Oxford Handbooks Online

Visualization of Political Networks

Jürgen Pfeffer

The Oxford Handbook of Political Networks
Edited by Jennifer Nicoll Victor, Alexander H. Montgomery, and Mark Lubell

Subject: Political Science, Political Behavior Online Publication Date: Jan 2017
DOI: 10.1093/oxfordhb/9780190228217.013.13

Abstract and Keywords

Network visualization and political networks have a long history, and some of the earliest and most effective network visualizations have been about power and influence. Now as in the past, network visualization is one of the most effective tools for both exploratory analysis and the communication of scientific results. This chapter discusses the rhetorical, technical, and aesthetic principles that underlie successful network visualizations. The chapter covers automated layout algorithms as well as layouts resulting from the substance of the network. Aspects of visualizing multivariate network data are also discussed. The use of additional visual elements such as color and size is deliberated. The various topics of this chapter are contrasted with issues resulting from human perception and with frequently encountered visualization challenges, such as those encountered when working with dense networks.

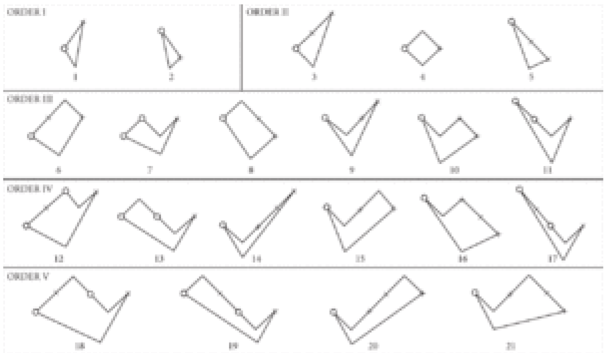
Keywords: network visualization, layout algorithms, substance-based layout, visual elements, human perception, colors

Introduction

The ability to visualize social structures is one of the most obvious advantages of social network analysis. In the context of political networks, researchers are able to use network drawings to visualize in a communicative way the relationships of politicians, parties, and interest groups, as well as political discourses on local, national, and international levels. In "The Development of Social Network Analysis," Freeman (2004) attributes the evolution and growth of network analysis to two factors: the development of metrics and the power of visualizations. This chapter discusses the latter. Visualizations can be very powerful for communicating ideas or research results. Since language and written text require sequential coding, and the amount of information that we can absorb in a given period of time is limited, pictures are an excellent means of communication and are processed *in parallel* at a high *bandwidth* by the human brain. Humans absorb visually thousands of scenes every day. However, we forget most of these visual impressions immediately. The challenge for successful network visualization is therefore twofold. First, one must create network drawings that can be absorbed easily and efficiently and that build on the human ability to process pictures quickly. "The faster the information is understood, the more effective the visualization is" (Krempel, 2005). Second, one must create compelling visualizations that capture the viewer's attention. Even though some network drawings have made it to museums, "efficient communication of information" (Tufte, 2001) with network drawings is not primarily an artistic act. Instead, a better knowledge of visualization algorithms, visual design elements, and human perception can help to create more successful network drawings. This is the motivation for this chapter.

Early network visualizations can be found in medieval times, depicting family trees of biblical figures or of noble families. Early on, visualizations of kin structures were related to policy issues, since they were used to identify possible marriage estoppels. Figure 1 shows a similarly motivated network picture from the nineteenth century. Macfarlane visualized the marriage constellations that were prohibited by British law (1883). At the same time, Hobson used network figures to show that the financial system in South Africa is tightly connected and controlled by a few men (figure 2). Interestingly, this is a very early representation of two-mode network data—that is, networks that consist of people and connections to affiliations (Borgatti and Everett, 1997). Even though network figures have been used over the course of centuries, modern-day network visualizations are intrinsically tied to Jacob Levy Moreno. Moreno introduced his sociograms in 1934 (figure 3), through which he standardized and formalized the drawing of network figures. The focus of this chapter is on sociograms as defined by Moreno, that is, node-and-link diagrams. Other ways of presenting network information, such as matrices (Doreian, Batagelj, and Ferligoj, 2005), are not discussed here. The interested reader is referred to Hennig et al. (2012), as well as Freeman (2000), for an overview of other methods.

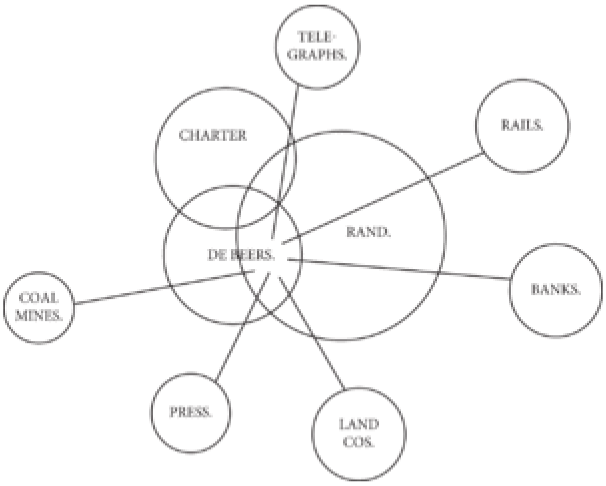
Figure 4 is a great example of how a network picture actually is worth more than a thousand words. Adamic and Glance (2005) analyzed the links and discussions of political bloggers before the US presidential election of 2004. The gray nodes on the right side (red nodes in online colored version) represent conservative bloggers and the black nodes on the left (blue nodes in online colored version) are their liberal colleagues. The network visualization makes it very clear that these two groups of bloggers actually form two tightly knit communities, with a smaller number of connections among members of the two different groups. As a matter of fact, 91 percent of all links in this network stay within the conservative or the liberal community. But just reporting this number would be much less intriguing than the network visualization, which is a perfect representation of a divided blogosphere, or to be more precise, a divided country.



Click to view larger

Figure 1: British marriage prohibition. Males (+), females (o). Lowest points are prohibited offspring.

