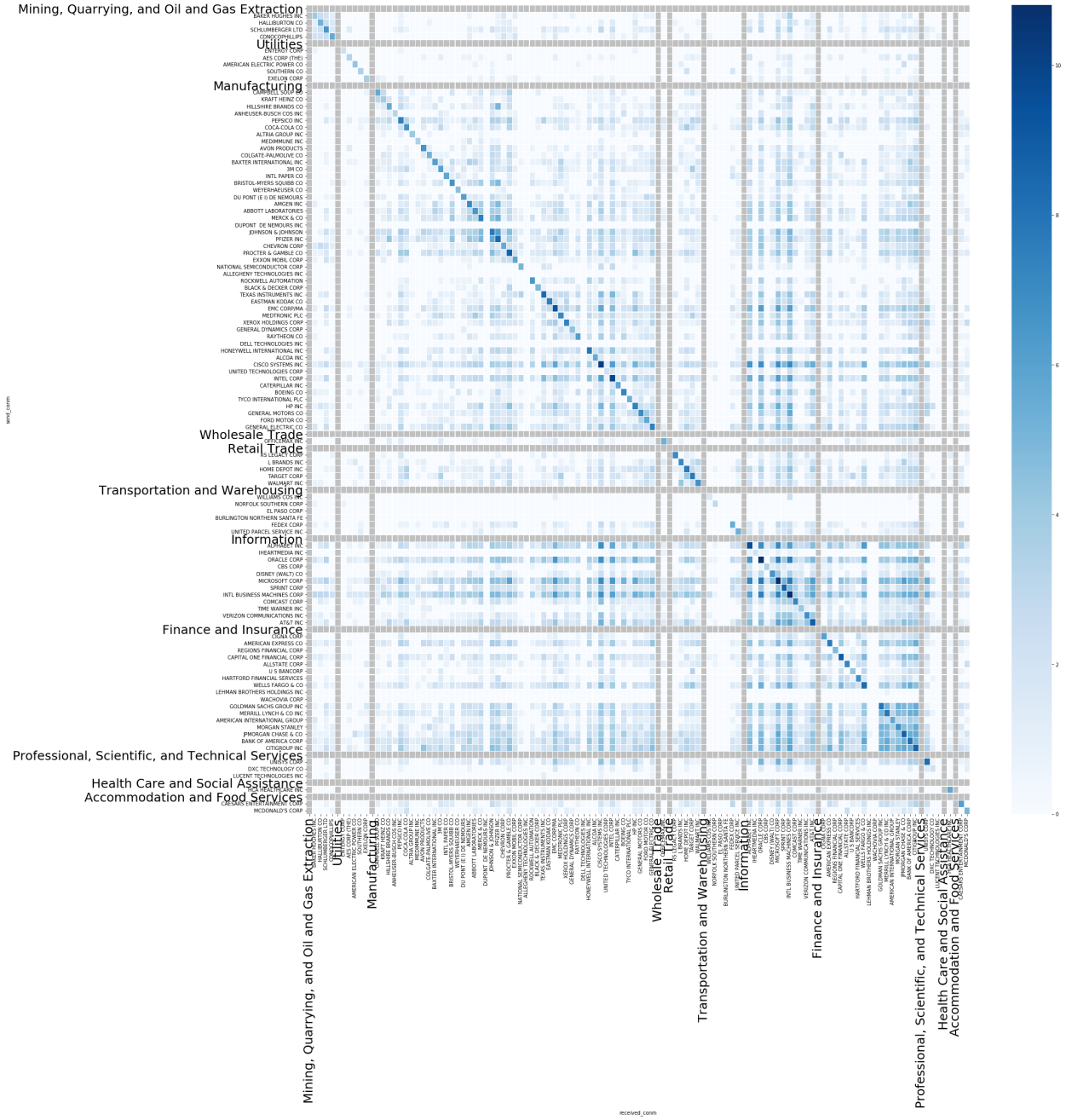


Figure A13 Connection heat map for S&P 100 firms in 2005

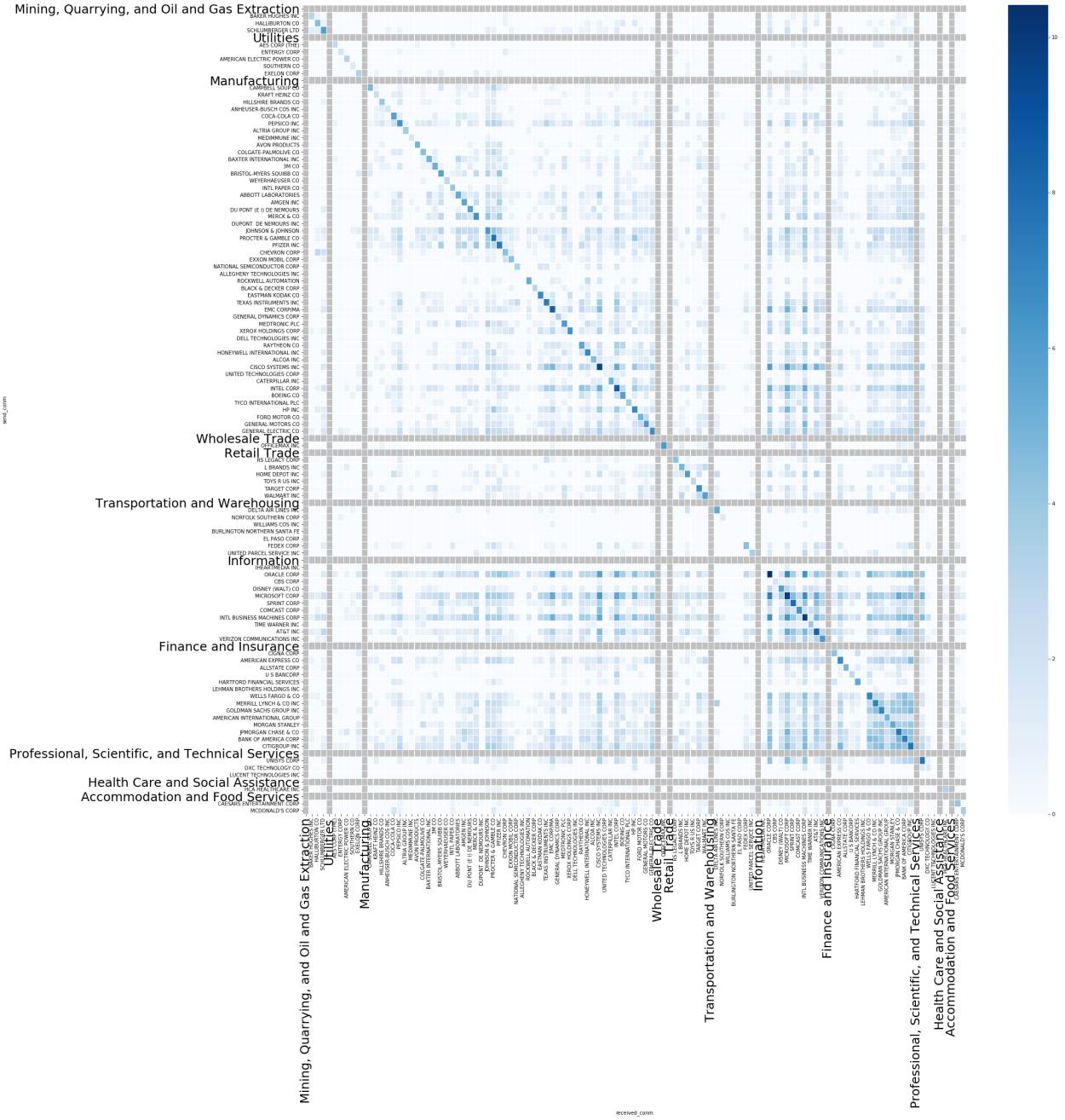


Figure A14 Connection heat map for S&P 100 firms in 2004

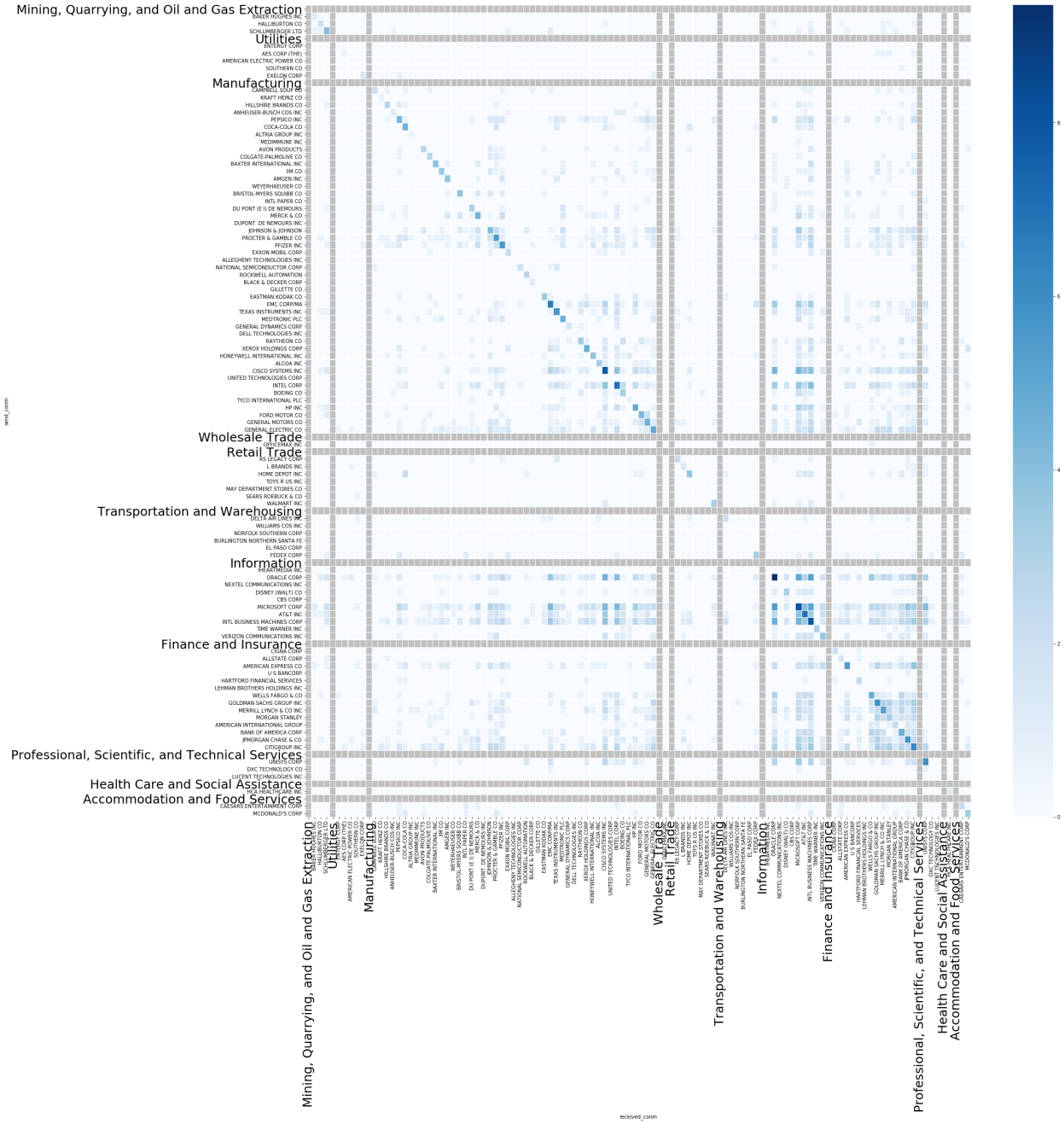


Figure A15 Industry Network Diagram for S&P 100 Firms in 2017

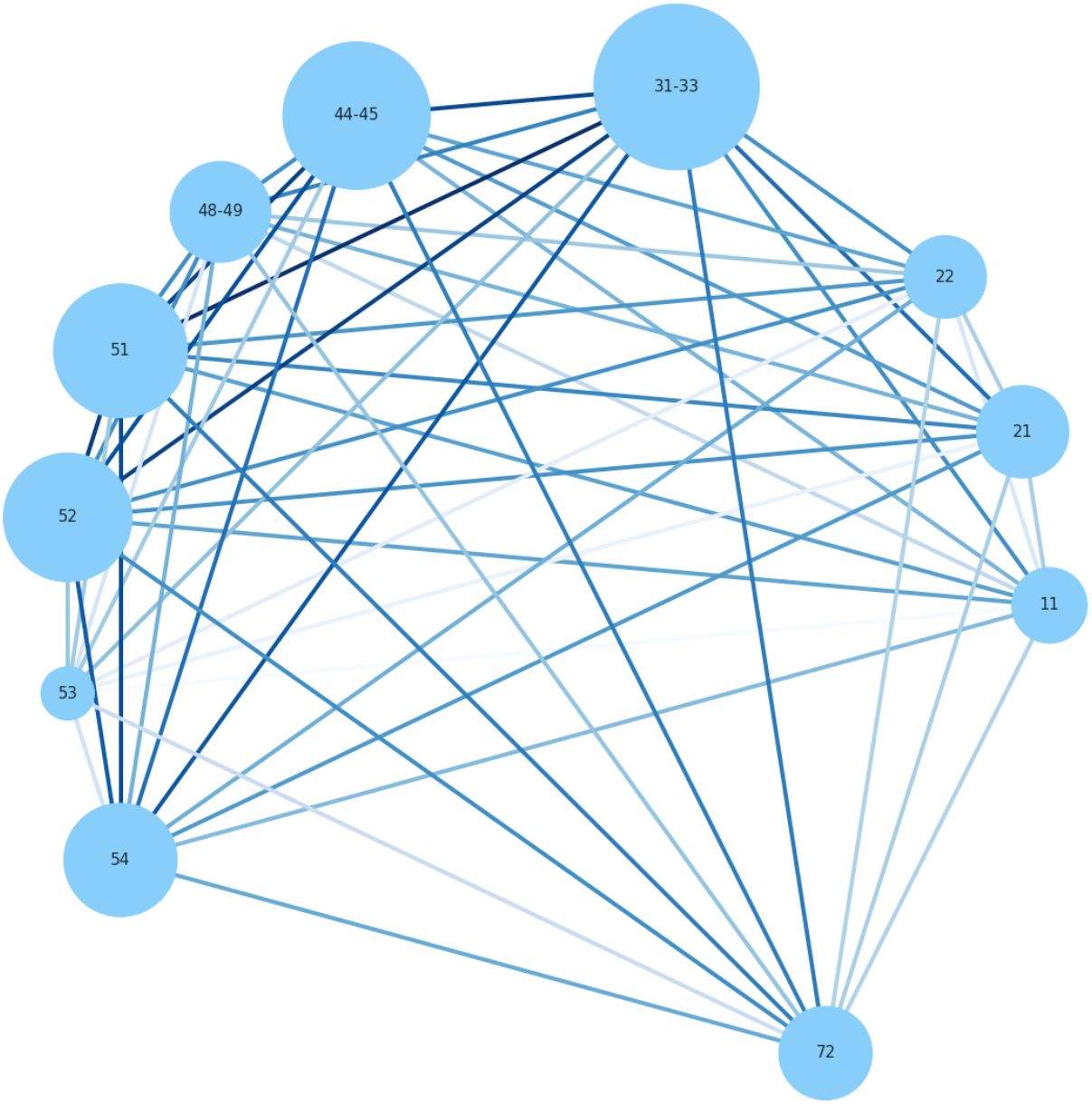


Figure A16 Industry Network Diagram for S&P 500 Firms in 2017

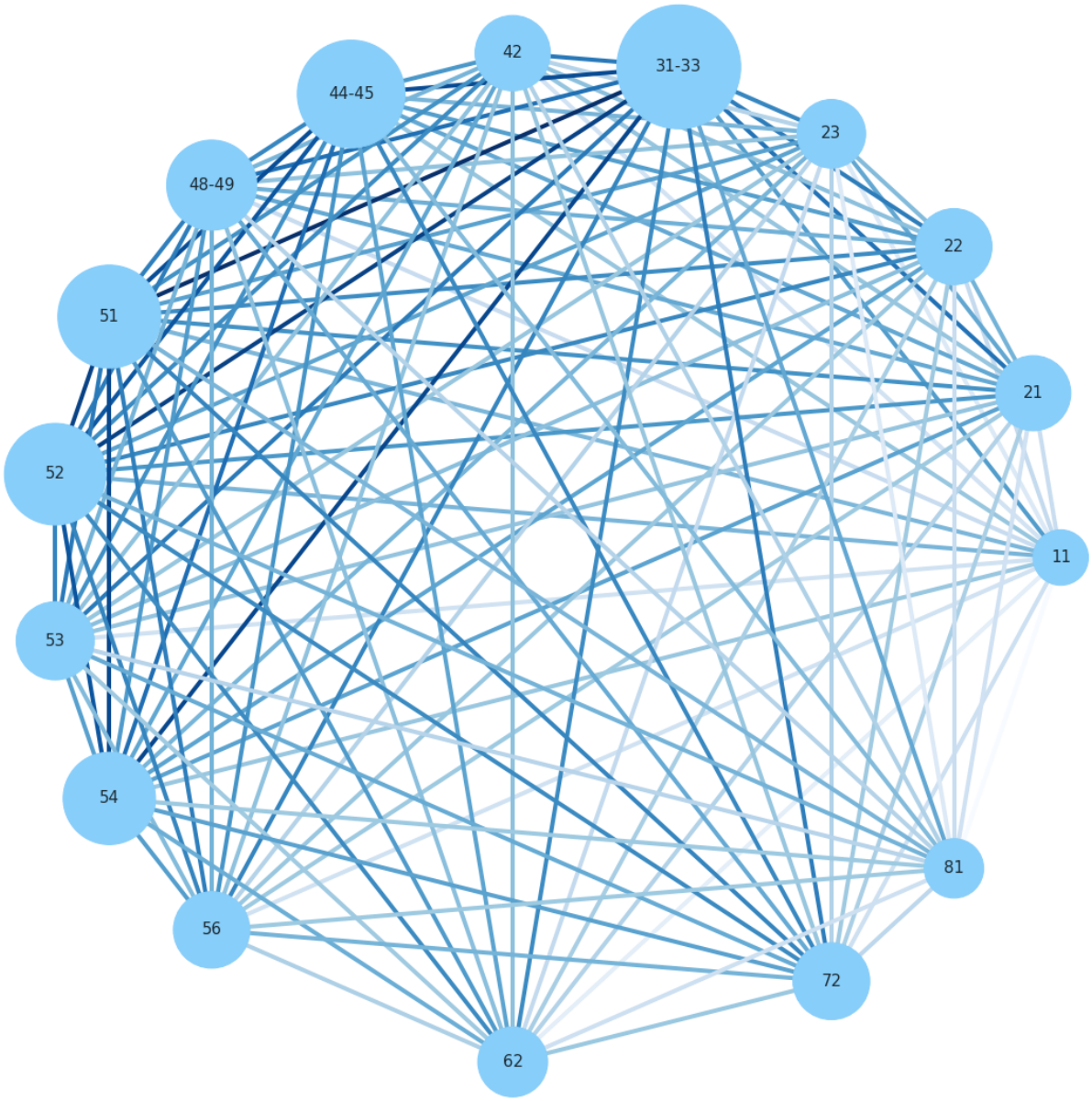


Figure A17 Industry Network Diagram for US Public Firms in 2017

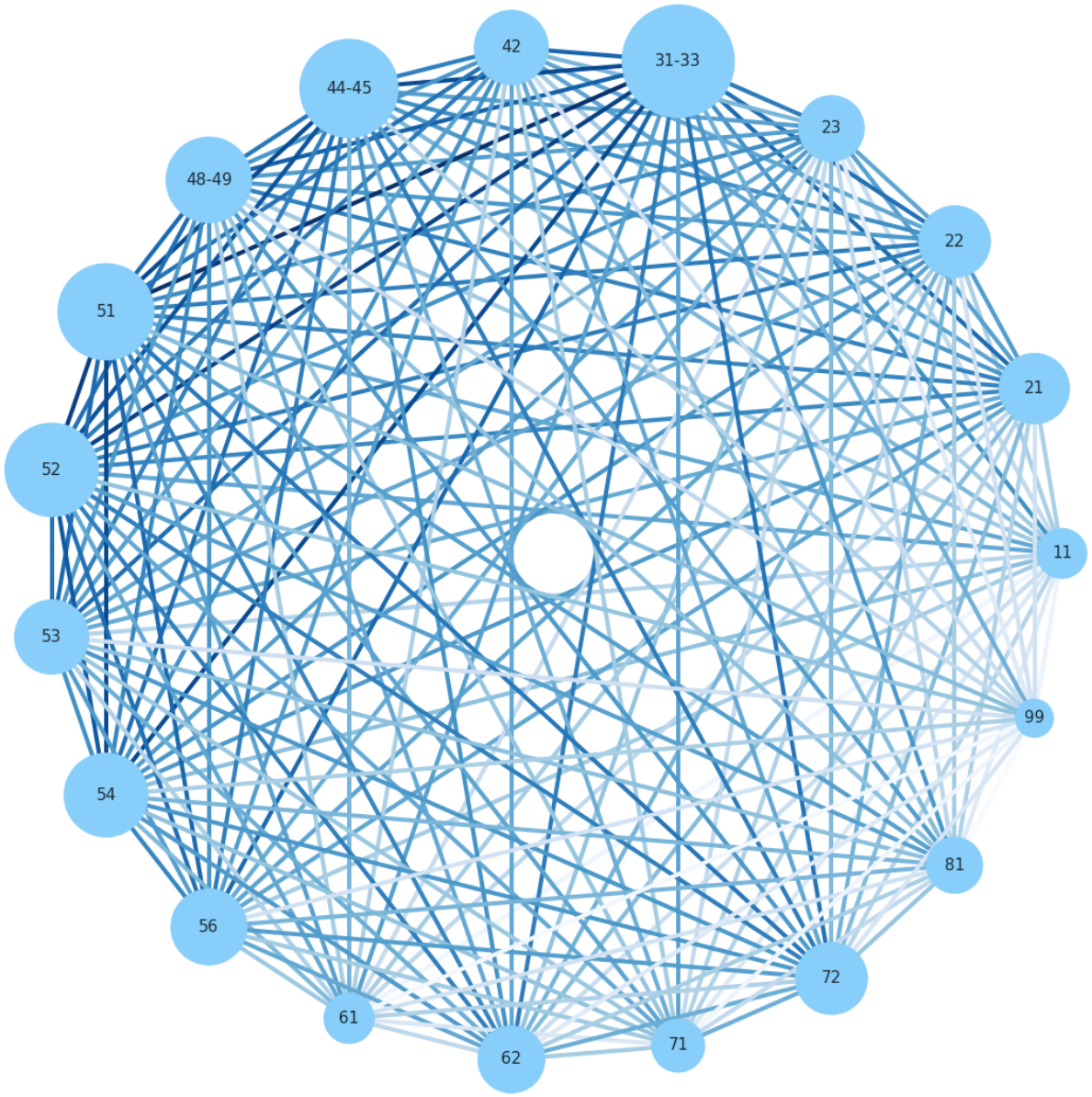


Figure A18 Industry Network Diagram for S&P 100 Firms in 2016

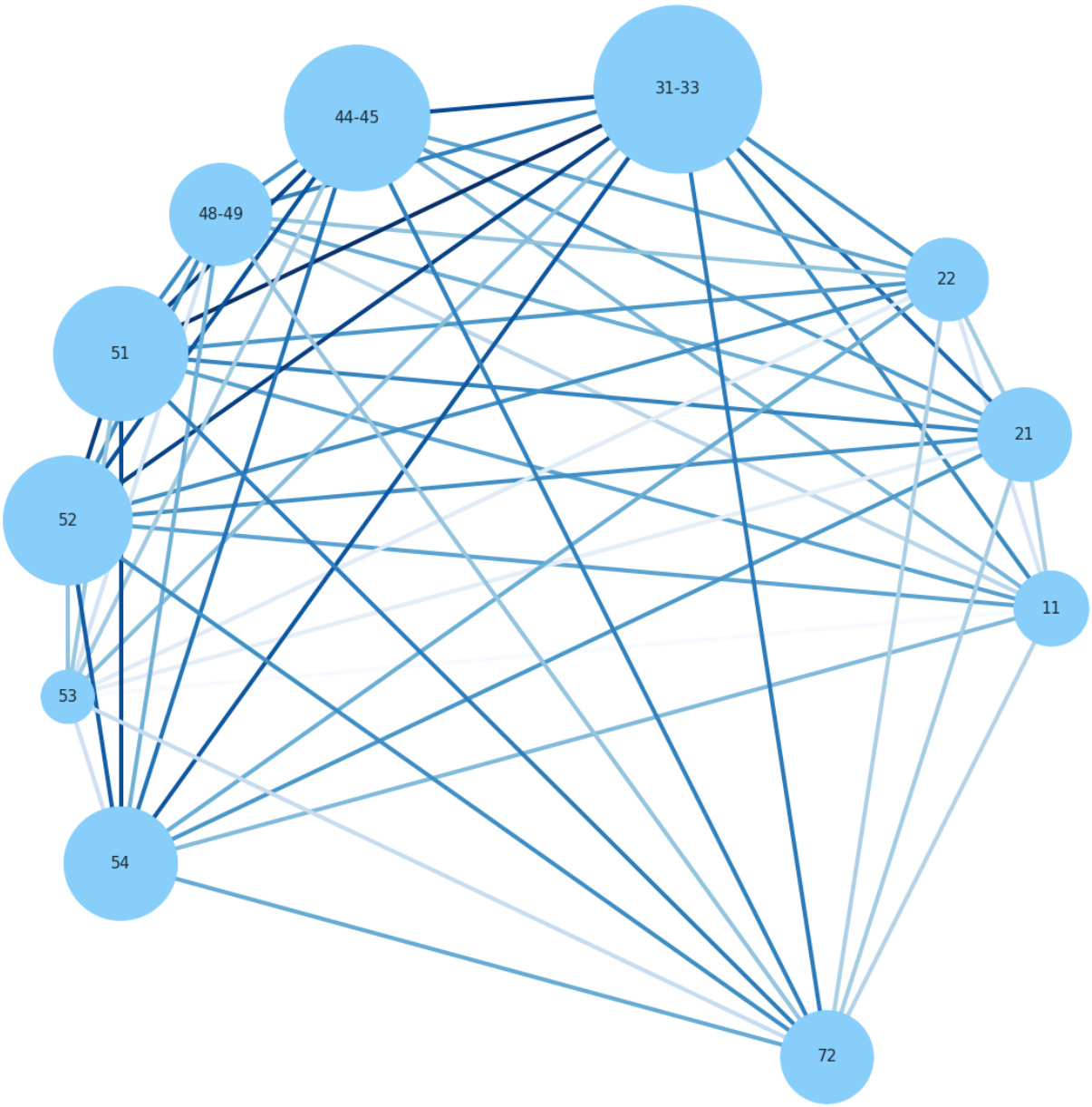


Figure A19 Industry Network Diagram for S&P 500 Firms in 2016

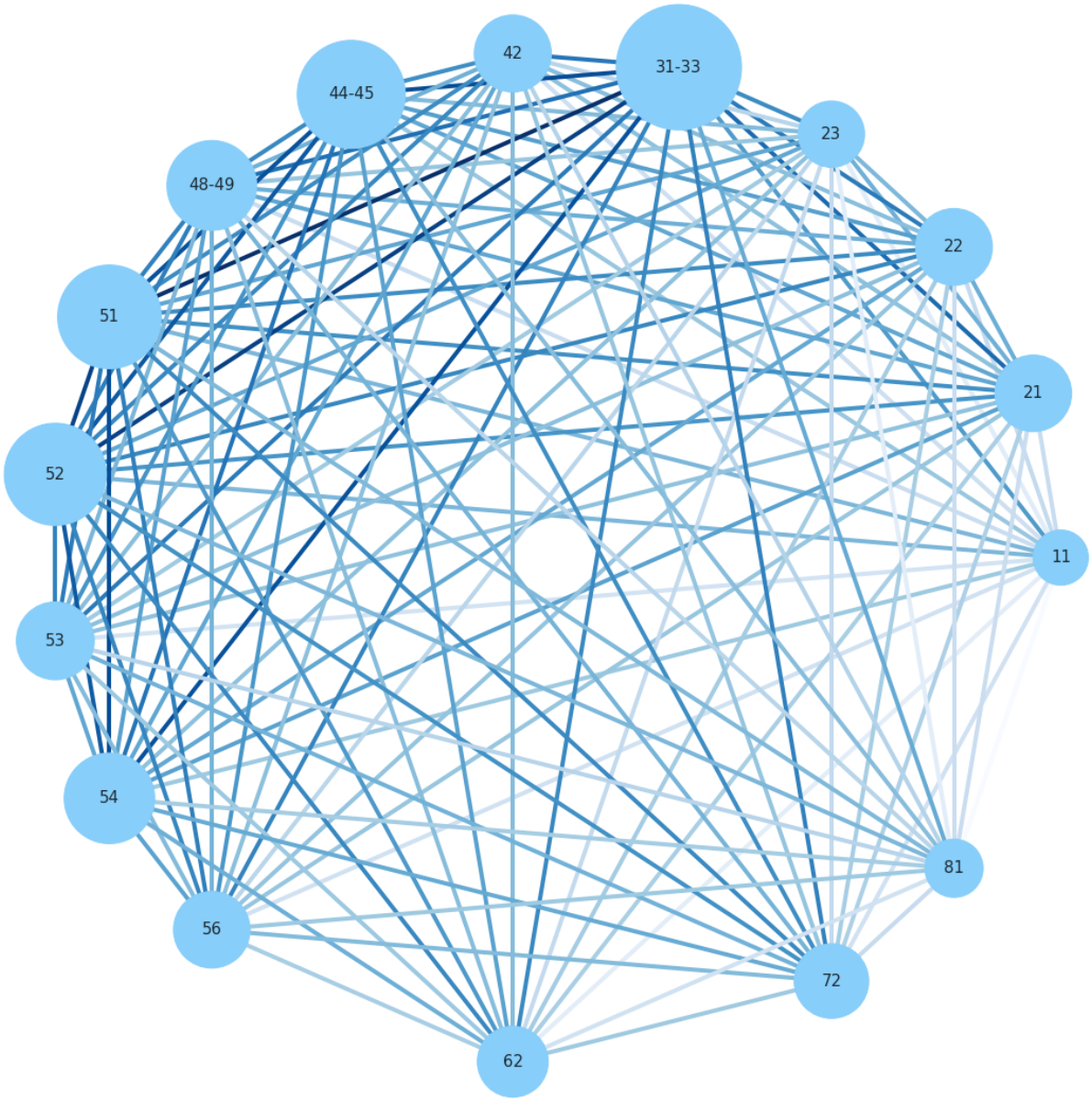


Figure A20 Industry Network Diagram for US Public Firms in 2016

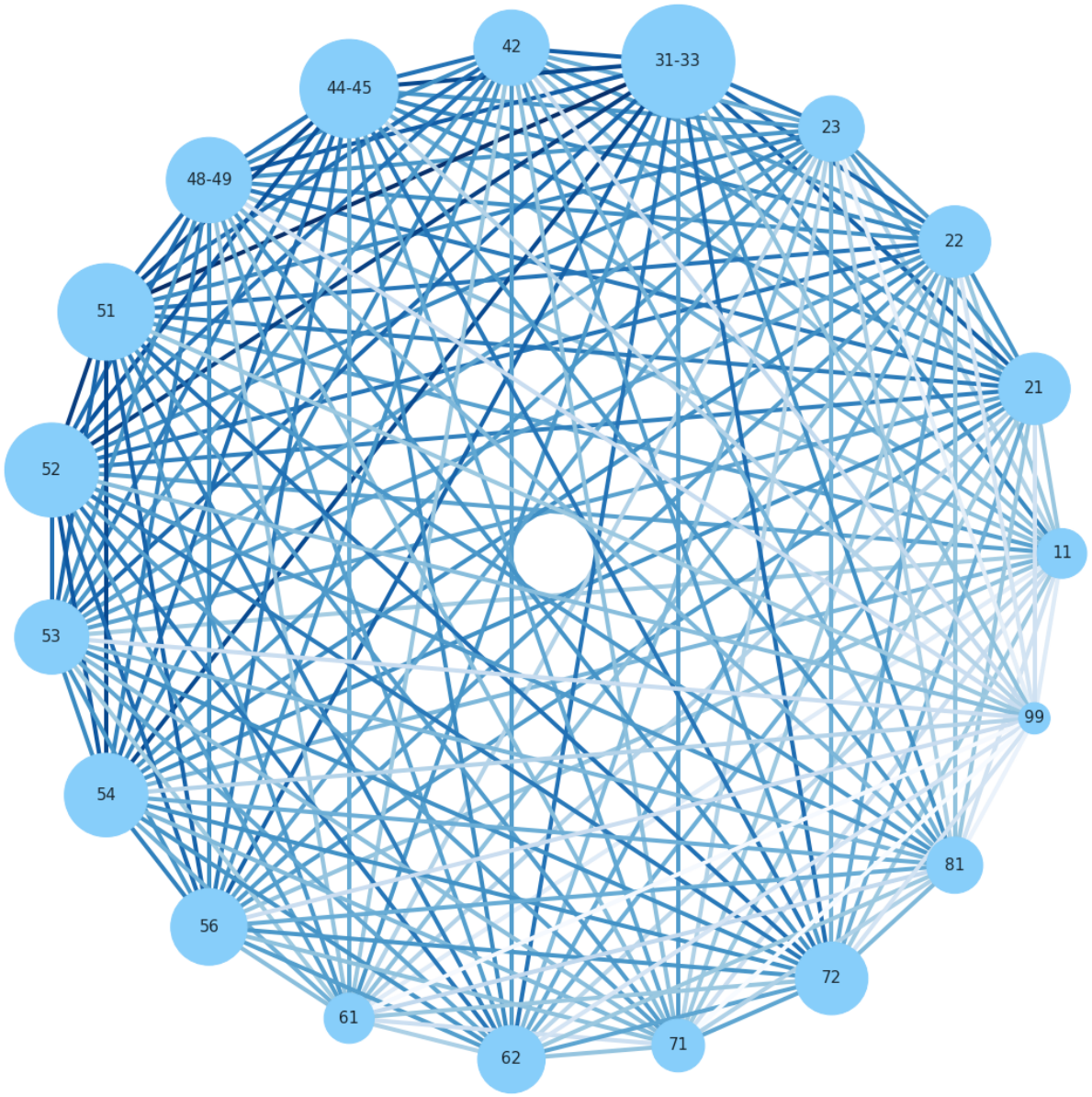


Figure A21 Industry Network Diagram for S&P 100 Firms in 2015

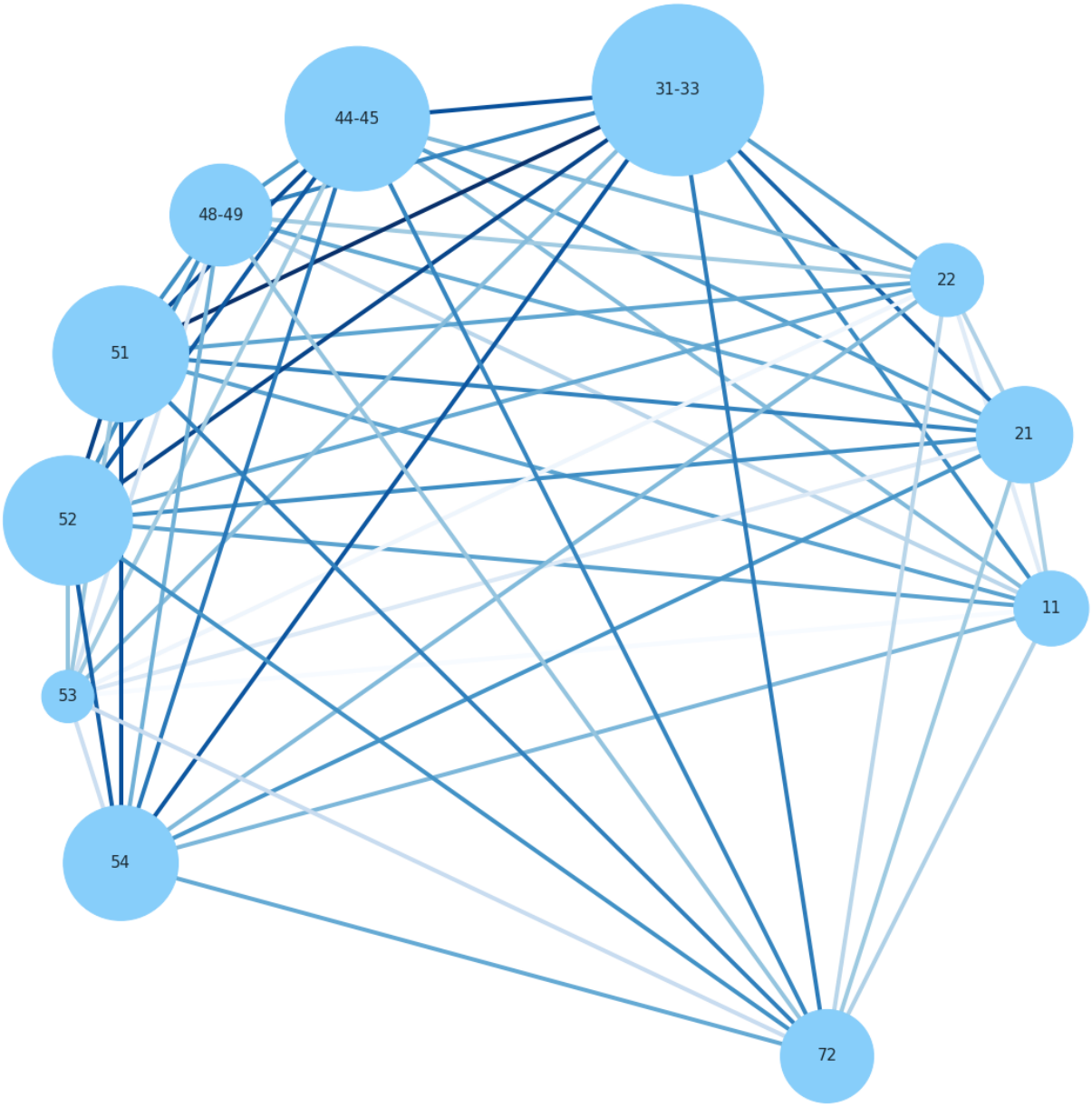


Figure A22 Industry Network Diagram for S&P 500 Firms in 2015

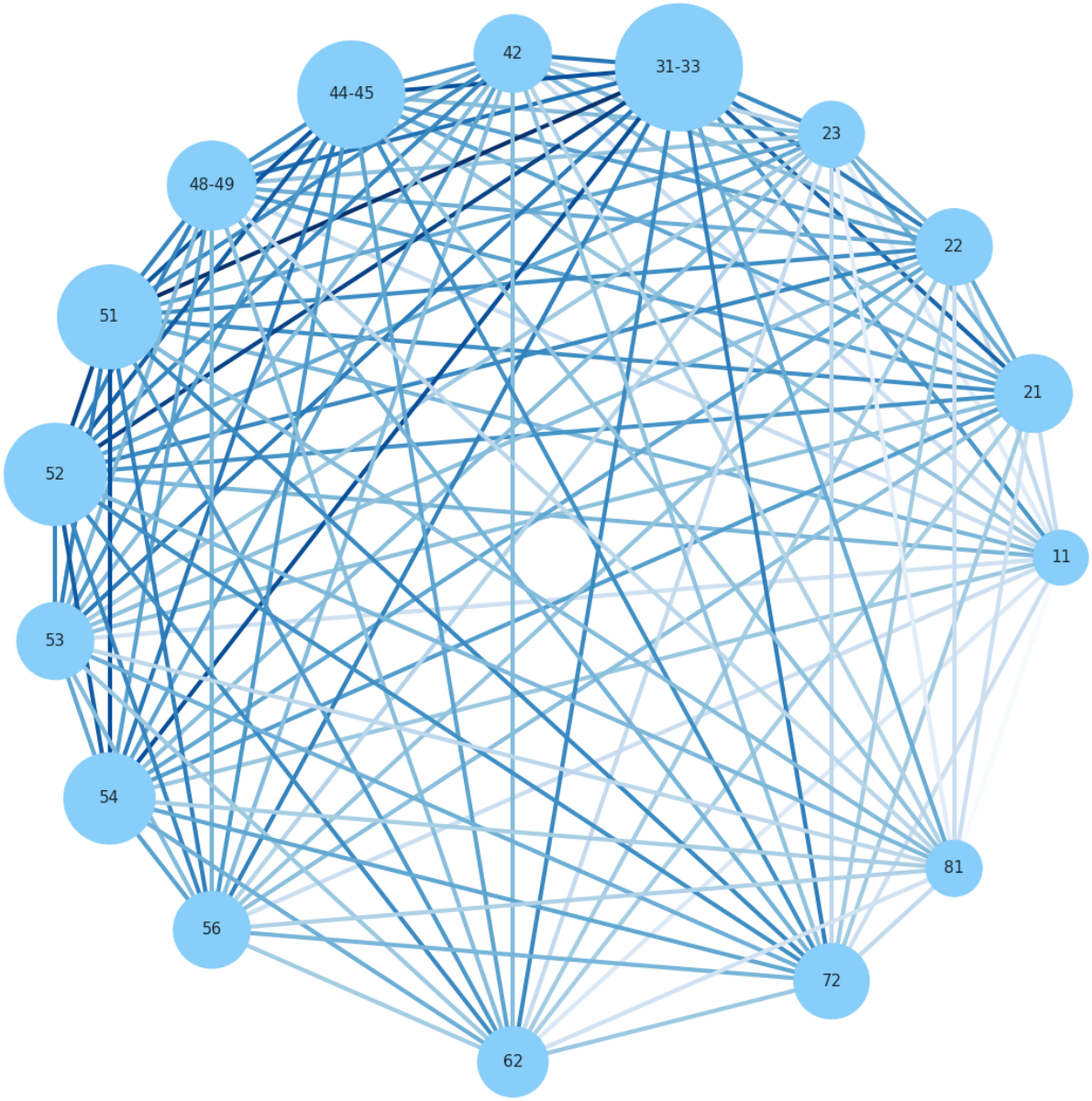


Figure A23 Industry Network Diagram for US Public Firms in 2015