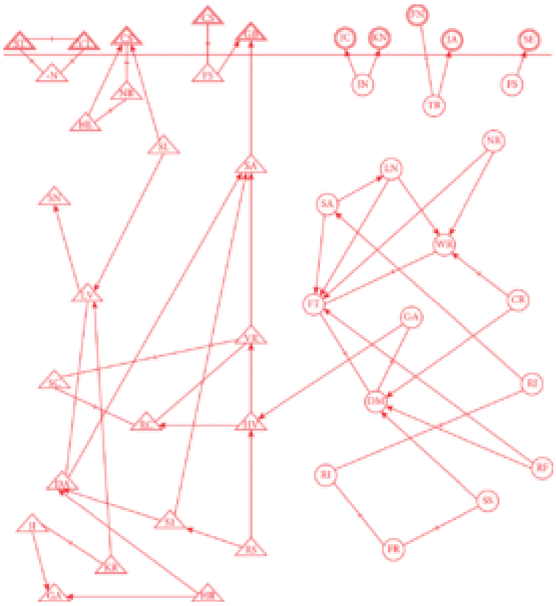
Reproduced from Macfarlane (1883).



Click to view larger

Figure 2: Interlocking corporate directorates showing the inner ring of South African finance.

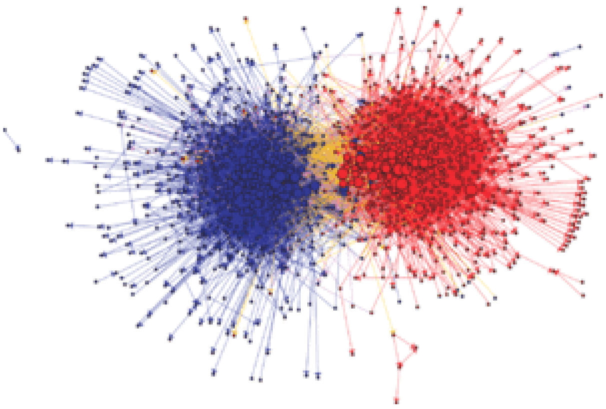
Reproduced from Hobson (1884).



Click to view larger

Figure 3: Girls (circles) and boys (triangles) in a school class. Links represent two best friends. Top line defines group border.

Reproduced from Moreno (1934).



Click to view larger

Figure 4: Political blogs in 2004 US election, red for conservative, blue for liberal. Blue/red links are hyperlinks between blogs within a political camp.

Reprinted with permission from Adamic and Glance (2005).

Similar illustrations that demonstrate political polarization can be found in studies that investigate votes in the US House of Representatives. Over time, analysis of changes in voting behavior within the last decades reveals the rise of partisanship (Andris et al., 2015). Other forms of networks of political actors can be social media networks; for example, visualizations are used to communicate the structure of Twitter interactions among politicians (Cherepnalkoski and Mozetic, 2015) or to provide a better understanding of the dynamics of online political communities (Chu, Wipfli, and Valente, 2011). For international politics, network figures can show countries' sanctions against other countries (Cranmer, Heinrich, and Desmarais, 2014), and Flandreau and Jobst (2005) use networks drawn on maps to show the grouping structure of the international monetary system of the late nineteenth century.

In all these examples, network figures are used to present research outcomes (i.e., to show something). However, network visualizations are not used only at the end of a research project. It is also possible to use pictures to explore data (i.e., to find something). Some researchers visualize their network data at a very early stage of the research process in order to get a first impression. A good picture of a network can help researchers gather information about the network data or reveal coding errors. Moreno also used his sociograms in an exploratory manner: “It is first of all a method of exploration” (1953, 96).

Network visualization is often a neglected part of network analytical projects. Yet drawing networks is a more complex act than merely clicking the “Draw” button of social network analysis tools at the end of a project. When drawing a picture of a social network visualization, aspects have to be considered that can be divided into three major components: substance, design, and algorithm (Brandes et al., 1999).

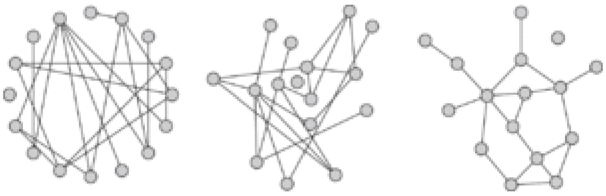
Substance combines all questions concerning the content of the data. The most interesting substance of network data is the information about the connections among the nodes of a network. But normally there is additional information about the actors of a network. If a network consists of political actors, most likely information such as gender, age, or party affiliation will also be available. If a network consists of companies, the data could be about revenue, profit, or number of employees. The interest groups in Box-Steffensmeier’s and Christenson’s analysis of the evolution of amicus curiae networks (2014) are coded with twelve industry codes. In addition to these “real-world” data, computed data can be generated for all nodes of a network, for example, centrality measures. When it comes to visualizing networks, it is important to remember that network data are almost always multivariate. Visualizing high-dimensional information “with clarity, precision, and efficiency” is the foundation for “graphical excellence” (Tufte, 2001).

Design is about mapping the substance of a network to graphic elements. Arranging the nodes (layout) to uncover the structure of the network is the most important task of network visualization. Nevertheless, it is also an interesting challenge to enrich the picture with additional information (substance). The major challenge of this chapter—and also of network visualization in general—is dealing with “effective translation of information to a system of visual elements” (Bertin, 1983). The following sections discuss various design elements of network pictures and review them for their suitability to represent different aspects of network data.

Algorithm ultimately deals with the computer-assisted realization of drawing network pictures. Algorithms must be implemented in network analytical tools, and they have to be efficient enough to deal with the size or complexity of the data and ideas. Another aspect is the effectiveness of an algorithm in respect to the substance of the researcher’s data and research questions. Is the algorithm helpful, or are distracting visual artifacts generated? Knowing about the inner workings of visualization algorithms can help in answering this question. Consequently, this chapter also discusses layout algorithms. Their development and optimization have been a focus of research for decades (Eades, 1984).

Position of Nodes: Network Layout

The layout of a network is the result of the process of positioning the nodes on the surface. This is normally done by layout algorithms that are implemented in network analytical software. Before discussing the concepts of layout algorithms, we must demonstrate the difference a *good* layout can make. Figure 5 shows three different layouts of "Padgett's Florentine Families Network" (Wasserman and Faust, 1994; Kent, 1978). The nodes represent Florentine families in the fifteenth century; the edges are drawn if there is a marriage relationship between two families. In the first (left) picture the nodes are positioned on a circle. In the second picture the position of the nodes is random. The third picture is drawn by applying a distance scaling layout algorithm (see later in this section). When comparing the righthand picture with the two pictures on the left, the viewer can ask about the differences among the figures and why one of them looks "better" or even "prettier."



Click to view larger

Figure 5: A circular, a random, and a springlike layout of the same network.

Many criteria exist to judge whether a graph layout is "good" or not (Fleischer and Hirsch, 2001; Battista et al., 1998). These criteria are "technical" but also "aesthetical," because when looking at many graph pictures, it turns out that these aspects often go hand in hand. Increasing the readability and clarity of the image of a network picture often also makes the picture more attractive. The list of criteria for assessing the quality of a network layout can be summarized by the following points:

- **Show structure.** If a network has a globally organized structure, the layout should uncover it; for example, if there are two almost separate groups, as in figure 4, then the picture should show this structural characteristic. Revealing structural characteristics of networks is the most important reason for applying layout algorithms.

- **Optimize distribution on the surface.** The nodes should be evenly distributed across the whole area and not sit on top of each other. This sounds like a straightforward criterion, but it is all but ignored in most layout algorithms, resulting in areas of the picture with a lot of nodes close to each other (most of the time at the center) surrounded by lots of empty space with almost no nodes.
- **Minimize line crossings, maximize angles, and optimize length of lines.** Line crossing are confusing because they make it more difficult to trace connections. Very acute angles produce overlapping lines, while obtuse angles increase the aesthetics of the layout and also the traceability. If the lines are unweighted, all of them should have approximately the same length in the picture. This is because shorter lines create the impression that nodes are closer to each other than they appear to be when they are connected with a long line.
- **Optimize path distances.** The path distances (how many steps are needed from one node to another one) should be represented in the distances of the nodes in the picture. Therefore, two adjacent nodes should be drawn near each other, while two path-distant nodes should be drawn at a larger distance from one another.

It is all but impossible to meet these criteria perfectly when dealing with complex networks of hundreds or thousands of nodes. Also, the criteria in some respects contradict one another. Nevertheless, they are essential for computer-assisted layout algorithms. As shown in the following discussion, automated algorithms try to optimize one or more of these criteria to create pictures that meet the user's expectation. Layout algorithms can be divided into two groups: distance scaling and classical scaling.

Distance scaling algorithms are very intuitive approaches. The crucial idea of these algorithms is the concept that connected nodes attract each other. That is why these algorithms are called "spring-embedded" (Eades, 1984). The left network (circular) of figure 5 can illustrate this idea. There is a node on the top left position of the circle, with a couple of connections that are all in the lower half of the network. These long connections can be seen as stretched out springs that pull the node on top toward its connected nodes. The best known approaches based on the spring idea were developed by Fruchterman and Reingold (1991) and Kamada and Kawai (1989). These two algorithms, or variations of them, are implemented in almost every network analytical tool. A similar algorithm was used by Andris et al. (2015) to show that the rise of partisanship over time based on congressional covoting pulls the members of the two parties apart. The main disadvantage of distance-scaling algorithms is their iterative process for finding an optimized layout. This means that two layouts of the same network will look different and the outcome of a layout calculation will depend on the starting position of the nodes.

Classical scaling approaches are multidimensional scaling (MDS) algorithms based on the classical MDS of Torgerson (1952). These algorithms produce two-dimensional representations of high-dimensional data (e.g., dissimilarities calculated from path distances of nodes). Classical scaling algorithms find a global maximum for every network. Regardless of the starting position of the nodes, different calculations with the same network arrive at the exact same outcome. This is algorithmically correct, but "wrong" from a readability point of view. Porter et al. (2006) use singular valued decomposition (SVD), a closely related method to MDS, to map members of the US Senate to a two-dimensional visualization. The global optimization of the structure implies the disadvantage of these algorithms, which can also be seen in the SVD figure of the senators: structural equivalent nodes (those nodes with connections to the same other nodes) lie exactly on top of each other, reducing the readability of the figure.

A fundamental problem of all traditional layout algorithms is the computational complexity that leads to long calculation times if the networks become larger. If a network consists of more than a few thousand nodes, then these algorithms need too much time to be practicable. However, in recent years new algorithms have been developed that overcome this limitation. For instance, Brandes and Pich (2007) developed Pivot MDS, which creates layouts based on MDS approximations. With this approach it is possible to calculate layouts of networks with millions of nodes. Combining such an efficient method with optimized distance-scaling approaches (Brandes and Pich, 2009) enables the drawing of *good* pictures under the criteria introduced above, even if the network gets very large.

The layout of a network should be assessed carefully, because the position of the nodes is the most dominant visual element for viewers of the figure when it comes to interpreting the picture (see later in this chapter). Certain positions are automatically seen as more important. For example, if a node is in the center of a round-shaped network, or if a node is on top of a hierarchically shaped network, then these nodes are inevitably associated with greater importance. Other interpretation artifacts are rooted in Gestalt theory (Koffka, 1935). Mcgrath, Blythe, and Krackhardt (1997) showed experimentally that people group nodes of the exact same network into different groups depending on the perceived visual grouping of the nodes. In other words, if nodes are combined and separated slightly from the rest of the network, people will see the group and not the details of the links.

Substance-Based Layout

Even though layout is the most important challenge in network visualization, and this challenge is often handed over to a layout algorithm to reveal the structure of the network, there are a couple of other ways to use position to communicate essential content of network data. Following is a discussion of four ways to lay out the nodes of a network, based on the substance of the network data. In all cases, contextual information of the network nodes is used for the positions. Substance-based layouts constrain the possibilities of using the node positions to express the network structure.

Predefined layout.

Sometimes the nodes of a network are connected to a physical real-world location. The most obvious location is geographic space. Flandreau and Jobst (2005) illustrate their findings about the continental differences in the international monetary exchange system of the late eighteenth and early nineteenth centuries by mapping the network onto a world map. The existence of many links within Europe and almost none within South America, but many connections from South America to Europe, is an easy concept to grasp without any additional coding or labeling of the nodes (the authors do not even use nodes; links are simply sourced and targeted within country borders). Using real-world locations for positioning nodes has been a common practice since the early days of modern network visualization. Roethlisberger, Dickson, and Wright (1939) observed friendship ties and cliques in a factory, and for their visualizations of the social structure the position of the nodes (workers) reflected the location of their workspaces. The visualization of locations can be helpful in answering questions related to the network data and the underlying social forces that are responsible for creating network ties—for example, that people who are located closer to each other have a higher probability of forming relationships (Festinger, Schachter, and Back, 1950).

Status.