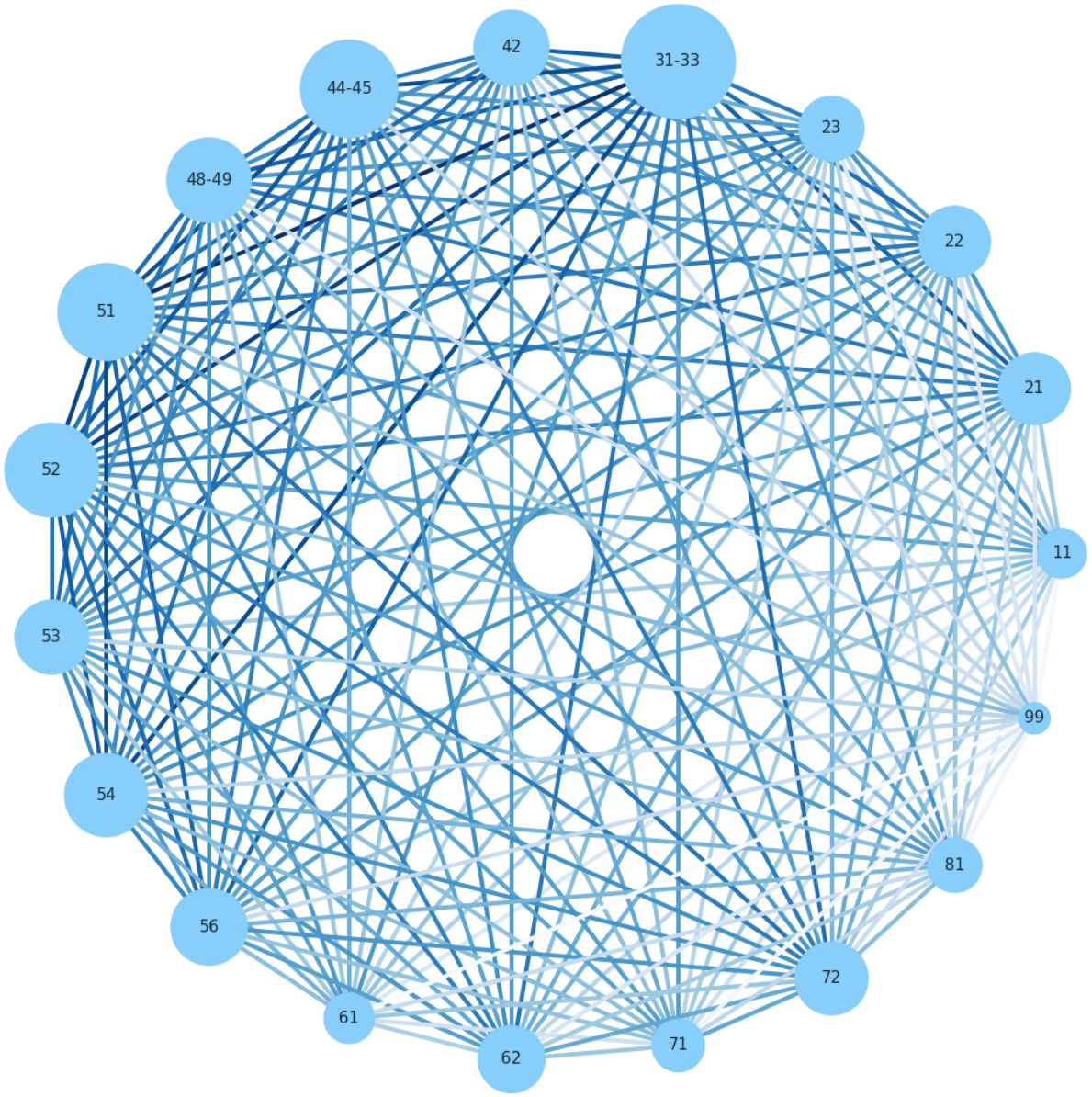


Figure A24 Industry Network Diagram for S&P 100 Firms in 2014

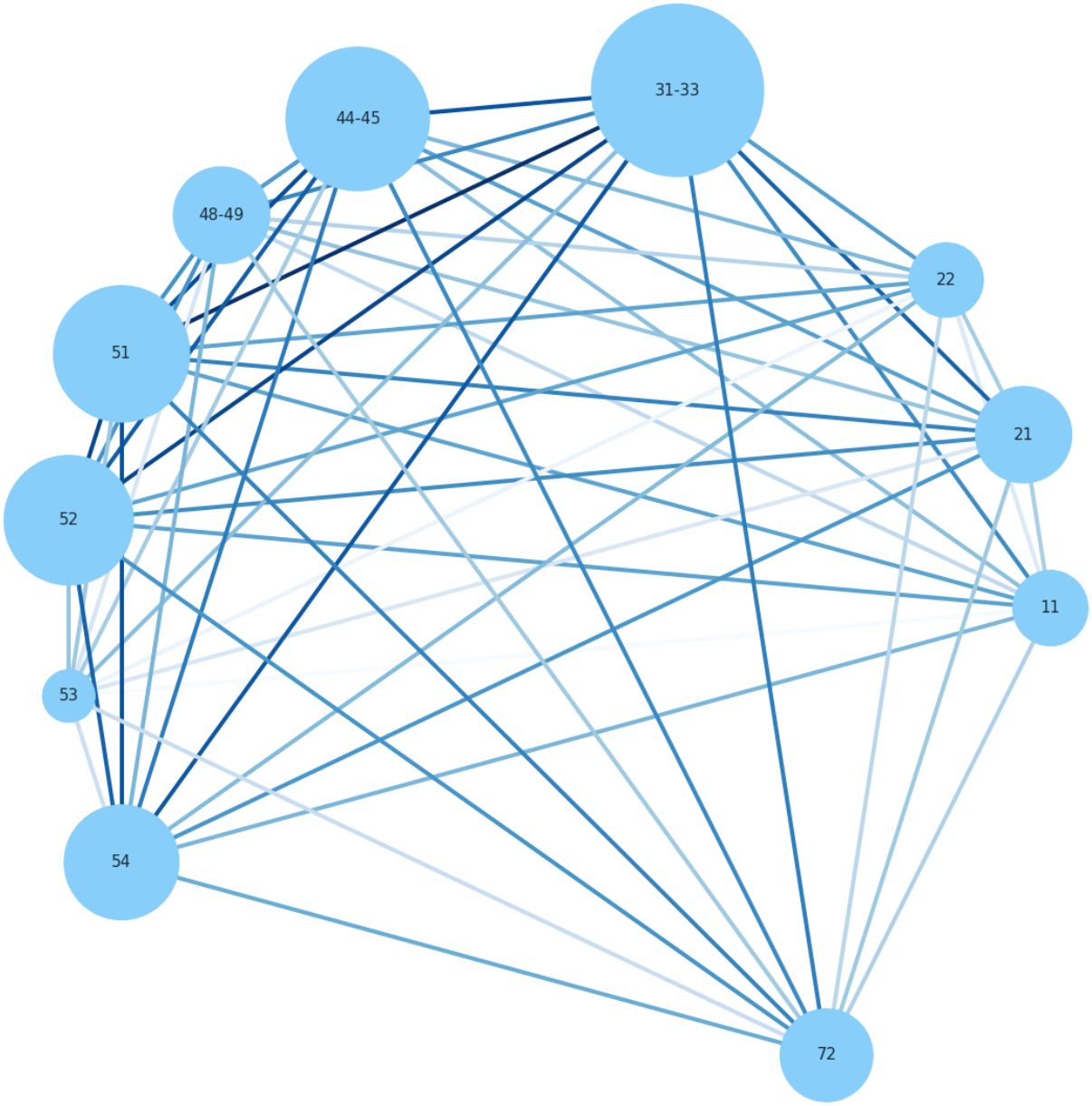


Figure A25 Industry Network Diagram for S&P 500 Firms in 2014

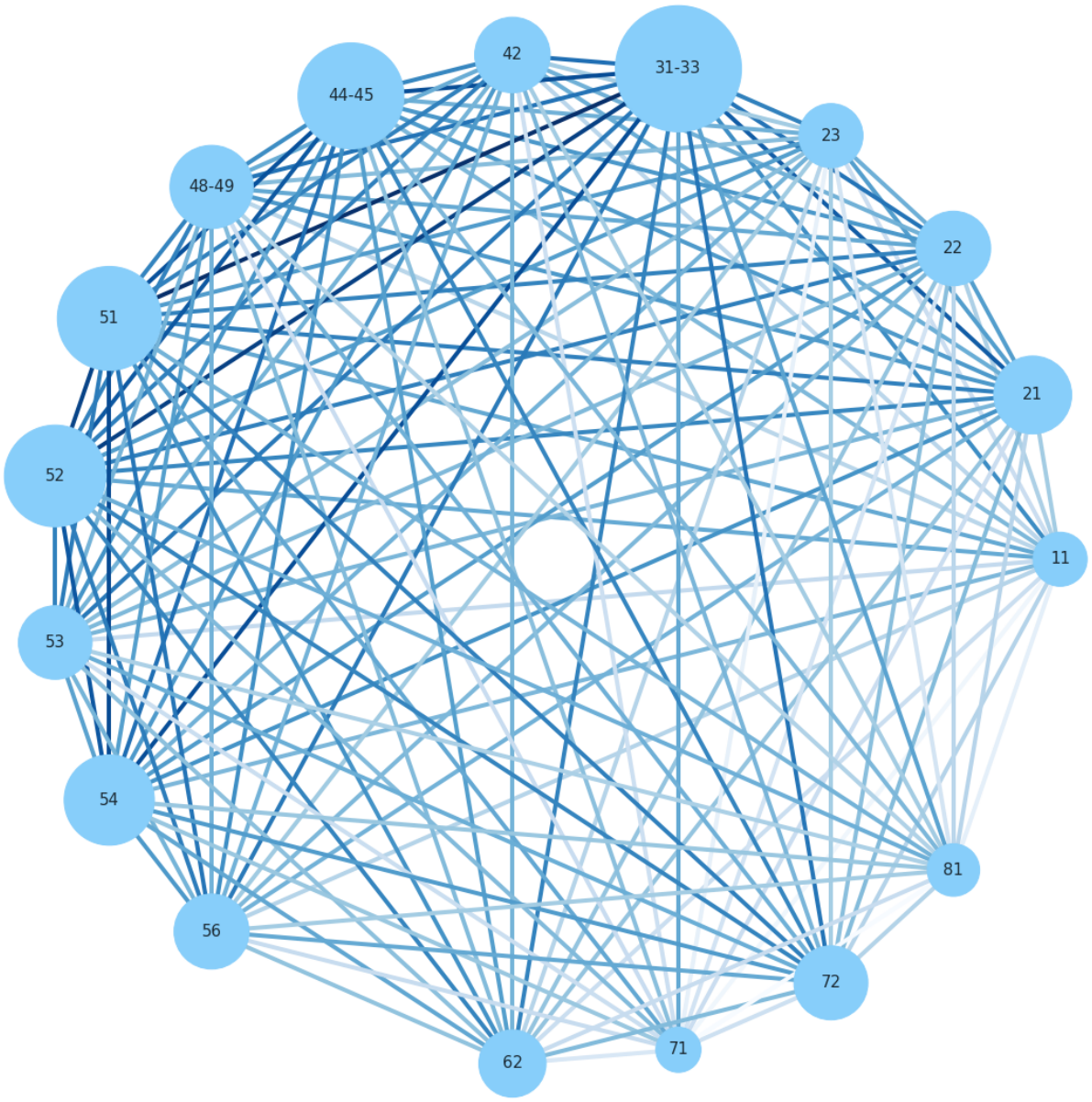


Figure A26 Industry Network Diagram for US Public Firms in 2014

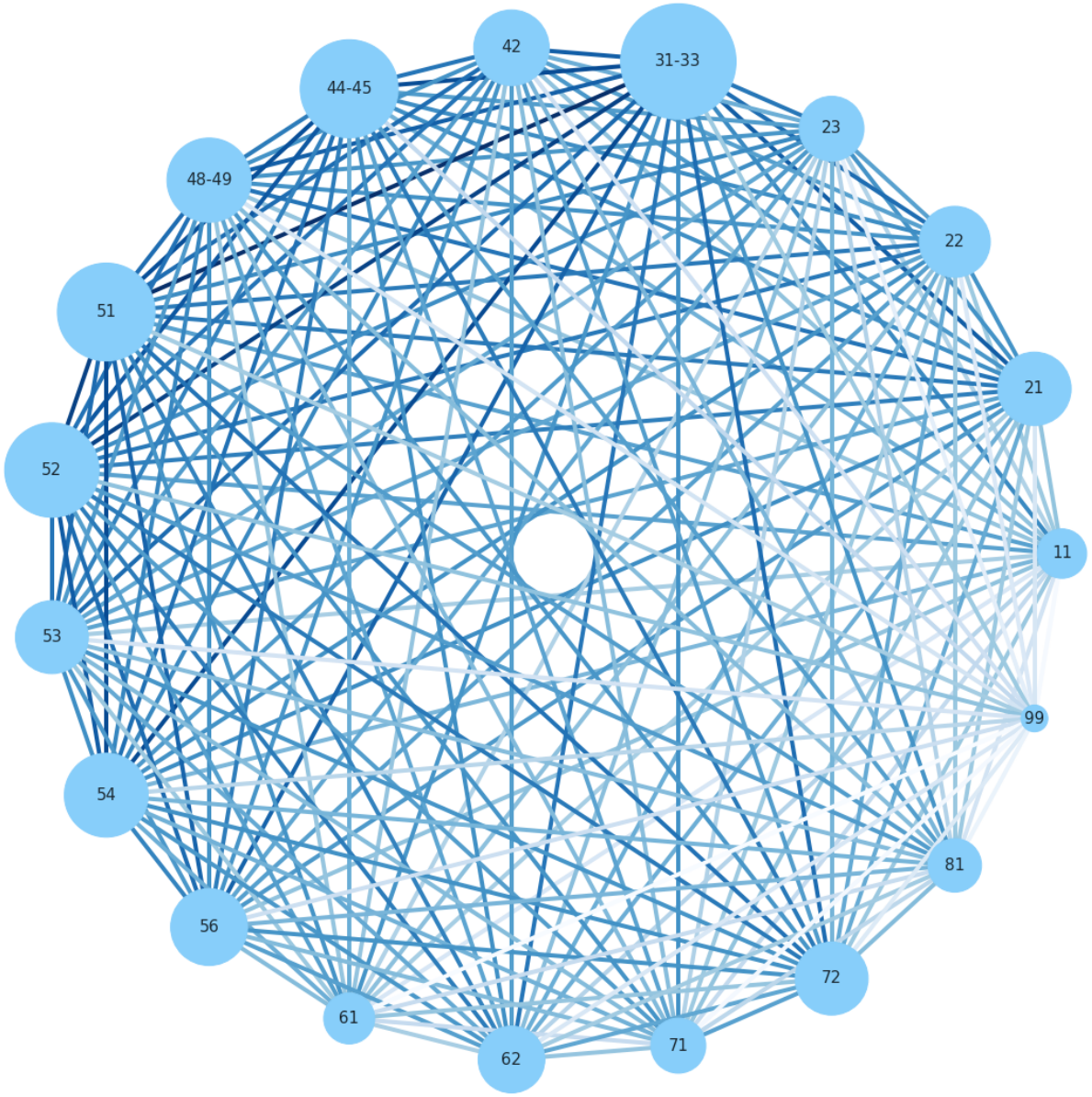


Figure A27 Industry Network Diagram for S&P 100 Firms in 2013

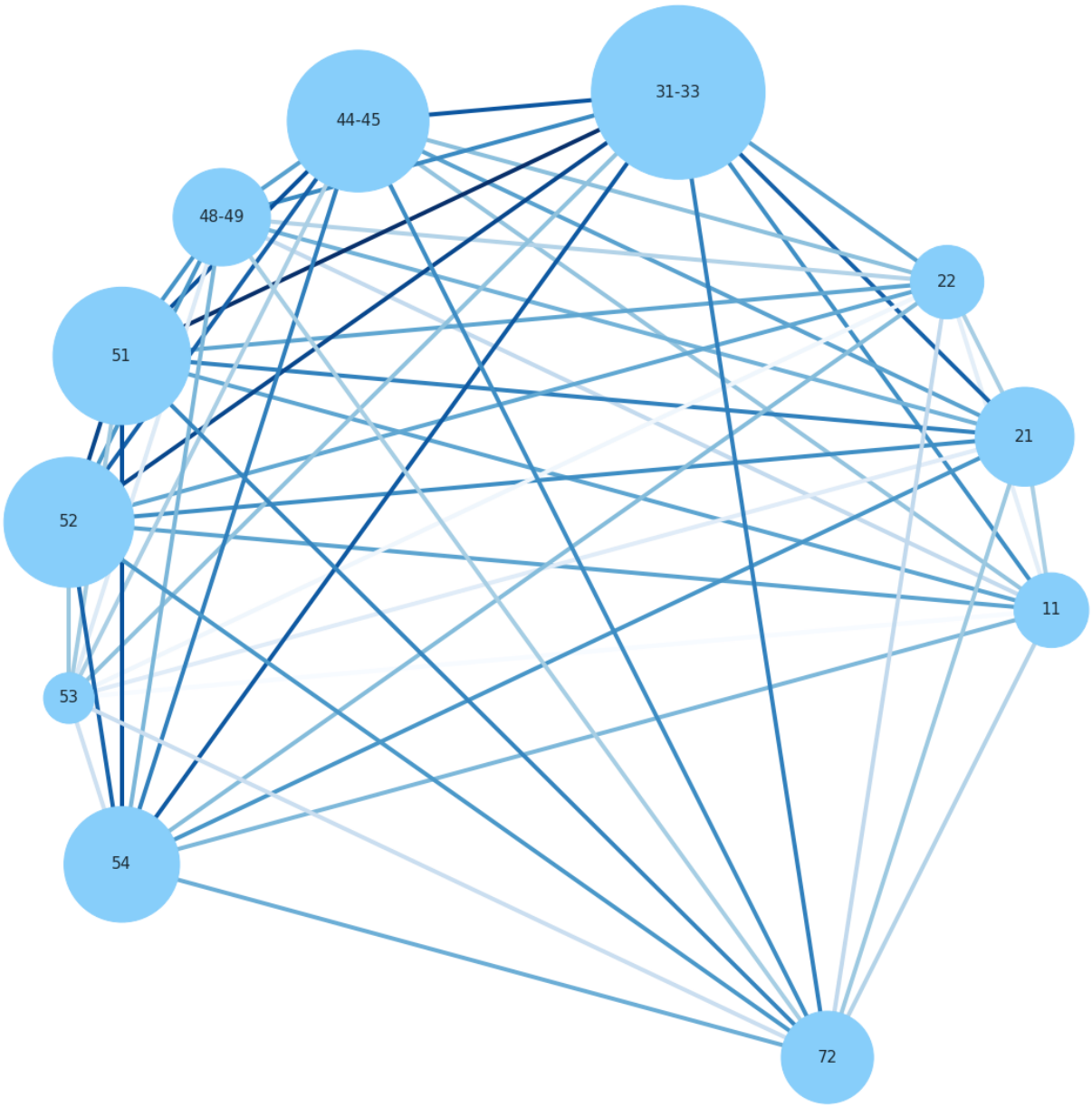


Figure A28 Industry Network Diagram for S&P 500 Firms in 2013

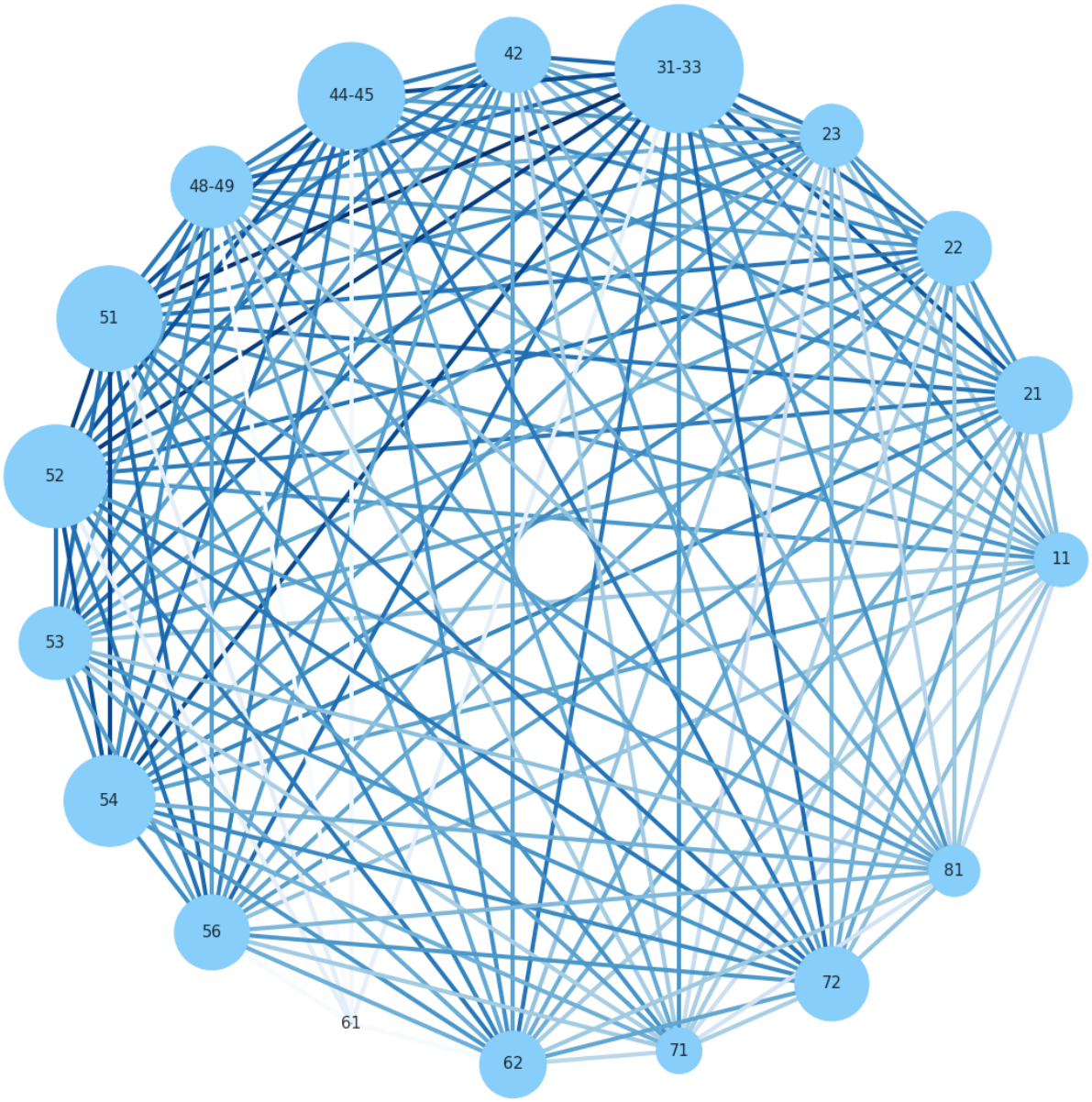