In administrations, companies, and other hierarchical organizations, the position of a person on an organizational chart is often used to visualize status. This is a very intuitive way of visualizing the formal hierarchy of an organization, especially because consumers of network figures bring their own experience and “graphical vocabulary,” which influence interpretation (McGrath and Blythe, 2004). In Western culture, the most important person is put at the top of a picture. The *importance* of a person as the determinant of his or her position can also be calculated from centrality scores or contextual information. In the illustrations of the famous network study of Sampson (1968) about a novitiate in a period of change, the position of the novices on the y-axis shows the sum of received positive and negative choices—the most liked novices are positioned at the top of the figure. Whyte (1943) used position to show the hierarchies of the social structure in an Italian slum. Krackhardt (1996) impressively showed the differences in importance of people in a company by comparing the formal organization chart with the advice network; both networks were visualized by putting the more important people at the top of the figure.

Centrality.

The idea of centrality visualization is similar to the status visualizations introduced in the previous paragraph, but the focus of centrality figures is to put the most important node in the middle of the figure—that is, to put the central nodes at the center (Brandes, Kenis, and Wagner, 2003). This approach can be useful when dealing with centrality metrics (Freeman, 1979), as it makes a lot of sense to put, for example, a node with the highest betweenness centrality[1] (Freeman, 1977) *between* the other nodes of the network. For the rest of the nodes, the centrality score is mapped to the distance from the center (pole) of the circular polar coordinate system. Additional concentric circles, standing for levels of centrality, can help with comparison and interpretation. It is important to notice that this is different from automated layouts, in which the most important nodes are also often in the center of the figure. In these cases, a central position results from the structure. Here, centrality could represent any network or non-network metrics. The first known visualizations in which centrality layouts were used in an elaborate way were created by Northway (1940). She used an acceptability measure to determine importance. Another common use of centrality layout can be found in the visualization drawings of ego networks in which a focal node (ego) is in the center and alter egos are placed around it, often based on emotional closeness. When qualitative network analysists (Hollstein, 2011) collect network data, they often ask their interviewees to manually create such circular layouts of their alters (Kahn and Antonucci, 1980). Related to the idea of centrality layouts is the network representation of Moody and Mucha (2013). They mapped the polarization modularity of the US Senate voting similarity network to the y-axis of a network visualization over time to the increasing polarization of US politics.

Attribute grouping.

We can also use attributes of the nodes of the network to structure the layout of the network. This is especially feasible when the network structure is dominated by an attribute; that is, actors are more likely to be connected within the groups and less likely to be connected between the groups. An example for group visualization is the sociogram by Moreno in figure 3. Moreno visualized boys on one side and girls on the other side of the picture. By doing so, he communicated the importance of gender for the connections between the actors of the network. This is assisted by the fact that just one line connects the two groups of nodes with each other. Hargittai, Gallo, and Kane (2007) arranged political bloggers on a circle and put liberal bloggers on the left half of the circle and conservative bloggers on the right half in order to analyze linkages between bloggers with different perspectives. In contrast, the visual grouping of bloggers in figure 4 is the result of a layout algorithm and was not predetermined by the political orientation attribute.

Graphical Elements for Mapping Information

Layout algorithms, as discussed in the previous section, are fundamental for drawing networks. However, drawing a network is more than positioning the nodes. The structural information about the nodes and their relationships is just one dimension of network data. If we focus on information about nodes, network variables are vectors of numbers, with one number for each node. Contextual network variables normally come in two different types. In political networks, *quantitative variables* can be percentage points of election results, donation dollars of companies, and so forth. Quantitative variables can also be the result of centrality calculations, for example, degree centrality (Freeman, 1979) in figure 4. *Nominal variables* represent categorical information, such as gender or party affiliation. The result of grouping algorithms (Newman and Girvan, 2004) can also be seen as a nominal variable, since every node is put in a category. Networks are often multivariate; that is, there are a couple of different variables for every node. Mapping these variables successfully and intuitively to a network drawing is key to a compelling network visualization. In order to do that mapping, we need to identify the list of graphical elements that can be used for network drawings.

The properties of an information visualization were defined by the French cartographer Jacques Bertin (1983) as the *retinal variables*: position, size, color value, color saturation, orientation, shape, and texture. These variables can be processed involuntarily by the human brain within a fraction of a second. For example, we are capable of *automatically* identifying the largest or the darkest node. Mackinlay (1986) ordered Bertin’s visual variables for different types of statistical data. Subsequently, the visual variables were adapted to network visualization (Pfeffer, 2013). Different visual elements are suitable for visualizing quantitative and nominal variables. Table 1 shows the ordering of visual network variables for quantitative and qualitative variables.

Table 1: Graphical Elements and Their Suitability for Mapping Quantitative and Nominal Variables

Type	**Best**				**Worst**
Quantitative	Position	Size	Saturation	Hue	Shape
Nominal	Position	Hue	Saturation	Shape	Size

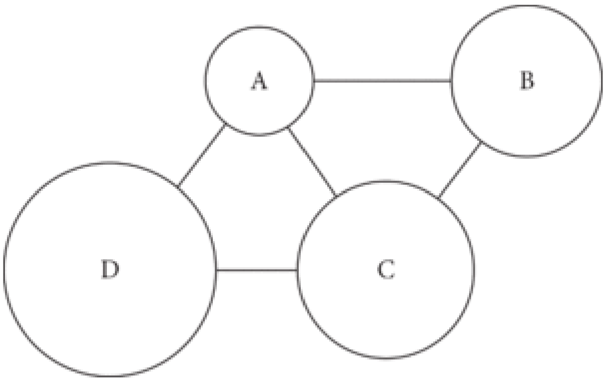
These are the elements discussed in this chapter. As mentioned previously, the most important element is position. As position is the strongest visual variable, it is worthwhile questioning whether the result of an automated layout actually shows enough information (e.g., the global structure of the network) to justify using this visualization element, or if substance-based layouts should be applied.

The visual element *shape* is straightforward, especially as the number of different shapes in network visualization tools is often limited to circles, squares, triangles, and other simple geometric figures. Although shape is often used to visualize different types of nodes—for example, people as circles and squares for affiliation in a two-mode network—the visual element *shape* is surprisingly weak in communicating information. People are not as good at distinguishing among different shapes when many nodes are visualized. The other elements of table 1, namely hue, saturation, and size, are discussed below.

Size

The graphical element *size* is normally used to mark *important* nodes. In figure 4, as well as in most other network visualizations, node size is used to visualize the result of centrality metric calculations. Visually, it is made very obvious that there are larger nodes and smaller nodes and that larger nodes are more important and smaller nodes are less important. However, there are several issues related to mapping a quantitative variable, say a centrality metric, to the size of the node.

Unfortunately, the obvious way to use a quantitative variable to map onto node size is not as straightforward as it might seem. Let us assume that we use circles; we could map a centrality score to the area of a circle. This would be a correct way of doing it. However, human perception puts a spoke in our wheel. It turns out that we are very good at comparing visual information as long as that information is one-dimensional; that is, it is presented as lines. Stevens (1975) defined the psycho-physical power law that describes the inconsistency between actual and perceived differences of two visual (or other) stimuli. Lodge (1981) quantified this perceptual error with experimental studies. Figure 6 and table 2 illustrate these findings. Node B actually is exactly double the size of node A in terms of circle area. However, most people would say that node B is not that big. Lodge (1981) found that on average, people would estimate the scaling factor of these two nodes as about 1.6, while node C, which is actually about 2.7 times larger than node A, is perceived as being only twice the size of node A. Putting the scaling difference not on the area but on the diameter or radius of a node overemphasizes the differences. The radius and the diameter are both one-dimensional information. Consequently, humans can decide accurately that node D has double the radius of node A, but (without mental arithmetic) would perceive the size difference as smaller than 4:1.



Click to view larger

Figure 6: Node size comparison of different scaling factors.

Table 2: Node Sizes with Actual and Perceived Differences in Size and Radius			
Node	**Actual Size**	**Perceived Size**	**Radius**
A	100%	100%	100%
B	200%	163%	141%
C	269%	200%	164%
D	400%	263%	200%

The underlying issue of the size confusion is the fact that network nodes are drawn as two-dimensional objects, while a variable such as degree centrality (Freeman, 1979) is just a one-dimensional variable. As a consequence, some authors (Brandes, Kenis, and Wagner, 2003) suggest using two variables for mapping node size, one for the horizontal and one for the vertical node size, for example, in-degree for the horizontal dimension and out-degree for the vertical dimension of oval-shaped nodes. Using two variables also increases the amount of information that can be communicated with a single network drawing. On the downside, differently shaped ovals take longer to decipher than simple circles or boxes.

Colors

The study of color perception is a small research field, but one with a long tradition (Wyszecki and Stiles, 2000). The human eye has three different color-sensitive photo receptors (cones)—for red, green, and blue (RGB)—that react to different wavelengths of light. The brain analyzes the intensity coming from these three types of receptors and, together with the brightness information coming from a second group of photo receptors (rods), colors are constructed. The famous German writer Johann Wolfgang von Goethe was among the first to be interested in perception of colors. In his “Theory of Colors” (1810), Goethe discussed and criticized Newton’s (1704) discovery that all colors come from colorless white light split up in a prism. More important, Goethe was interested in the perception of colors and introduced his color wheel with complementary colors. Even though Goethe’s theories were criticized by natural scientists early on, and the colors on his color wheel are not well distributed in terms of perceptual differences, his wheel is widely known and used. However, in the early twentieth century Albert Henry Munsell published a color system that set the standard for modern understanding of colors. Munsell’s color system (1912) uses three dimensions: hue (color tint), chroma (saturation), and value (brightness). The great advantage of this system is that it is easy to create perceptually uniform distributed differences in a set of colors and calculate with colors; for example, it is easy to get a color with different saturation levels for color gradients representing quantitative information, which is rather tricky in standard RGB and CMYK (cyan, magenta, yellow, and black) color systems used for monitors and printers. HSL (hue, saturation, lightness) and HSV (hue, saturation, value) are modern color schemes used in graphic tools (e.g., Adobe Photoshop) that are based on Munsell’s colors.

For network visualization, it is important to know that there are two different ways of using colors. Following Munsell, color hue represents the different color tints (e.g. red, green, blue). Color saturation is the intensity of the color (e.g. light or dark red). These two different aspects of color are used for two different types of network variables. Color hue is primarily used for ordinal data. In figure 4, Democratic and Republican bloggers are drawn in different colors, as are the interest groups based on industry codes from Box-Steffensmeier and Christenson (2014). We could also use hue to mark two groups of nodes in two-mode networks (Borgatti and Everett, 1997), for example, politicians as circles and committees as boxes. If we now assume that we do not want to draw more important bloggers in figure 4 with larger nodes, we could use color saturation to represent this quantitative variable; that is, we could draw nodes with high in-degree in dark red/blue and nodes with low in-degree in a very light red/blue. For both parties, we could create a color gradient to cover different in-degree levels.

The use of colors faces a couple of obstacles. For readers of the printed book version of this chapter, the blogger network in figure 4 reveals the most obvious limitation of using colors: many journals and books are printed only in black and white, so colors are transformed to grayscales, thus losing most of their communicative power. Even if it is compelling to create network drawings with colors, it is important to be capable of visualizing informative and compelling network drawings without colors. Also, colors often have culturally loaded meanings. One reason figure 4 is memorable is the traditional color coding of the nodes and links representing political parties: red for conservatives and blue for liberals. Using colors that are counterintuitive or unfamiliar makes reading and understanding incomparably harder and can easily lead to frustration caused by visually jumping back and forth between the figure and the legend. But colors are globally not unambiguous, and the interpretation of colors is culturally dependent. In most European countries, the color red is used for Social Democratic parties and not for conservatives, as it is in the United States. Comparing the visualizations of the blogger network of Adamic and Glance (2005) with figures of European political networks on Twitter from Ausserhofer and Maireder (2013) can lead to false associations. Or imagine the terms “positive,” “negative,” “love,” “nature,” “death,” or “money.” Certain colors tend to be automatically associated with these terms. But all of these terms are mapped onto different colors in different parts of the world.

Finally, it is important to know that almost 10 percent of people are color-blind—mostly men—and the vast majority of them have problems distinguishing between red and green. Color vision deficiencies are normally inherited and are untreatable. Color-blind people are incapable of seeing the color hue differences of the affected colors.