Figure A29 Industry Network Diagram for US Public Firms in 2013

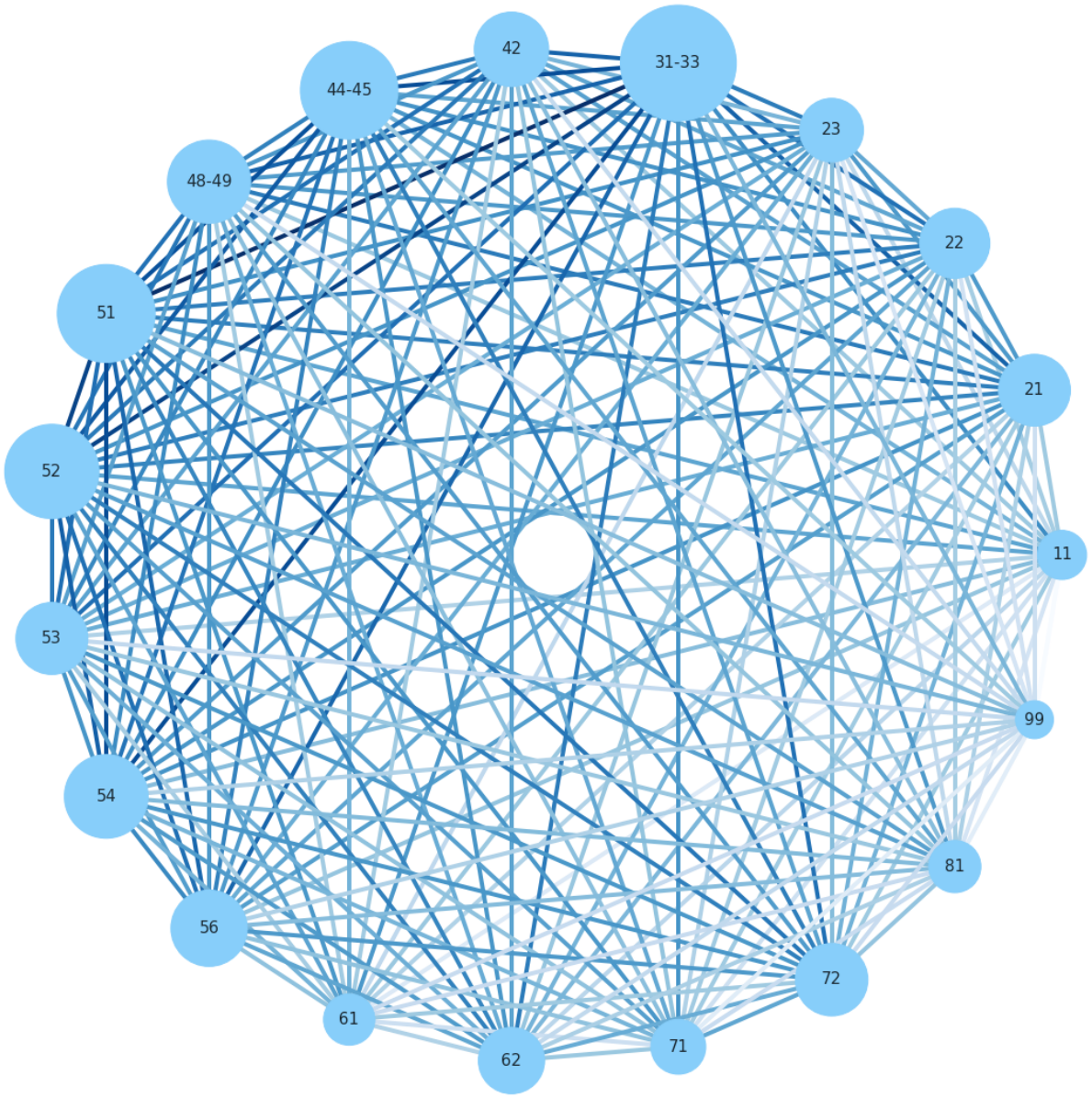


Figure A30 Industry Network Diagram for S&P 100 Firms in 2012

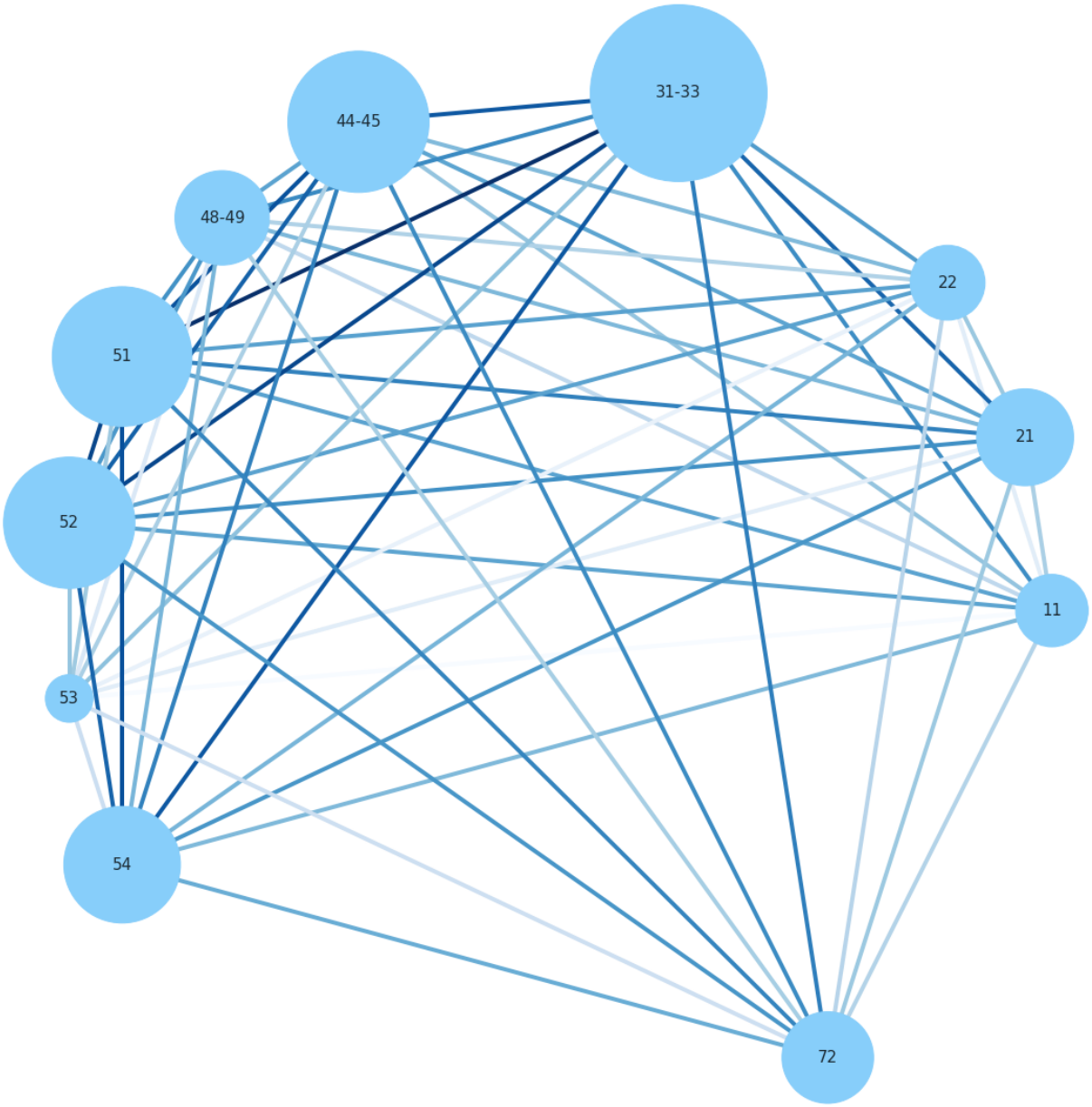


Figure A31 Industry Network Diagram for S&P 500 Firms in 2012

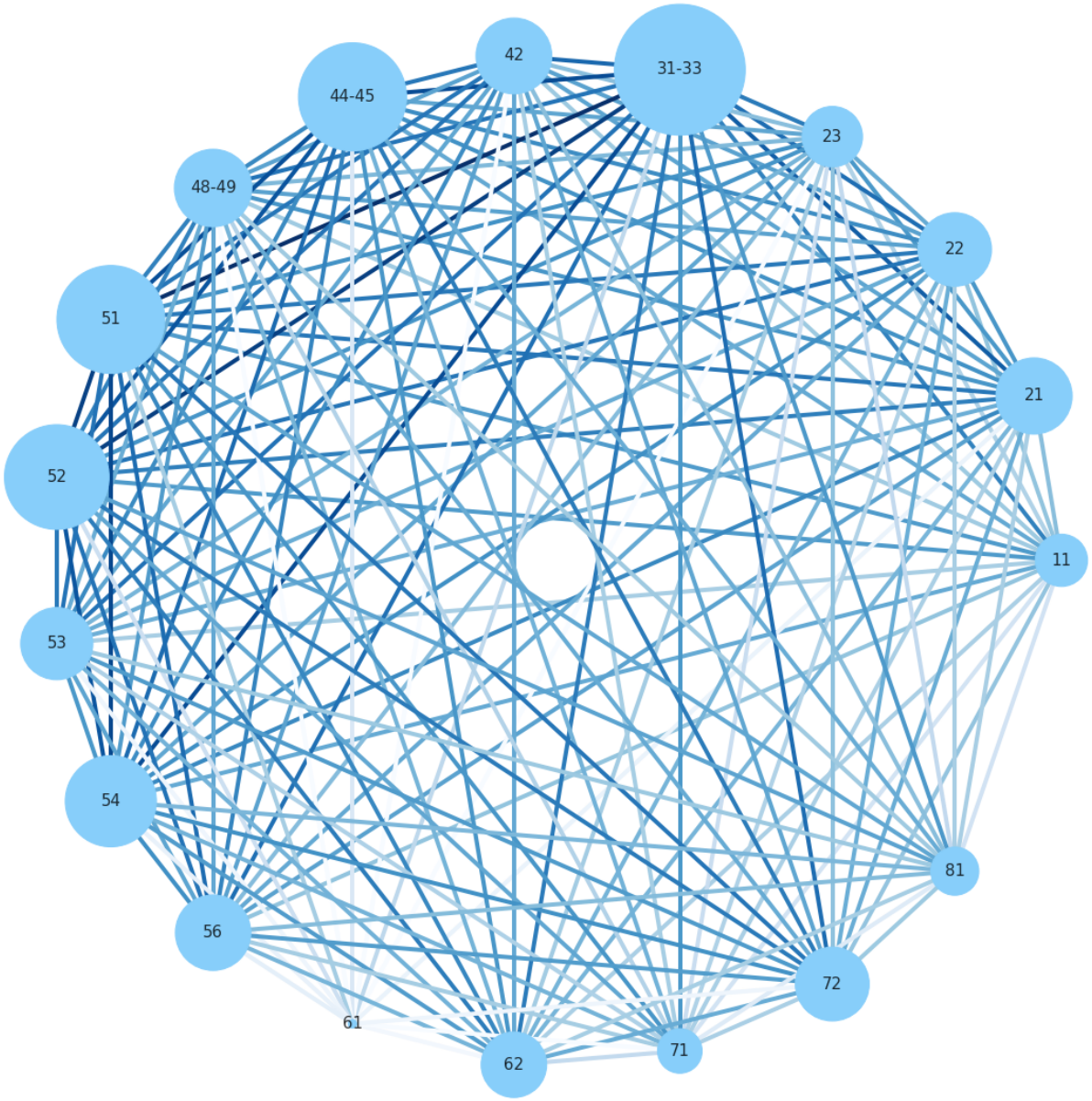


Figure A32 Industry Network Diagram for US Public Firms in 2012

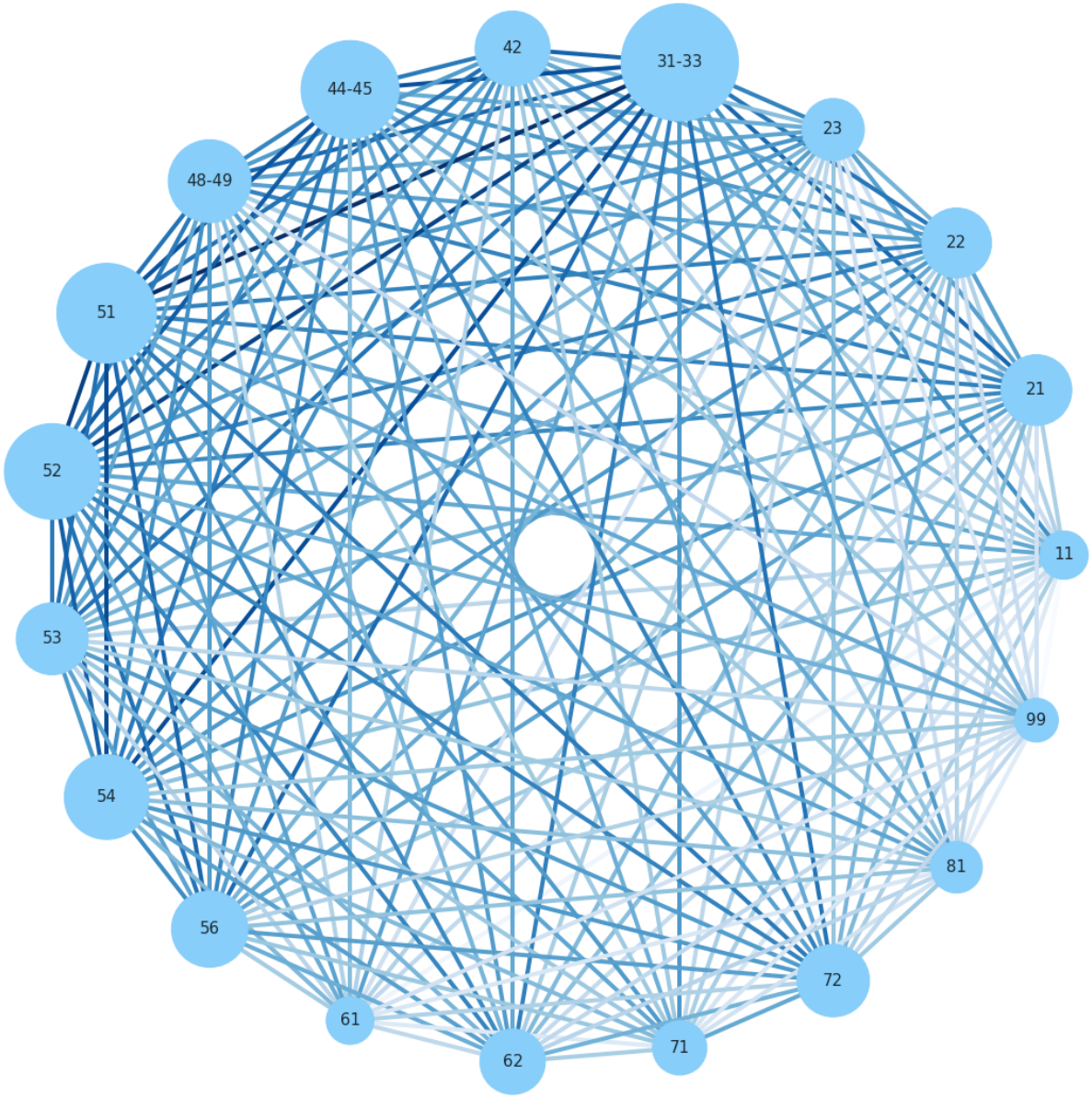


Figure A33 Industry Network Diagram for S&P 100 Firms in 2011

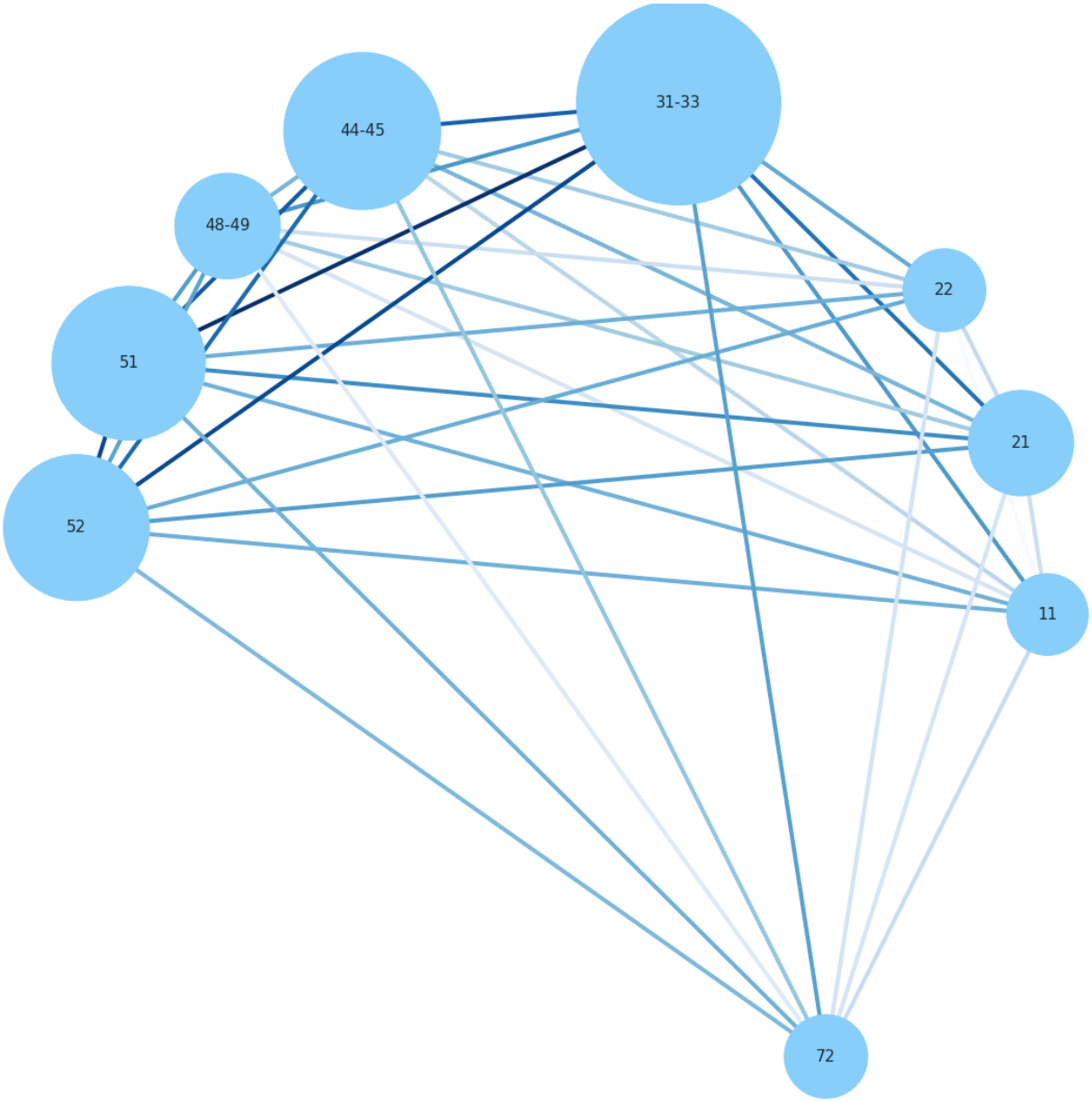


Figure A34 Industry Network Diagram for S&P 500 Firms in 2011