Links

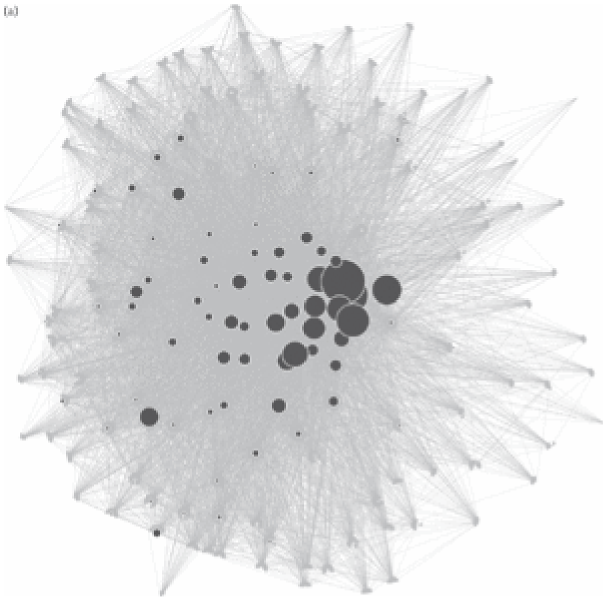
The above-mentioned graphical elements can also be used to map information about the links of a network to a network figure. In sociograms, links representing connections between pairs of nodes are drawn as lines between those nodes. A simple line connecting two nodes communicates binary univariate information; that is, there is a link between these two nodes. This is sufficient as long as the network is unweighted and undirected. If the network is directed, two bits of information need to be coded with the link, as connections can go in either or both directions of the dyad. Moreno (1934; figure 3), Adamic and Glance (2005; figure 4), and most other researchers use arrow heads at the end of a line to indicate the target of a directed edge (arc). The case of directed links between a pair of nodes that go in both directions (e.g., blog A links to blog B and blog B links to blog A) can be depicted by different visual elements. In figure 4, arrow heads can be found on both ends of the connecting link, while Moreno (1934) omitted the arrow heads for bidirectional links and added a small line at a right angle to the middle of the link. The most straightforward approach to handling bidirectional directed links is to just omit the arrow heads, leaving a simple line.

Layout.

In the aforementioned criteria about the quality of network visualization, *minimizing line crossings* was mentioned. However, this criterion was discussed solely from the perspective of rearranging nodes (that are connected to these lines). Lines can also be manipulated directly to optimize a network layout. The most common visual feature in this context is using curved instead of straight lines. While an exclusive usage of curved lines, as seen in many contemporary network visualizations, might be appealing from an aesthetic perspective, it decreases readability and increases visual complexity in most instances. More sophisticated approaches try to optimize a network layout by routing links individually to reduce visual complexity, for example, by avoiding drawing links in a way that they overlap with nodes. For dense networks, link bundling is another approach to increasing the readability of a network figure. This is accomplished by merging links that all go from one area of the network to another area. By removing some level of already hard to read detail from the visualization—that is, which nodes are connected to which other nodes—we render additional macro information to the network, showing areas that are connected.

Size.

In weighted networks, links are not just defined by their absence or presence, but by link weights. Link weights can come from two sources. First is direct observations; for example, the link weights in Subramanian and Wei's (2007) represent the volume of trade between pairs of countries. Second is from two-mode network data (Borgatti and Everett, 1997) in the case of shared activities or affiliations. Connections among politicians can be drawn based on the overlap of participation in committees and panels, with the number of shared affiliations as link weight. In any case, handling these different weights is straightforward and can be seen in figure 7; larger weights result in stronger lines. However, drawing different-sized links is limited to small and sparse networks. In figure 7a no differences in link weights are identifiable.



Click to view larger

Figure 7: Different representations of an international trade network. Nodes are countries or continents. Different width of links represents trade volume.

Figures created with data from Subramanian and Wei (2007): a) original network with all links; b) global threshold visualization; c) local threshold visualization; d) partially aggregated network; e) focus/context visualization of network.

Color hue.

Using colors to mark different types of links was used early on. Moreno (1934) drew red links for positive ("attraction") and black links for negative ("repulsion") relationships. Adamic and Glance (2005) use colors on links to show hyperlinks connecting blogs within political camps (see figure 4). The difference between these two examples is that Moreno mapped dyadic information to the links, while Adamic and Glance derived link colors from source and target nodes of the links. In both cases, link colors can be very useful to help interpretation or to highlight interesting actors or areas of the network. In Moreno's visualization of a football team (1934; not shown in this article), the colored links reveal very clearly one unpopular player. In dense networks, such as in figure 4, it is impossible to follow or interpret single links; instead, links can create visual areas that can be important for interpretation of the analysis. In Adamic and Glance's (2005) figure, colored areas consisting of hundreds of links amplify the segregation of the political camps. Another approach of using node information to color links is very similar. The grouping of the nodes is not derived from contextual information, but rather the result of clustering (Newman, 2006). In all these cases, colors for links are used to distinguish among different classes of relations. Along the lines of using different color hues to color different groups of nodes, we applied the same logic to links.

Color saturation.

To visualize quantitative information of links, we can turn to color saturation. The thicker links in figures 7b and 7d can also be drawn in black, while the thinner links are in different scales of gray. This is particularly useful in very dense networks. It is also important to sort the lines so that thicker lines are drawn on top of thinner lines, so that they are actually visible. The network for committees and subcommittees of the House of Representatives in Porter et al. (2006) serves as a good example of using sorted grayscaled links to emphasize important links in a dense and weighted network.

Reducing Visual Complexity

Looking at figure 4, but also at many other policy networks—for example, the congressional voting network (Andris et al., 2015) or the world trade network from Subramanian and Wei (2007)—one can see one of the major challenges to visualizing networks: handling dense networks, in other words, a large number of links. On the upside, these links can be used as visualization elements. This is the case in the blogger network in figure 4; hundreds of red and blue links create colored regions in the background of the nodes. This *macro* approach is reasonable when, as in the blogger network, the message of the network visualization is also on the macro level. If the level of interest is more detailed on the node level, approaches are needed that reduce the visual complexity of network visualizations.

As a rule of thumb, networks having at most two to three times more links than nodes are favored in terms of readability and aesthetics. This is much less than we find in many empirically collected networks. Consequently, reducing visual complexity is intrinsically tied to removing links (but also nodes). While removing parts of the network can result in a much easier to read network figure, it is important to realize that doing this could change the structure of a network and create visual artifacts. Comparing different reduction strategies or different removal thresholds can help mitigate this risk.

For the following approaches, we use the world trade network data from Subramanian and Wei (2007). The network (figure 7a.) represents 157 countries (nodes) with trade relationships (edges) among them. As one can easily see, there are a lot of lines in this figure, since almost every country has trade relations to a large number of other countries.

Global threshold.

When the network is weighted (each line is described by a value), we can remove all lines lower than a defined value (threshold). This threshold can be derived from the data (e.g., more than $1 billion trade volume) or from the readability of the figure (increase threshold until a clear structure is visible). In figure 7b the latter has been done. As one can see, the number of nodes is also reduced. This results from the fact that deleting lines with small line values may create isolates (nodes with no remaining links).

Normally, such isolates are removed from the network visualization. The focus of such an approach is clearly on the center of the network, since only the most important nodes and links remain in the figure. In a detailed analysis of their blogger network, Adamic and Glance (2005) removed weak links, which made the political separation even clearer.

Local threshold.

An alternative approach would be to remove all but the most important lines for each node. With this approach, all nodes are preserved. The disadvantage of this approach is that globally important connections get lost. For example, in figure 7c we just show the top two links for each country. However, the third most important link of the United States has a much higher weight than all of the links of most other countries.

Node aggregation.

Another approach for reducing the complexity of a network can be used when a grouping attribute of the nodes is available. In our case, countries can be grouped to continents. Consequently, we can merge multiple country nodes (and their links) to a single continent node. In figure 7d we did this for all countries except Germany and the United States. Node aggregation can also be applied after calculating communities in networks (Newman, 2006). Porter et al. (2006) identified communities in the US House of Representatives based on common committee memberships. They then aggregated the communities to one node each and showed the connections among the communities. Moreover, as a single group-node now represented multiple politicians, the authors mapped sets of attributes of these actors to pie charts, which replace the nodes in the network visualization.

Focus/context.

If the viewer's attention should be directed to a specific part of the network while the rest of the network should still be visible for orientation, a focus/context visualization can be used. Figure 7e represents a zoom into a specific area of figure 7c, while the overall network is still visible, so that we can see which part of the network is shown. Here we focus on the subnetwork of South American countries and their connections, showing that most of them have their strongest connections within their own continent.

Summary

This chapter discussed various aspects of visualizing political networks. We discussed layout algorithms, but also argued that visualizing informative networks consists of more steps than optimizing the layout of the network nodes. Being aware of graphical elements such as size, shape, and color and knowing about how to use them to represent different forms of variables in a visualization can create informative network drawings that communicate a large amount of information. Besides the technical and perceptual challenges, the value of a visualization could be evaluated by its narrative quality. A compelling network visualization, like all other information visualizations, tells a story to support scientific findings. Moreover, it is always important to remember that network visualization is about visualizing information, not just data. Excellent network graphics “induce the viewer to think about the substance rather than the methodology” (Tufte, 2001).

References

Adamic, L. A., and Glance, N. (2005). “The Political Blogosphere and the 2004 U.S. Election: Divided They Blog.” In *Proceedings of the 3rd International Workshop on Link Discovery*, pp. 36–43. New York: Association for Computing Machinery. doi: 10.1145/1134271.1134277.

Andris, C., Lee, D., Hamilton, M. J., Martino, M., Gunning, C. E., and Selden, J. A. (2015). “The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives.” *PLOS ONE* 10(4): e0123507. doi:10.1371/journal.pone.0123507.

Ausserhofer, J., and Maireder, A. (2013). “National Politics on Twitter: Structures and Topics of a Networked Public Sphere.” *Information, Communication & Society* 16(3): 291–314.

Battista, G. D., Eades, P., Tamassia, R., and Tollis, I. G. (1998). *Graph Drawing: Algorithms for the Visualization of Graphs*. 1st ed. Upper Saddle River, NJ: Prentice Hall PTR.

Bertin, J. (1983). *Semiology of Graphics: Diagrams, Networks, Maps*. Madison: University of Wisconsin Press.

Borgatti, S. P., and Everett, M. G. (1997). “Network Analysis of 2-Mode Data.” *Social Networks* 19(3): 243–269.

Box-Steffensmeier, J. M., and Christenson, D. P. (2014). “The Evolution and Formation of Amicus Curiae Networks.” In “*Political Networks*,” special issue, Social Networks 36 (January): 82–96. doi:10.1016/j.socnet.2012.07.003.

Brandes, U., Kenis, P., Raab, J., Schneider, V., and Wagner, D. (1999). “Explorations into the Visualization of Policy Networks.” *Journal of Theoretical Politics* 11(1): 75–106.

Brandes, U., Kenis, P., and Wagner, D. (2003). “Communicating Centrality in Policy Network Drawings.” In *IEEE Transactions on Visualization and Computer Graphics* 9 (2): 241–253.

Brandes, U., and Pich, C. (2007). “Eigensolver Methods for Progressive Multidimensional Scaling of Large Data.” In *Proceedings of the 14th International Symposium on Graph Drawing (GD ’06)*, pp. 42–53. Karlsruhe, Germany.

Brandes, U., and Pich, C. (2009). “An Experimental Study on Distance-Based Graph Drawing.” In *Graph Drawing*, edited by I. G. Tollis and M. Patrignani, pp. 218–229. Lecture Notes in Computer Science 5417. Berlin, Heidelberg: Springer.

Cherepnalkoski, D., and Mozetic, I. (2015). “A Retweet Network Analysis of the European Parliament.” In *Proceedings of the 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pp. 350–357 Bangkok.

Chu, K.-H., Wipfli, H., and Valente, T. W. (2011). “Using Visualizations to Explore Network Dynamics.” *Journal of Social Structure: JOSS* 14 (4). **https://www.cmu.edu/joss/content/articles/volume14/ChuWipfliValente.pdf**.

Cranmer, S. J., Heinrich, T., and Desmarais, B. A. (2014). “Reciprocity and the Structural Determinants of the International Sanctions Network.” In *Political Networks*, special issue, Social Networks 36 (January): 5–22. doi:10.1016/j.socnet.2013.01.001.

Doreian, P., Batagelj, V., and Ferligoj, A. (2005). *Generalized Blockmodeling*. Cambridge, UK: Cambridge University Press.

Eades, P. (1984). “A Heuristic for Graph Drawing.” *Congressus Numerantium* 42 (11): 149–160.

Festinger, L., Schachter, S., and Back, K. (1950). “The Spatial Ecology of Group Formation.” In *Social Pressure in Informal Groups*, edited by L. Festinger, S. Schachter, and K. Back, ch. 4. Cambridge, MA: MIT Press.

Flandreau, M., and Jobst, C. (2005). “The Ties That Divide: A Network Analysis of the International Monetary System, 1890–1910.” *Journal of Economic History* 65(4): 977–1007. doi:10.1017/S0022050705000379.

Fleischer, R., and Hirsch, C. (2001). “Graph Drawing and Its Applications.” In *Drawing Graphs*, edited by M. Kaufmann and D. Wagner, pp. 1–22. Lecture Notes in Computer Science 2025. Berlin, Heidelberg: Springer. **http://link.springer.com/chapter/10.1007/3-540-44969-8_1**.

Freeman, L. C. (1977). “A Set of Measures of Centrality Based on Betweenness.” *Sociometry* 40: 35–41.