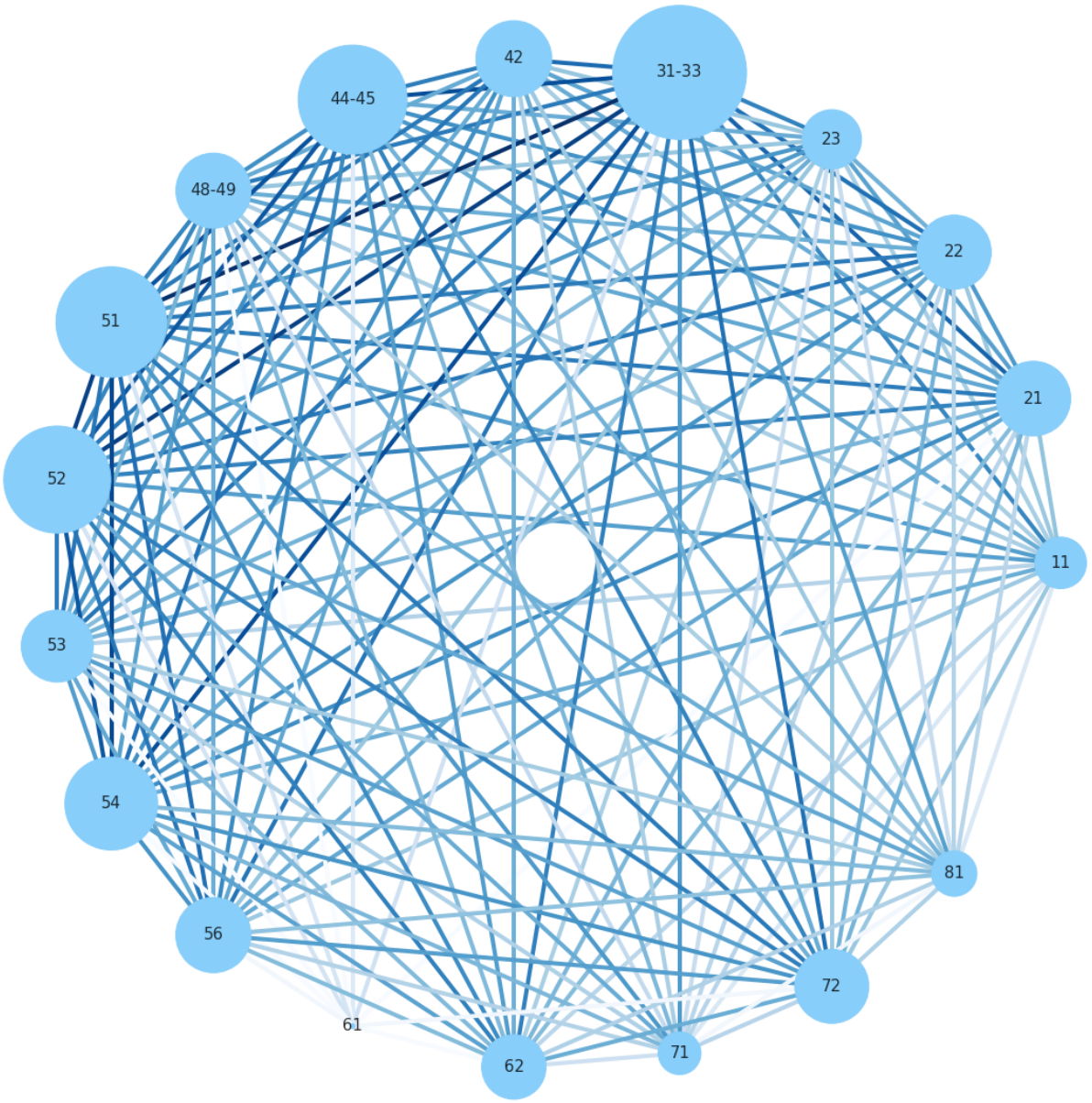


Figure A35 Industry Network Diagram for US Public Firms in 2011

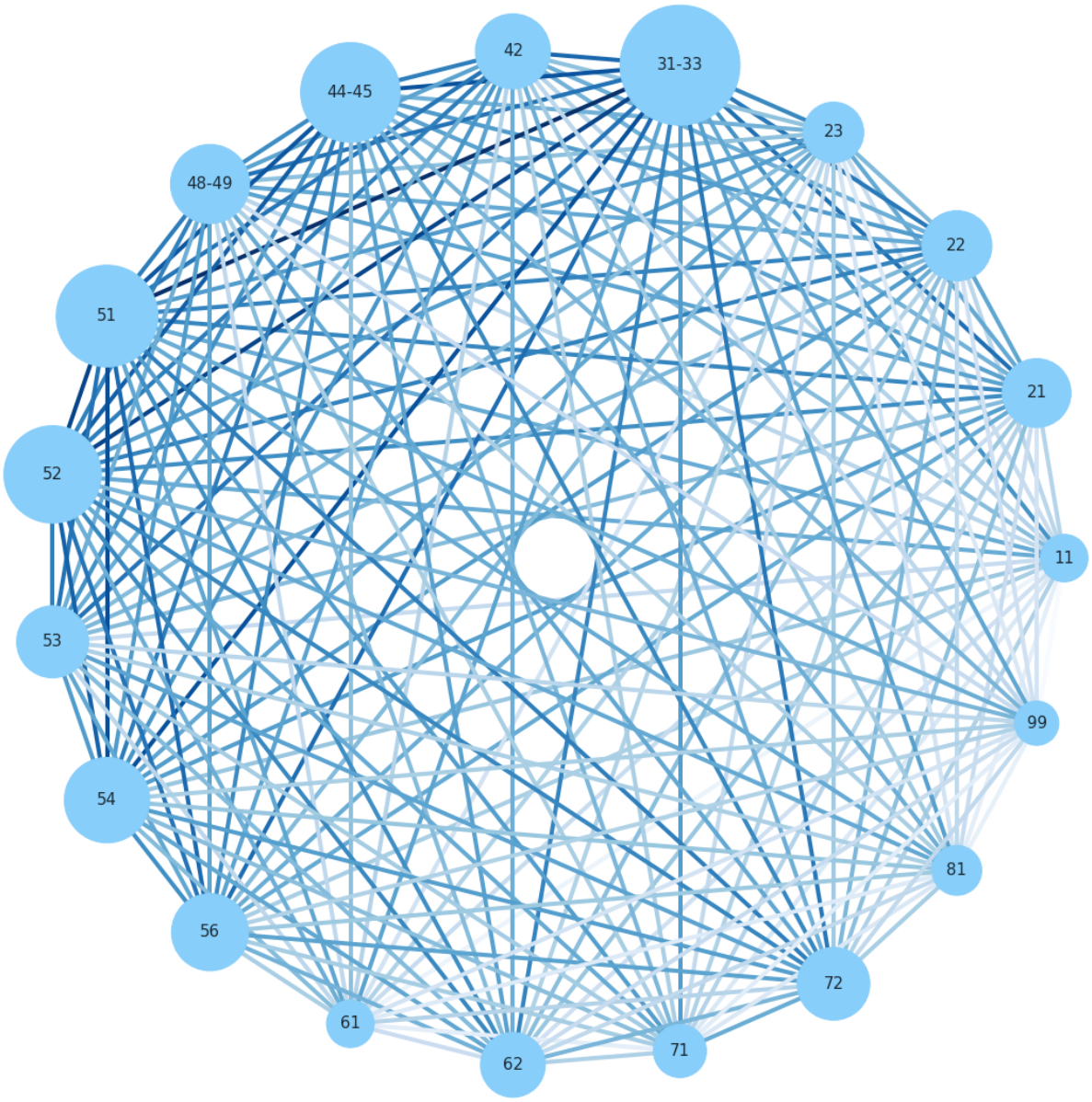


Figure A33 Industry Network Diagram for S&P 100 Firms in 2010

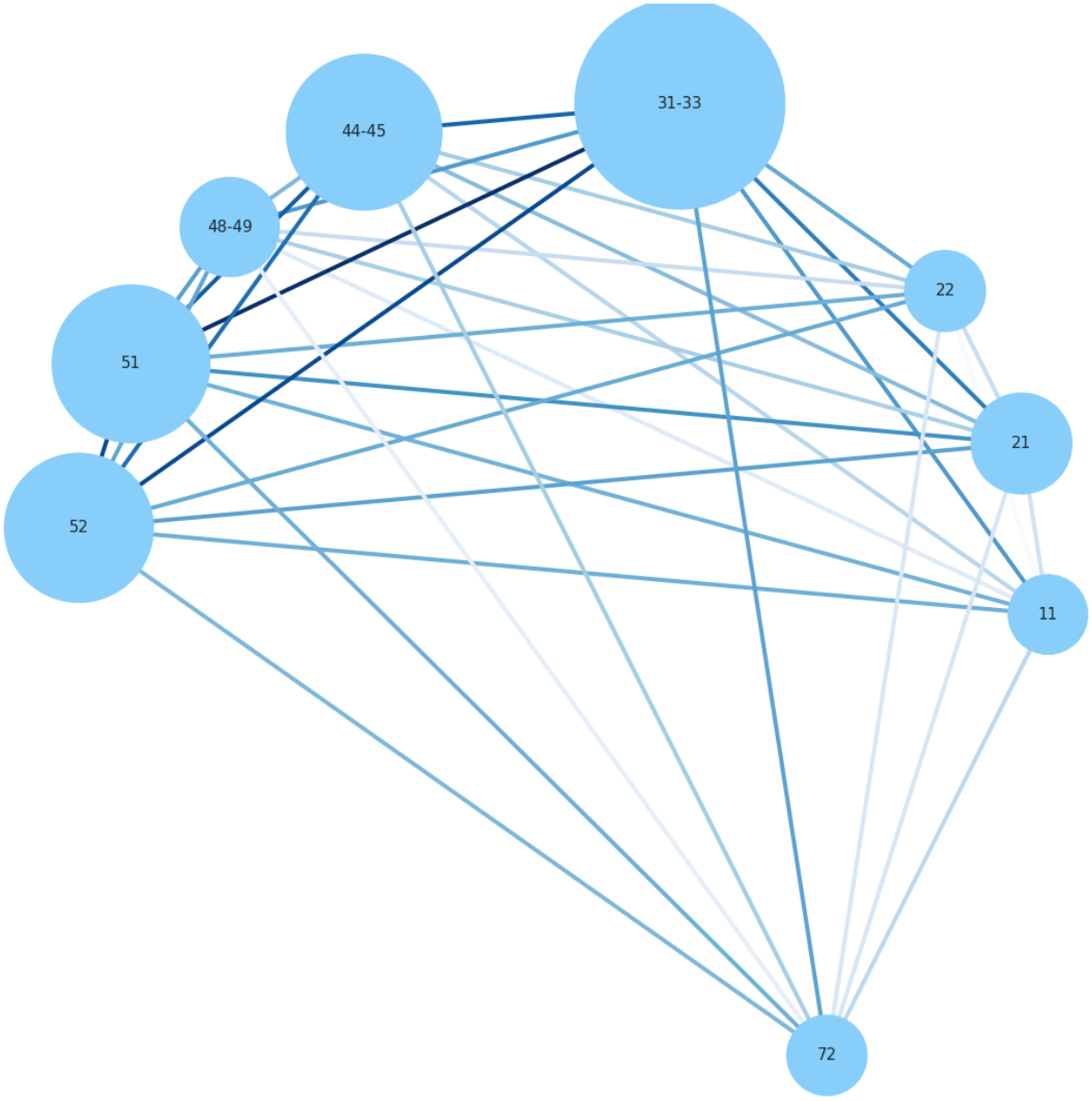


Figure A34 Industry Network Diagram for S&P 500 Firms in 2010

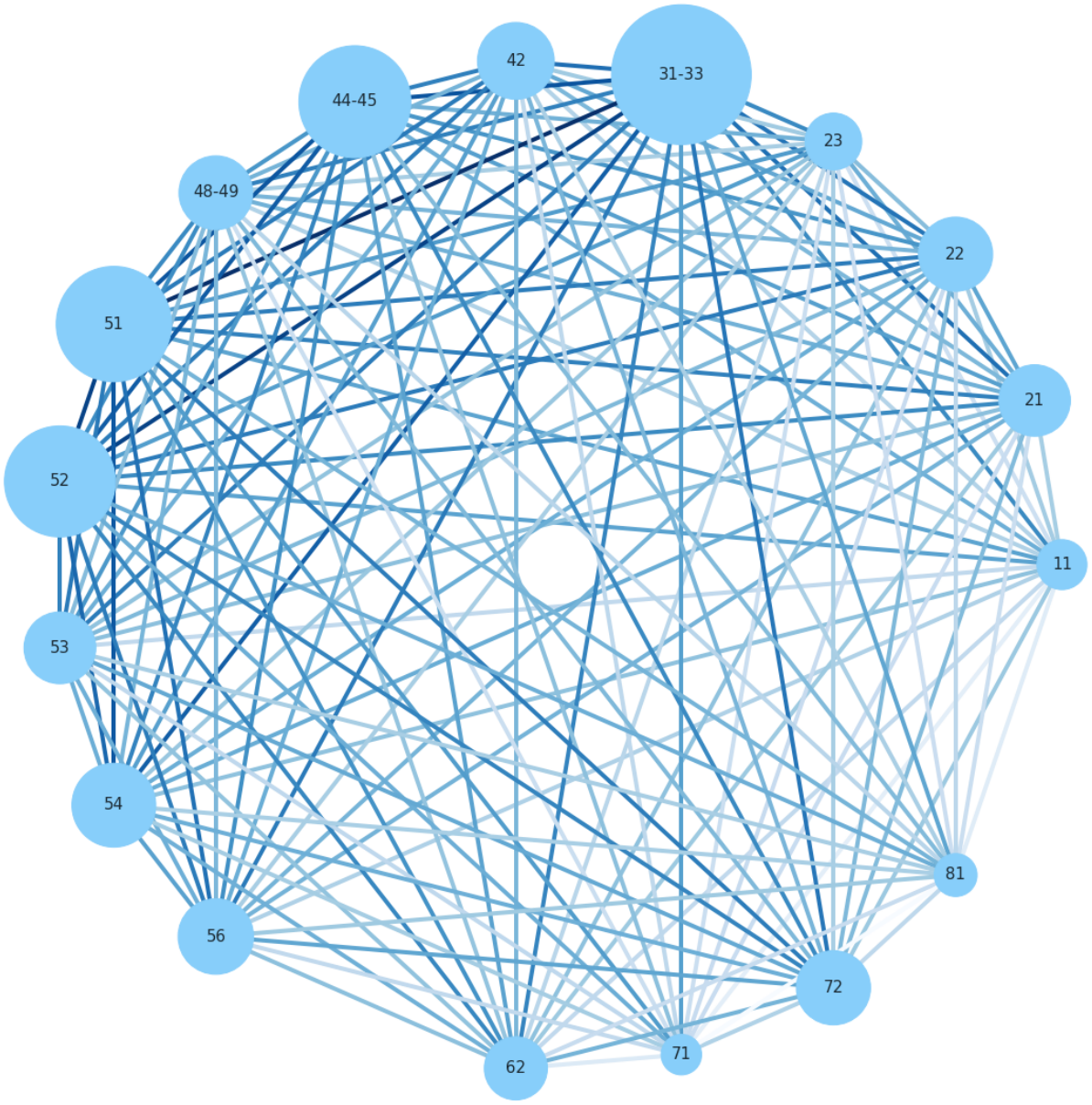


Figure A35 Industry Network Diagram for US Public Firms in 2010

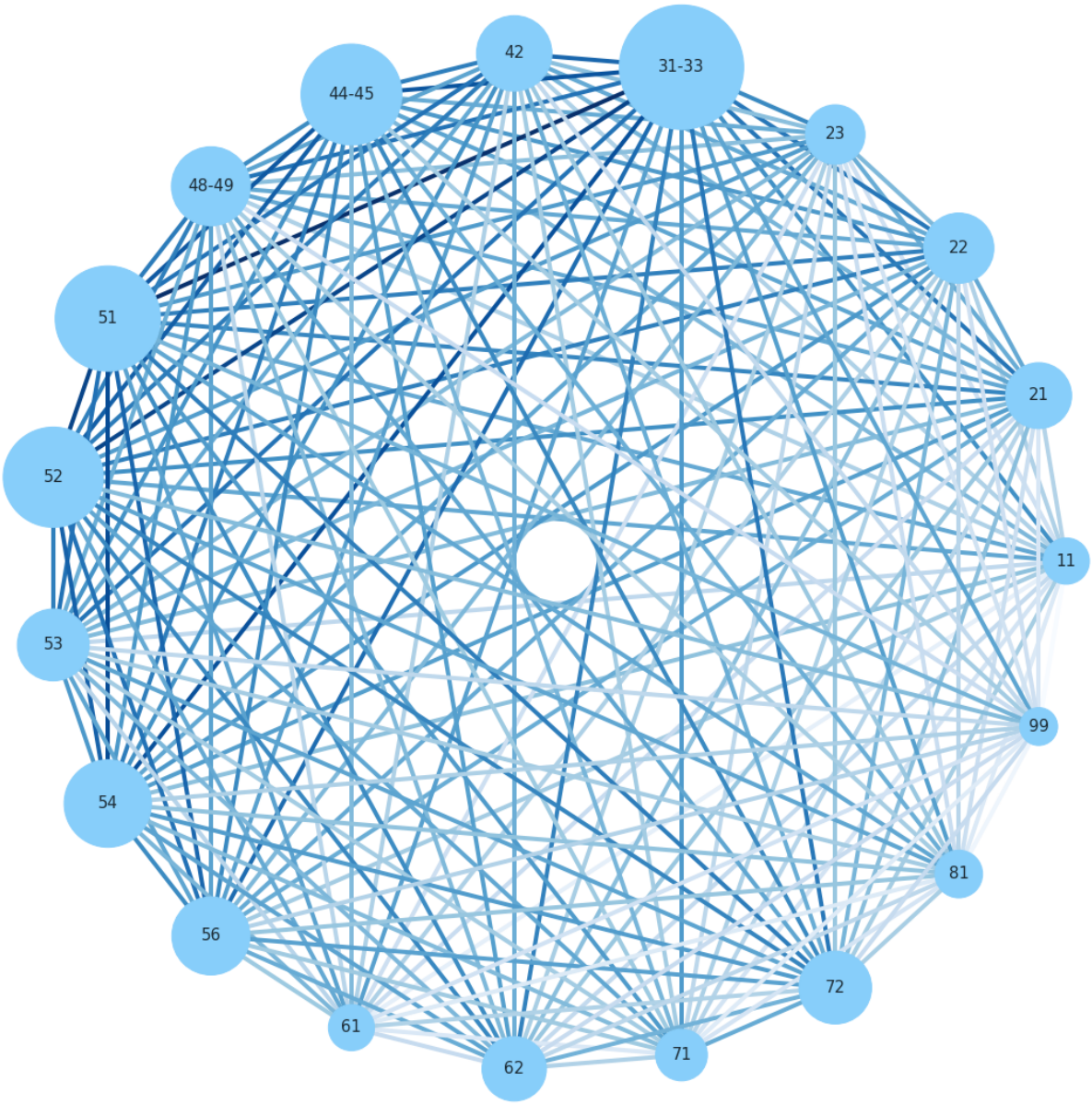


Figure A36 Industry Network Diagram for S&P 100 Firms in 2009

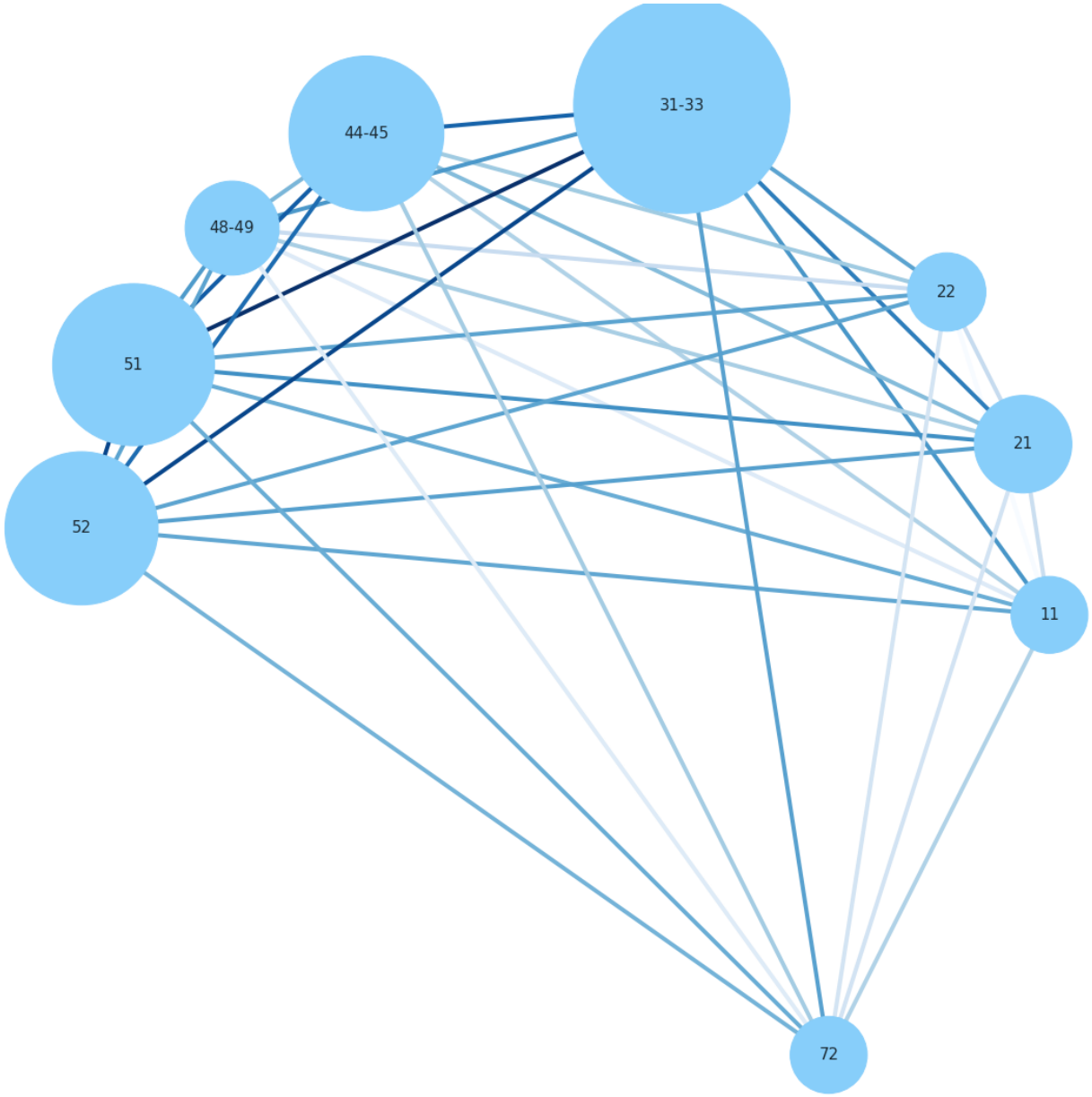