Freeman, L. C. (1979). “Centrality in Social Networks: Conceptual Clarification.” *Social Networks* 1(3): 215–239.

Freeman, L. C. (2000). “Visualizing Social Networks.” *Journal of Social Structure* 1(1). **http://www.cmu.edu/joss/content/articles/volume1/Freeman.html**.

Freeman, L. C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver, BC: Empirical Press; North Charleston, SC: BookSurge.

Fruchterman, T. M. J., and Reingold, E. M. (1991). “Graph Drawing by Force-Directed Placement.” *Software: Practice and Experience* 21(11): 1129–1164.

Hargittai, E., Gallo, J., and Kane, M. (2007). “Cross-Ideological Discussions among Conservative and Liberal Bloggers.” *Public Choice* 134(1–2): 67–86. doi:10.1007/s11127-007-9201-x.

Hennig, M., Brandes, U., Pfeffer, J., and Mergel, I. (2012). *Studying Social Networks: A Guide to Empirical Research*. Frankfurt: Campus Verlag.

Hobson, J. A. (1884). *The Evolution of Modern Capitalism: A Study of Machine Production*. London, New York: Allen & Unwin.

Hollstein, B. (2011). “Qualitative Approaches.” In *The SAGE Handbook of Social Network Analysis*, edited by J. Scott, P. J. Carrington, pp. 404–417. London: SAGE Publications.

Kahn, R. L., and Antonucci, T. C. (1980). “Convoys over the Life Course: Attachment, Roles, and Social Support.” In *Life-Span Development and Behavior*, edited by Paul B. Baltes and Orville G. Brim, pp. 254–287. New York: Academic.

Kamada, T., and Kawai, S. (1989). “An Algorithm for Drawing General Undirected Graphs.” *Information Processing Letters* 31: 7–15.

Kent, D. (1978). *The Rise of the Medici: Faction in Florence, 1426–1434*. Oxford, New York: Oxford University Press.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace & World.

Krackhardt, D. (1996). “Social Networks and the Liability of Newness for Managers.” In *Trends in Organizational Behavior*, edited by C. L. Cooper and D. M. Rousseau, pp. 159–173. Weinheim, Germany: John Wiley & Sons.

Krempel, L. (2005). *Visualisierung komplexer Strukturen: Grundlagen der Darstellung mehrdimensionaler Netzwerke*. Auflage 1. Sonderband Auflage. Frankfurt am Main: Campus Verlag.

Lodge, M. (1981). *Magnitude Scaling, Quantitative Measurement of Opinions*. Beverly Hills, CA: Sage Publications.

Macfarlane, A. (1883). “Analysis of Relationships of Consanguinity and Affinity.” *Journal of the Anthropological Institute of Great Britain and Ireland* 12: 46–63. doi: 10.2307/2841840.

Mackinlay, J. (1986). “Automating the Design of Graphical Presentations of Relational Information.” *ACM Transactions on Graphics* 5(2): 110–141. doi:10.1145/22949.22950.

McGrath, C., and Blythe, J. (2004). “The Effects of Motion and Hierarchical Layout on Viewers’ Perceptions of Graph Structure.” *Journal of Social Structure* 5(2):. **http://www.cmu.edu/joss/content/articles/volume5/McGrathBlythe/**.

McGrath, C., Blythe, J., and Krackhardt, D. (1997). “The Effect of Spatial Arrangement on Judgement and Errors in Interpreting Graphs.” *Social Networks* 19(3): 223–242.

Moody, J., and Mucha, P. J. (2013). “Portrait of Political Party Polarization.” *Network Science* 1(1): 119–121. doi:10.1017/nws.2012.3.

Moreno, J. L. (1934). *Who Shall Survive? A New Approach to the Problem of Human Interrelations*. Washington, DC: Nervous and Mental Disease Publishing Co.

Moreno, J. L. (1953). *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama*. New York: Beacon House.

Munsell, A. H. (1912). “A Pigment Color System and Notation.” *American Journal of Psychology* 23(2): 236–244. doi:10.2307/1412843.

Newman, M. E. J. (2006). “Modularity and Community Structure in Networks.” *Proceedings of the National Academy of Sciences of the United States of America* 103(23): 8577–8582. doi:10.1073/pnas.0601602103.

Newman, M. E. J., and Girvan, M. (2004). “Finding and Evaluating Community Structure in Networks.” *Physical Review E* 69(2): 026113. doi:10.1103/PhysRevE.69.026113.

Newton, I. (1704). *Opticks: Or, a Treatise of the Reflections, Refractions, Inflections, and Colors of Light*. London: Royal Society.

Northway, M. L. (1940). “A Method for Depicting Social Relationships Obtained by Sociometric Testing.” *Sociometry* 3(2): 144–150. doi:10.2307/2785439.

Pfeffer, J. (2013). “Fundamentals of Visualizing Communication Networks.” *Communications, China* 10(3): 82–90. doi:10.1109/CC.2013.6488833.

Porter, M. A., Mucha, P. J., Newman, M. E. J., and Friend, A. J. (2006). “Community Structure in the United States House of Representatives.” *Chaos: An Interdisciplinary Journal of Nonlinear Science* 16(4): 041106. doi:10.1063/1.2390556.

Roethlisberger, F. J., Dickson, W. J., and Wright, H. A. (1939). *Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago*. Cambridge, MA: Harvard University Press.

Sampson, S. F. (1968). “A Novitiate in a Period of Change: An Experimental and Case Study of Social Relationships.”, Thesis (Ph.D.), Cornell University.

Stevens, S. S. (1975). *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*. John Wiley & Sons, New York.

Subramanian, A., and Wei, S.-J. (2007). “The WTO Promotes Trade, Strongly but Unevenly.” *Journal of International Economics* 72: 151–175.

Torgerson, W. S. (1952). “Multidimensional Scaling.” *Theory and Method* 17: 401–419.

Tufte, E. R. (2001). *The Visual Display of Quantitative Informations*. 2d ed. Cheshire, CT: Graphics Press.

Wasserman, S., and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, MA: Cambridge University Press.

Whyte, W. F. (1943). *Street Corner Society: The Social Structure of an Italian Slum*. Chicago: University of Chicago Press.

Wyszecki, G., and Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae*. 2d ed. New York: Wiley-Interscience.

Notes:

([1]) Betweenness centrality (Freeman 1977) measures the extent to which a node is an intermediate on the shortest path connections between pairs of other nodes.

Jürgen Pfeffer

Bavarian School of Public Policy, Technical University of Munich