Figure A37 Industry Network Diagram for S&P 500 Firms in 2009

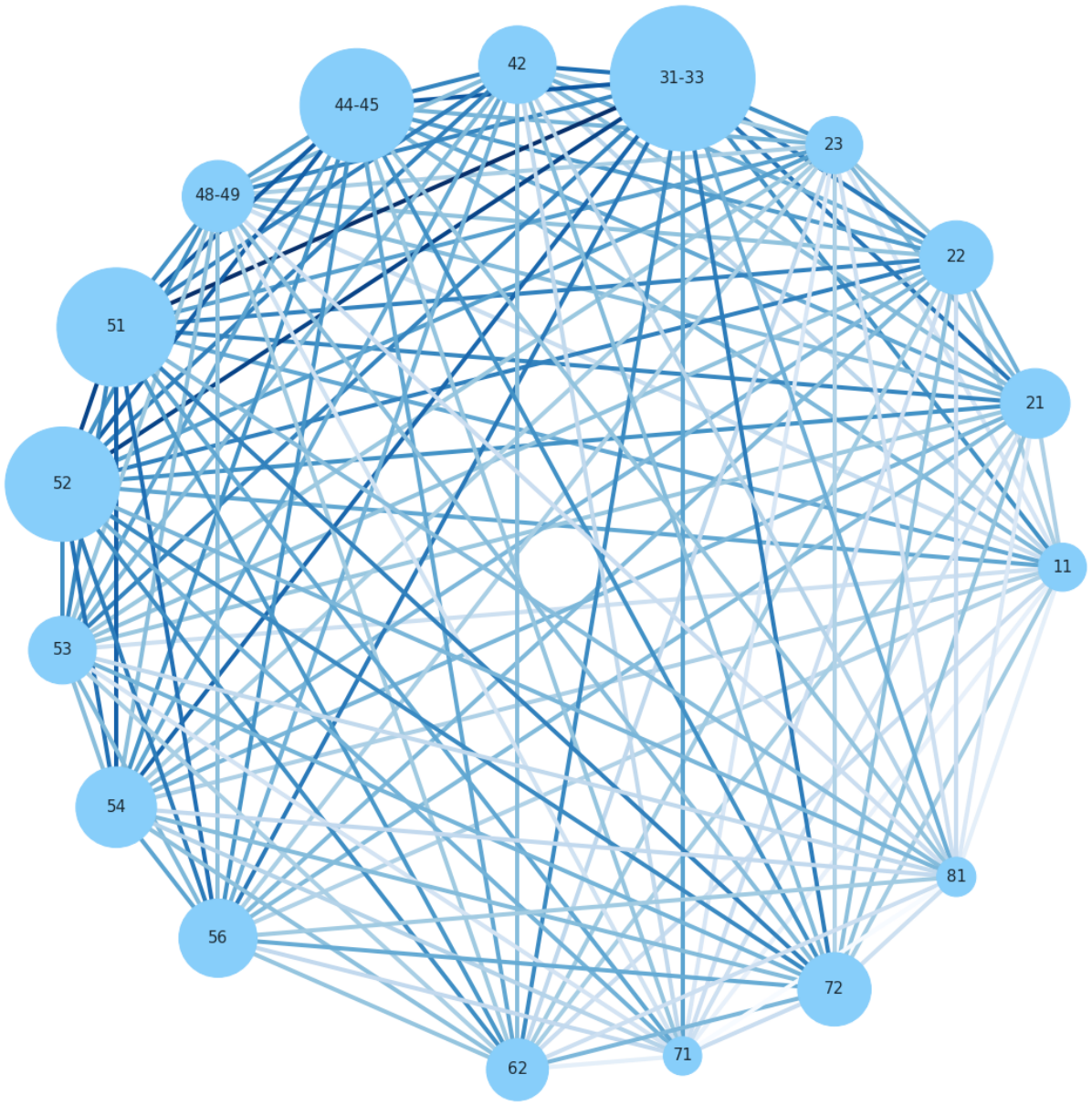


Figure A38 Industry Network Diagram for US Public Firms in 2009

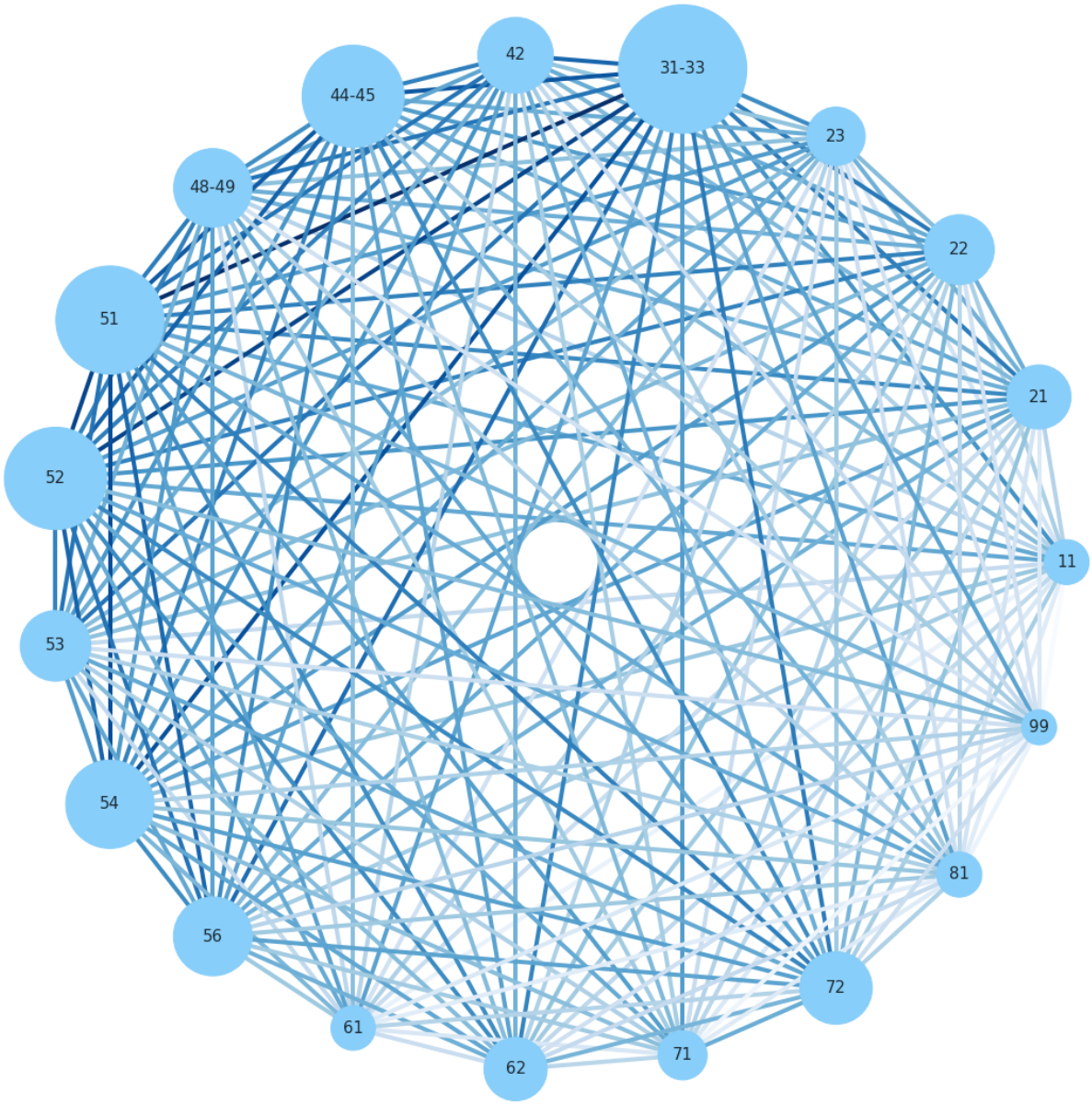


Figure A39 Industry Network Diagram for S&P 100 Firms in 2008

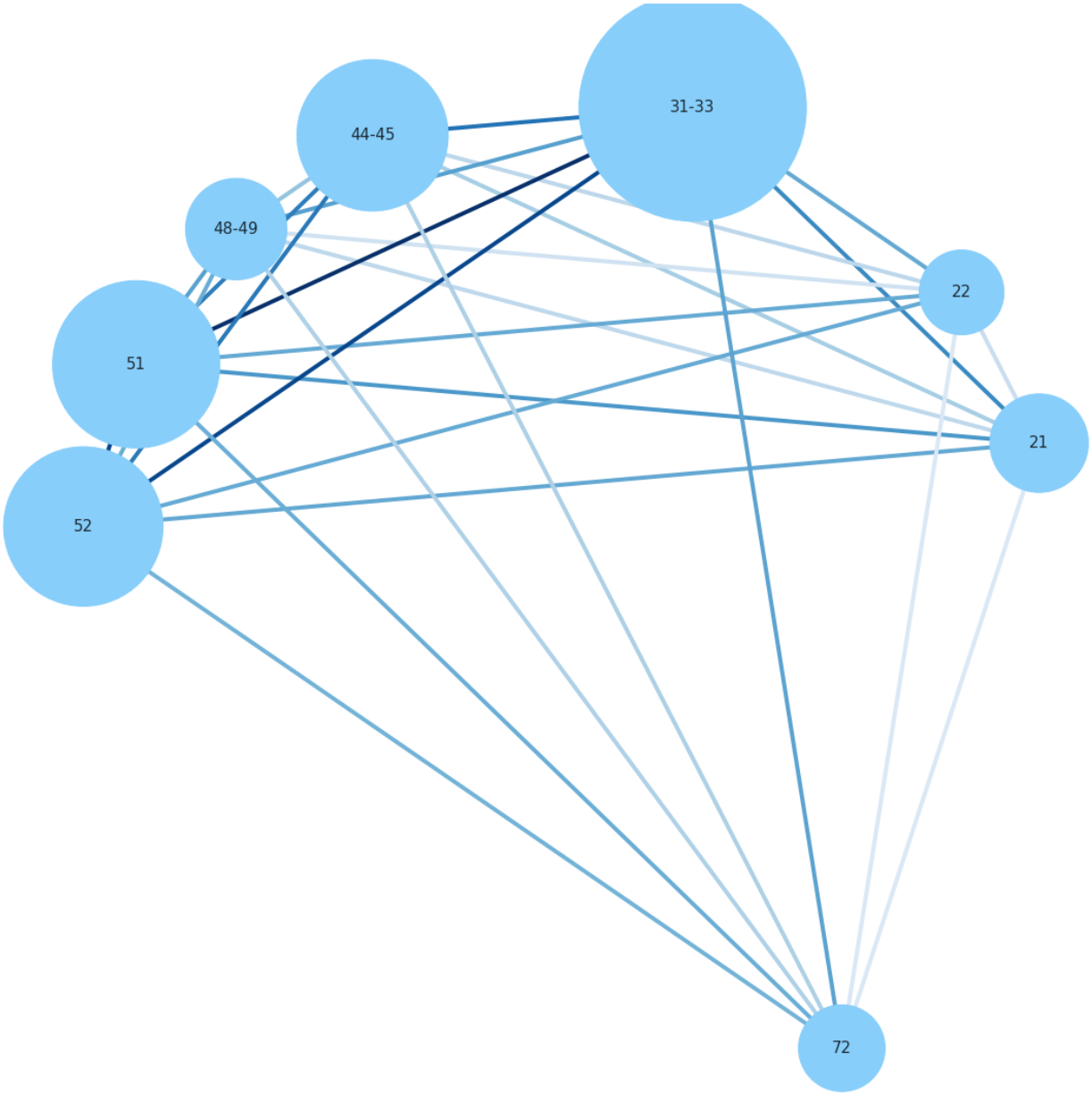


Figure A40 Industry Network Diagram for S&P 500 Firms in 2008

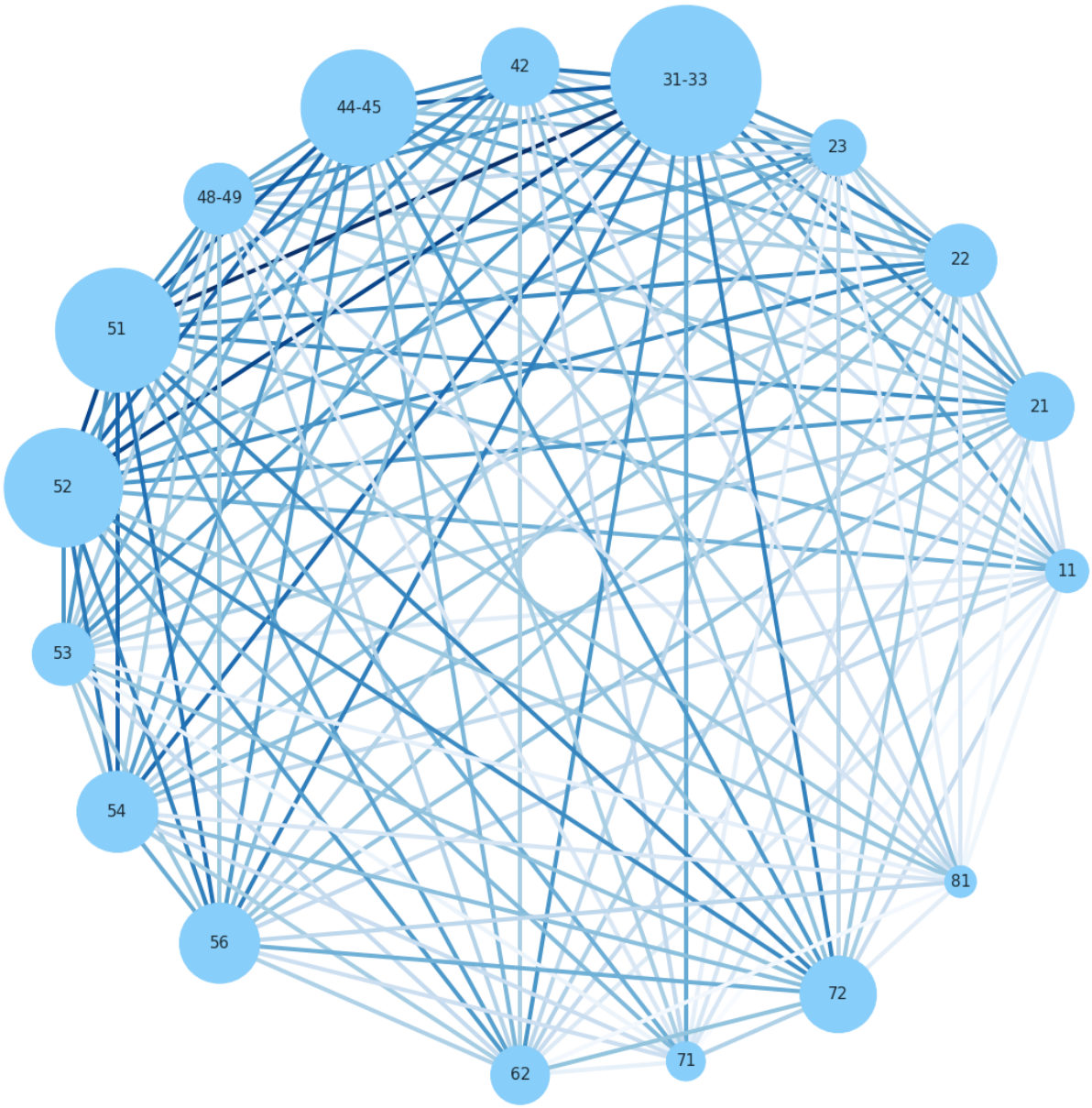


Figure A41 Industry Network Diagram for US Public Firms in 2008

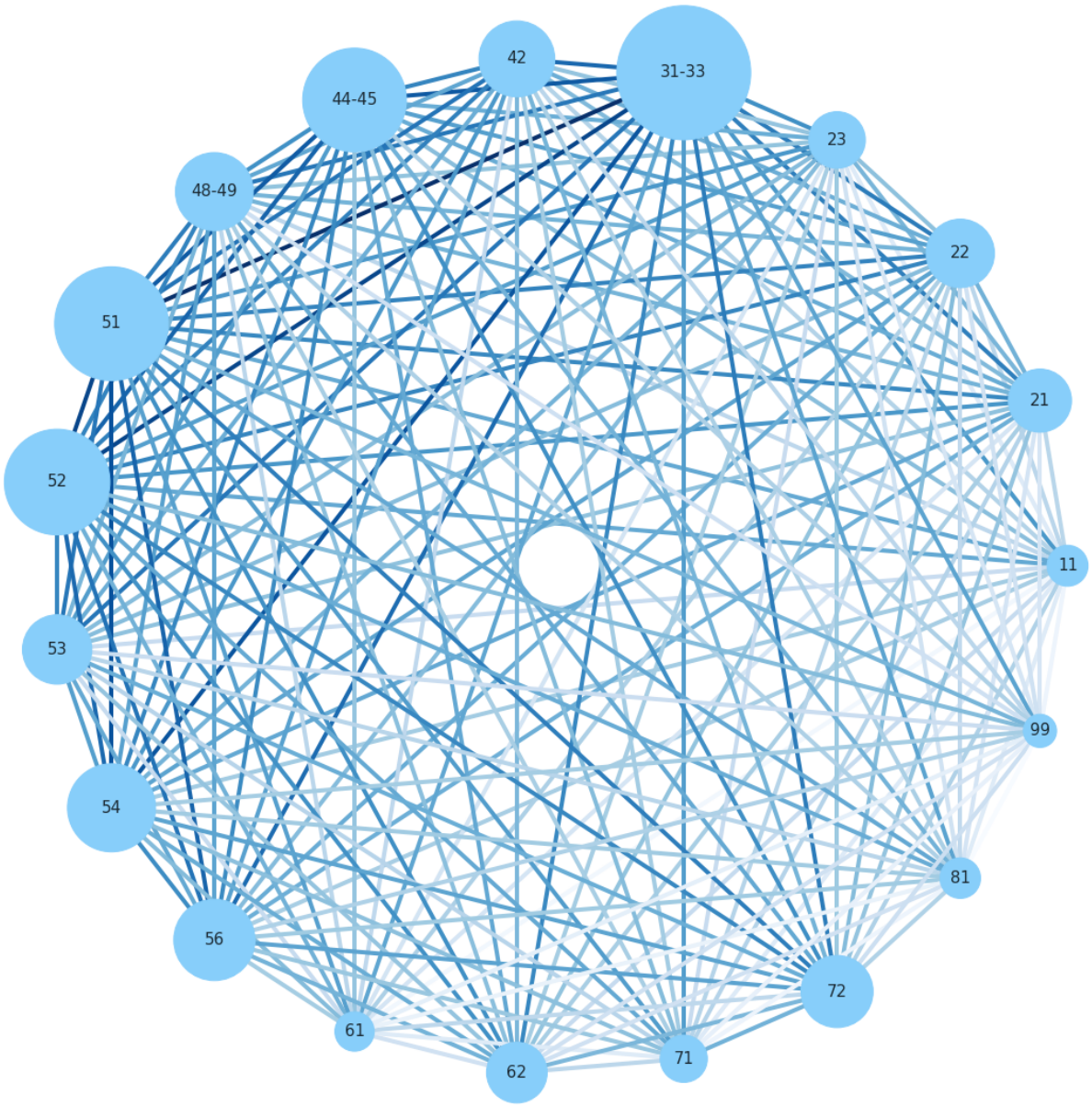


Figure A42 Industry Network Diagram for S&P 100 Firms in 2007

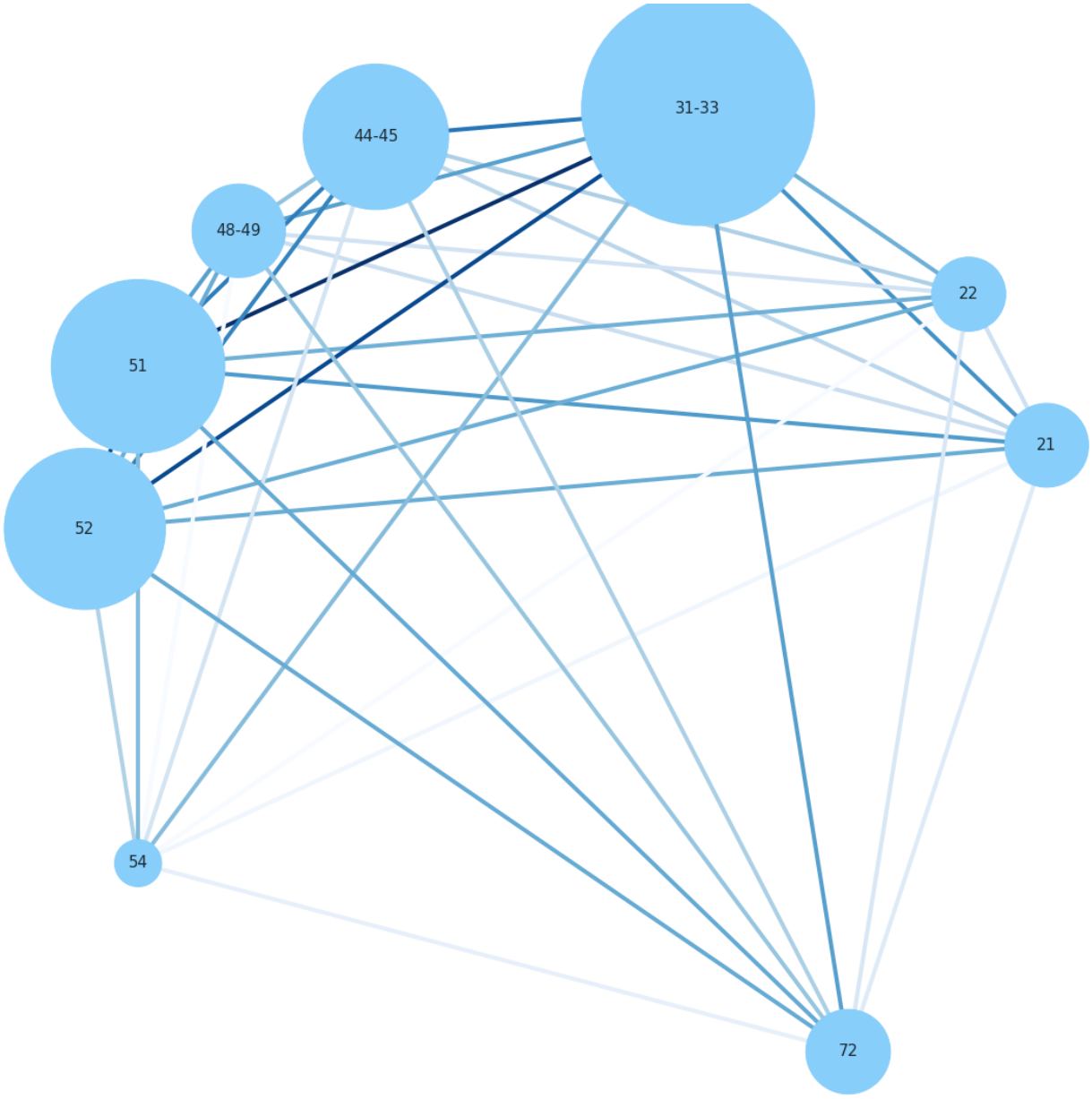


Figure A43 Industry Network Diagram for S&P 500 Firms in 2007

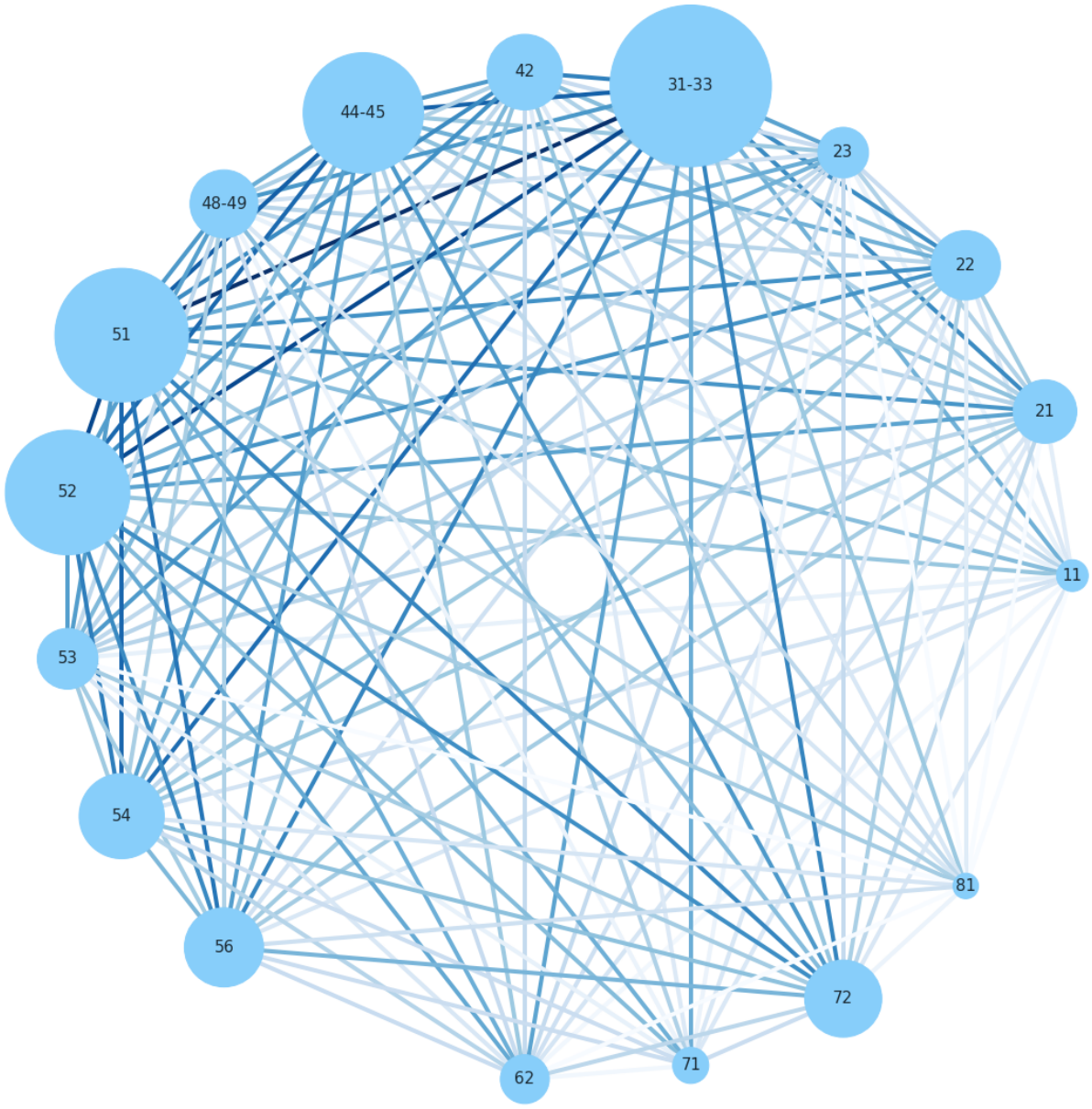


Figure A44 Industry Network Diagram for US Public Firms in 2007

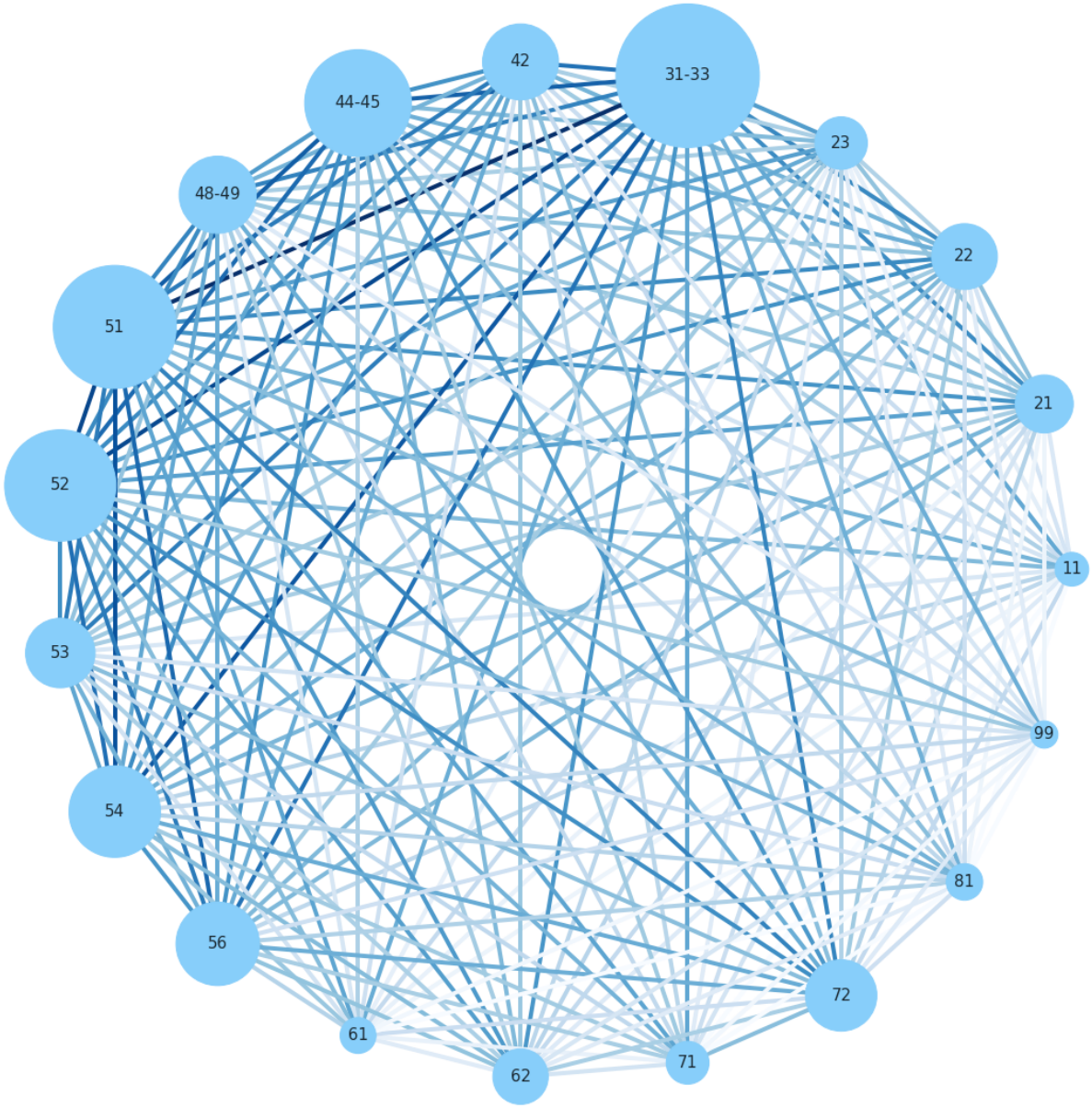


Figure A45 Industry Network Diagram for S&P 100 Firms in 2006

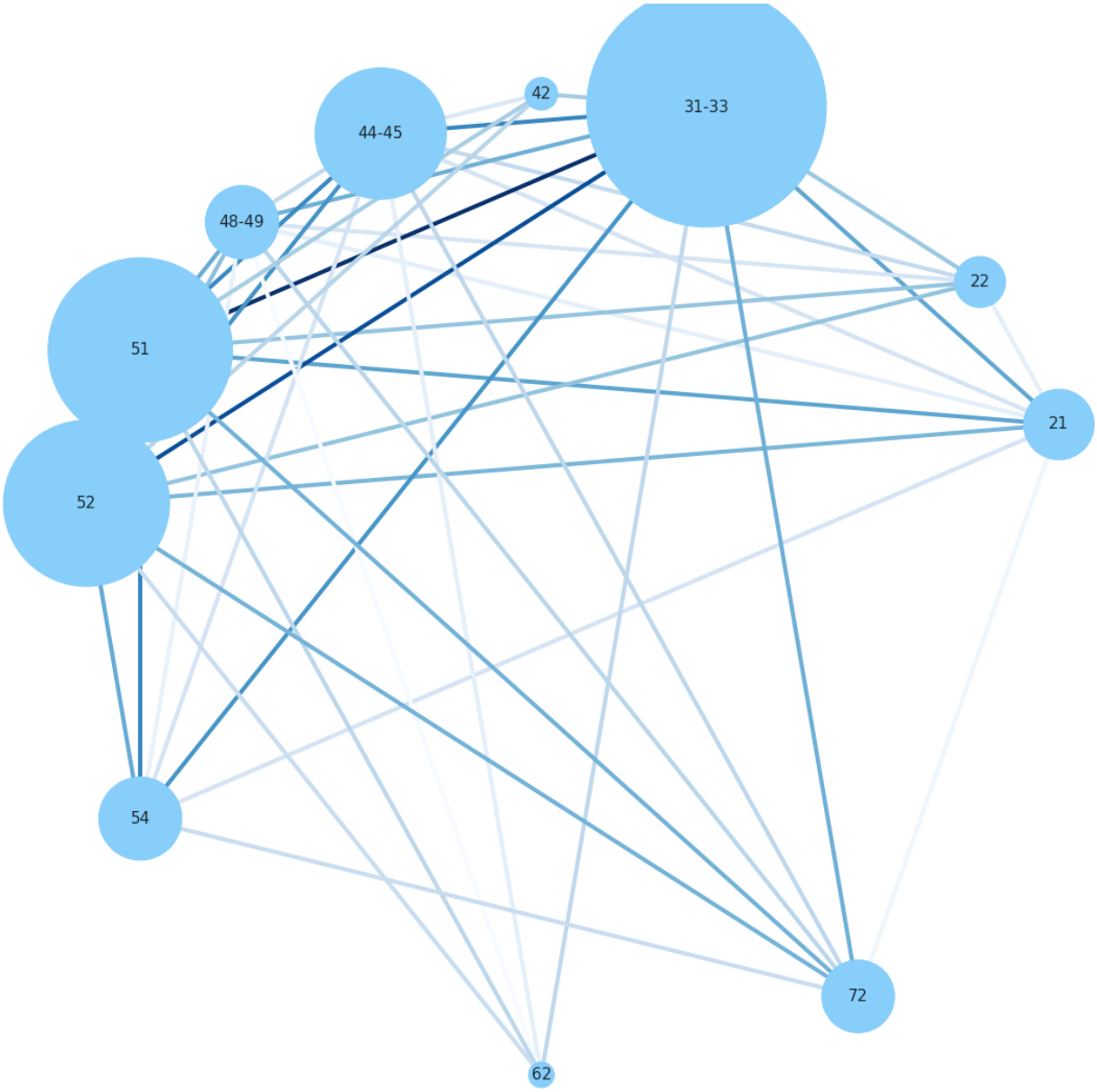


Figure A46 Industry Network Diagram for S&P 500 Firms in 2006

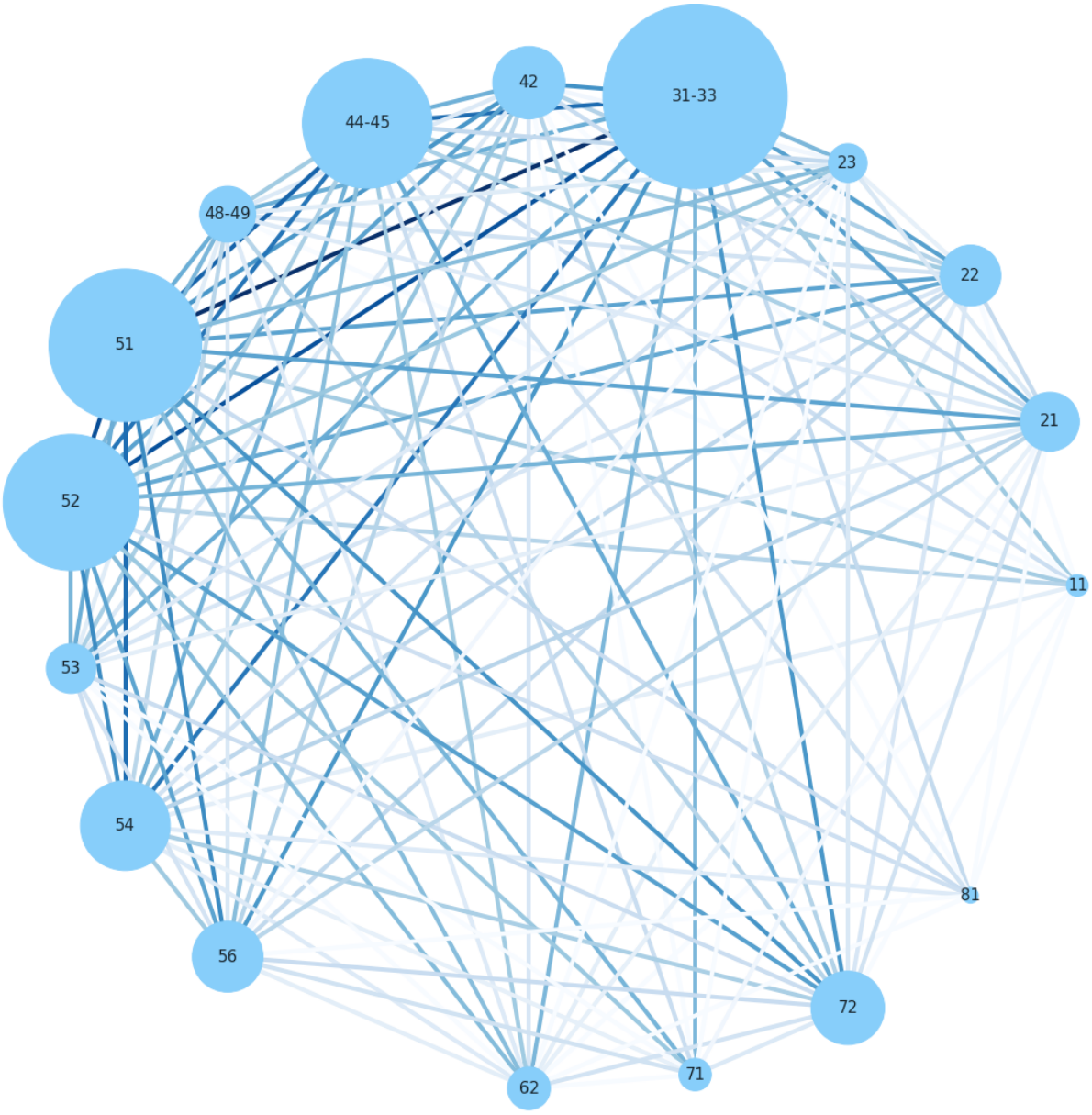


Figure A47 Industry Network Diagram for US Public Firms in 2006

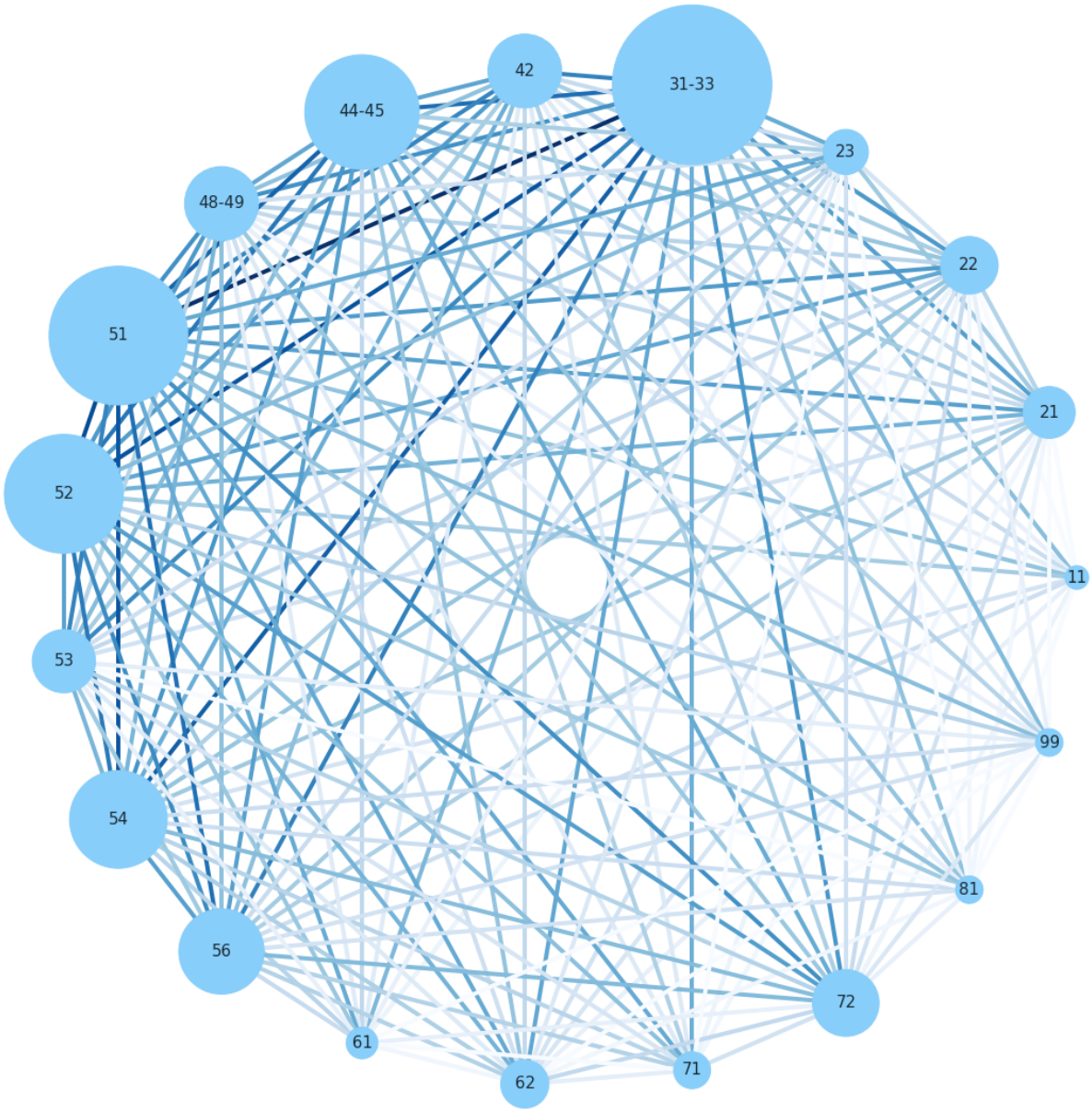


Figure A48 Industry Network Diagram for S&P 100 Firms in 2005

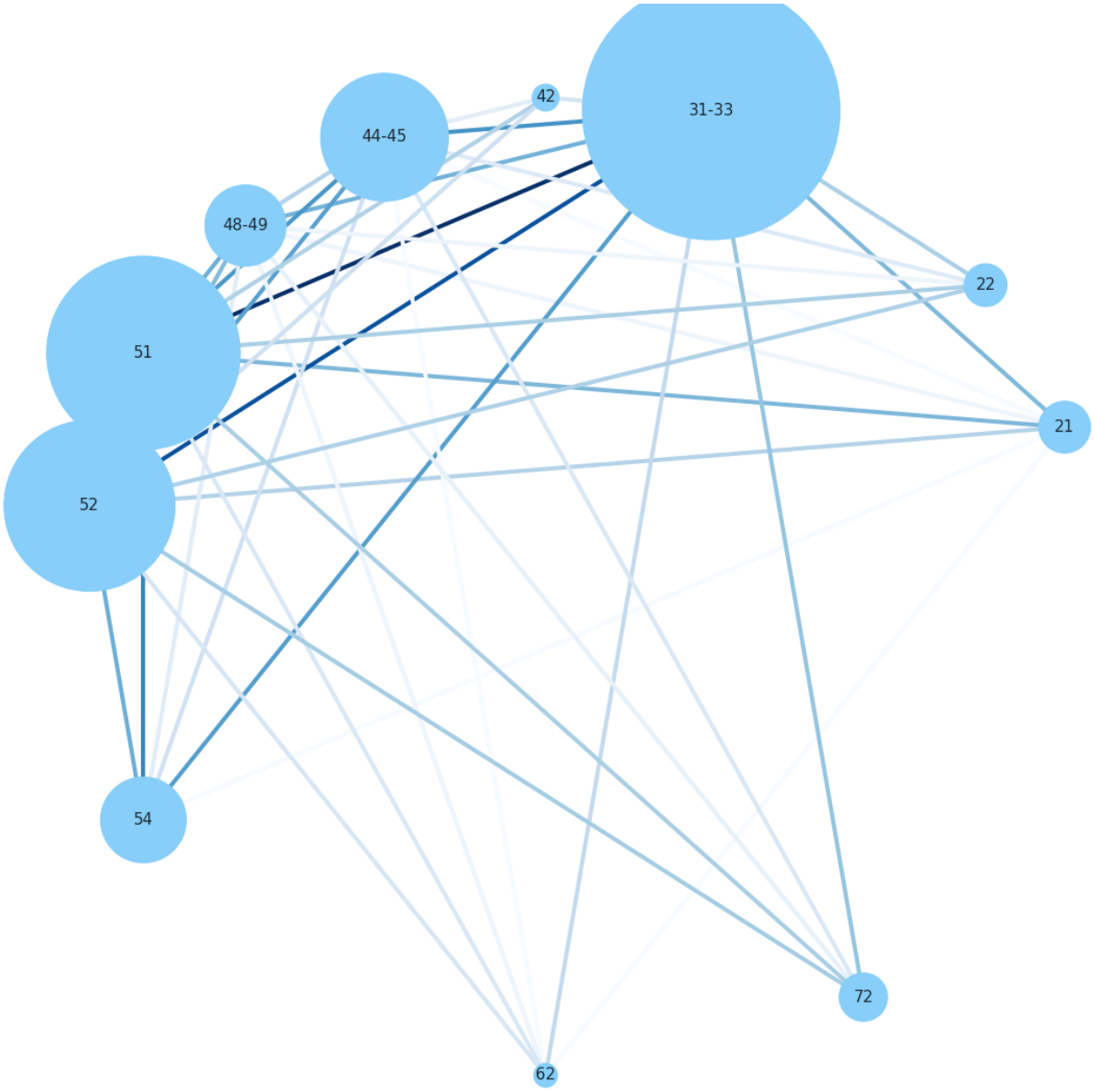


Figure A49 Industry Network Diagram for S&P 500 Firms in 2005

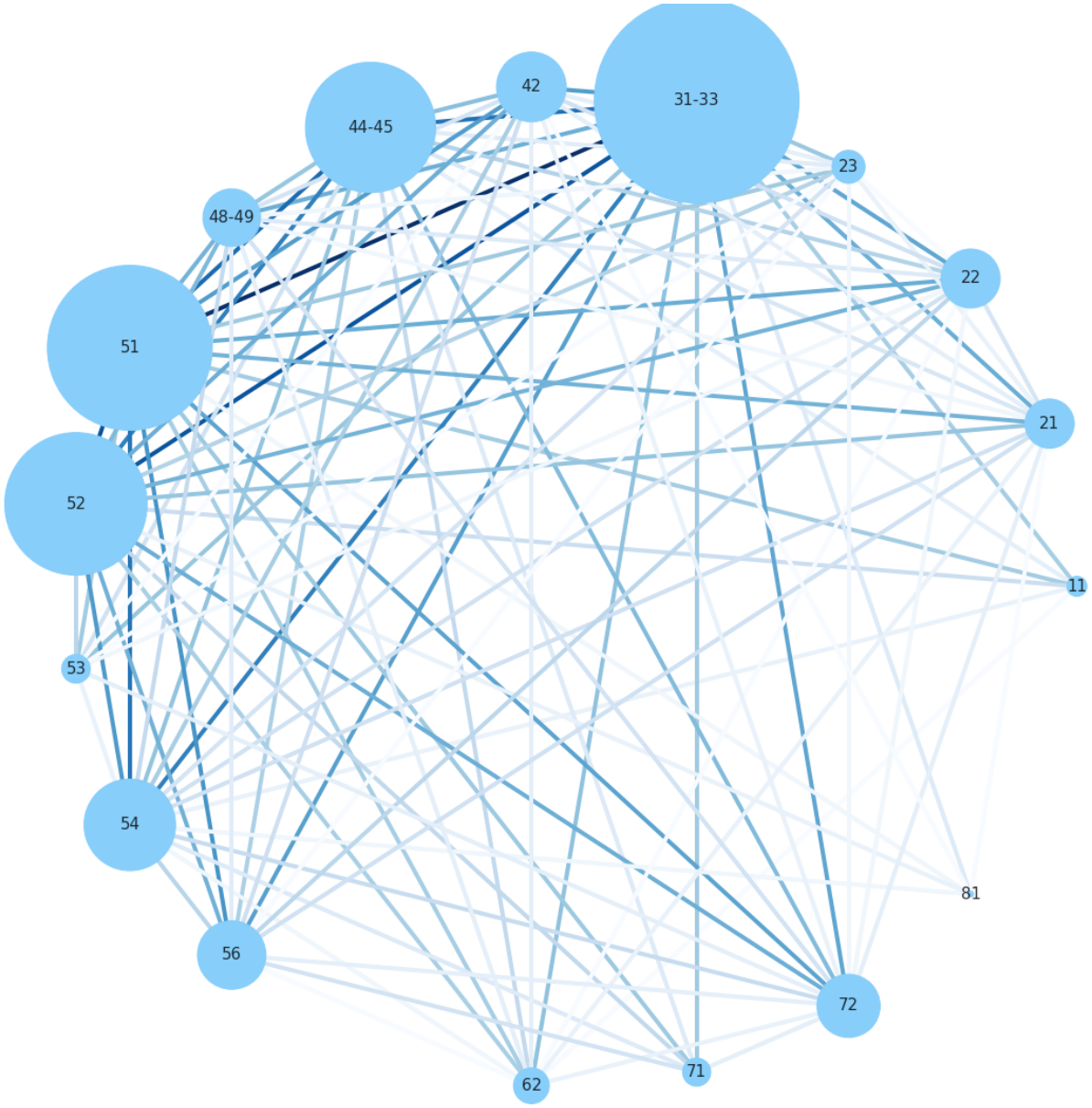


Figure A50 Industry Network Diagram for US Public Firms in 2005

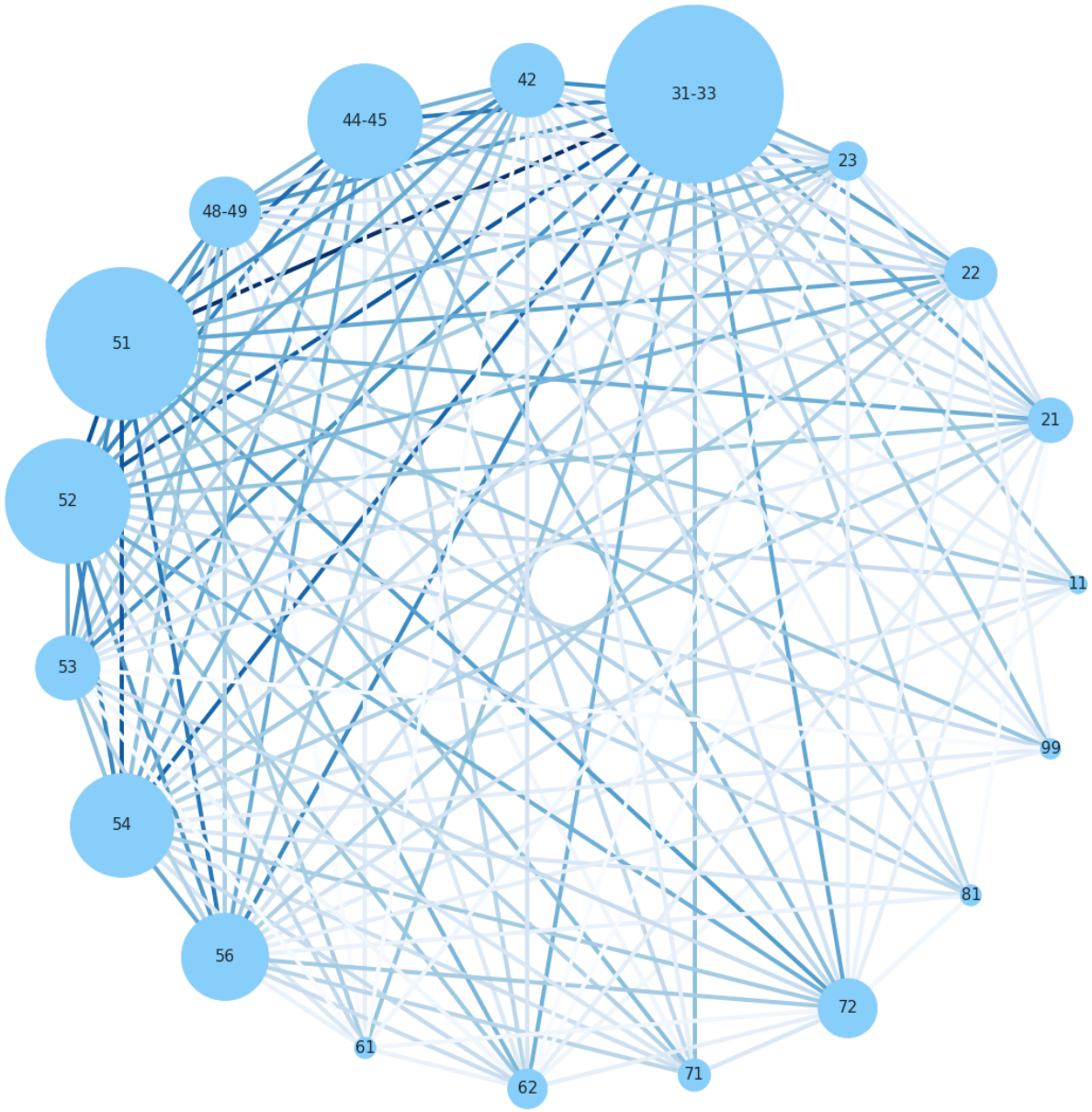


Figure A51 Industry Network Diagram for S&P 100 Firms in 2004

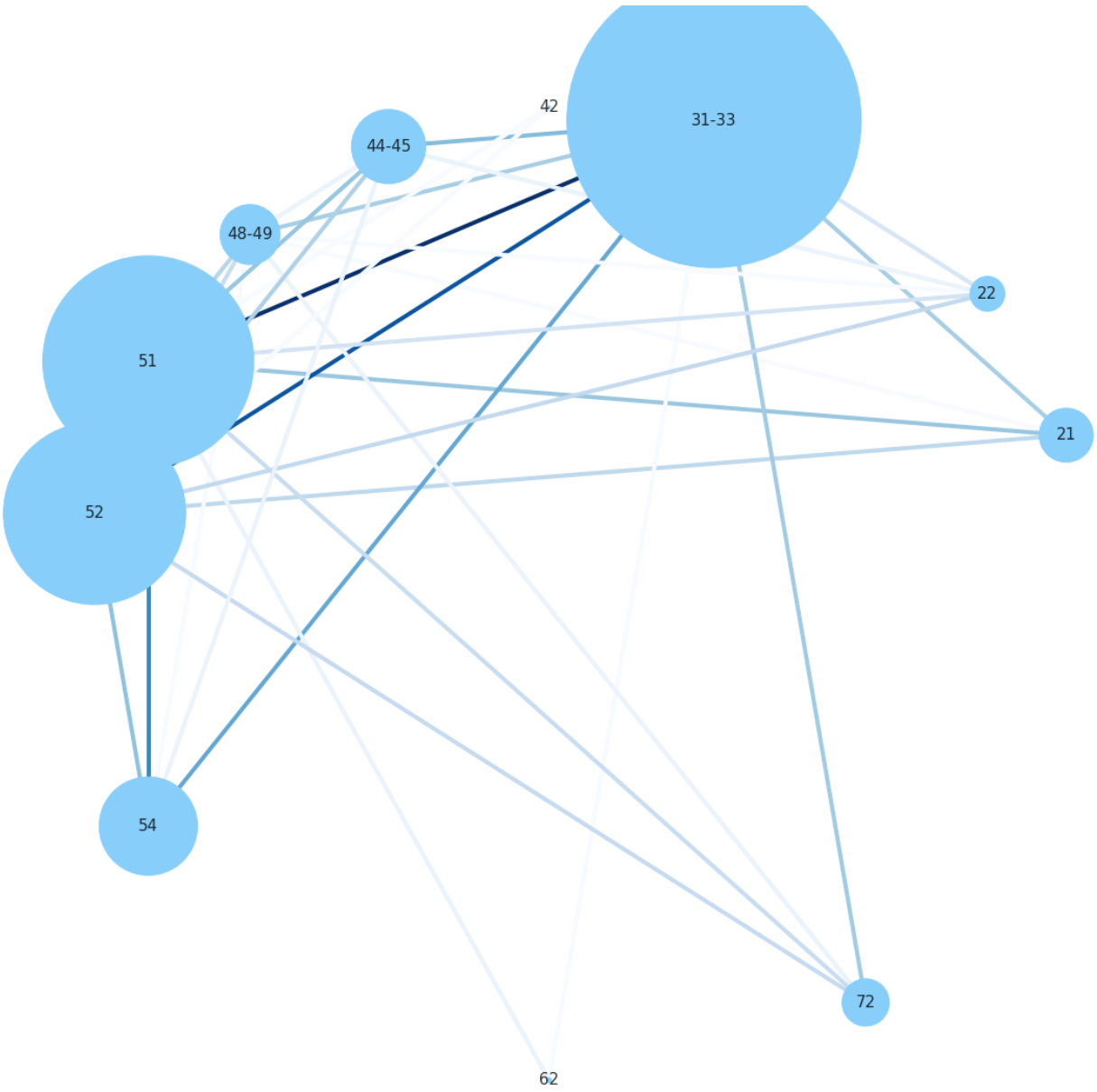


Figure A52 Industry Network Diagram for S&P 500 Firms in 2004

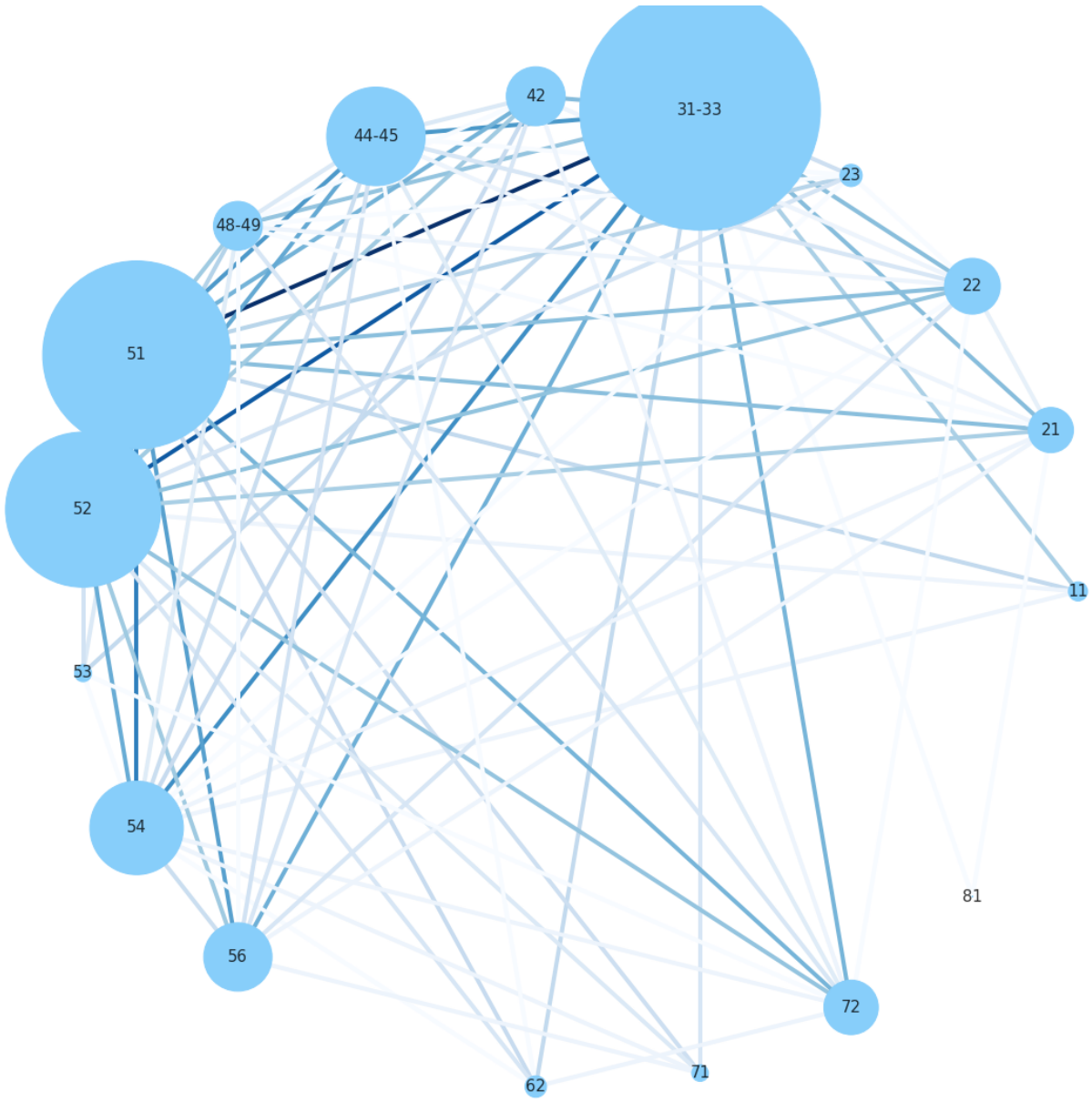


Figure A53 Industry Network Diagram for US Public Firms in 2004

