ELSEVIER

Social Networks 19 (1997) 223–242

SOCIAL
NETWORKS

The effect of spatial arrangement on judgments and errors in interpreting graphs

Cathleen McGrath [a], Jim Blythe [b,1], David Krackhardt [a]

[a] *Heinz School of Public Policy and Management, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA*

[b] *School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA*

Abstract

The spatial arrangement of social network data in graphs can influence viewers' perceptions of structural characteristics such as prominence, bridging and grouping. To study the extent of this effect, we conducted an experiment with 80 graduate students. Each student viewed three of five different spatial arrangements of the same network. We found that viewers' perceptions of structural features of the network changed as the spatial arrangement of the network changed. © 1997 Elsevier Science B.V.

Keywords: *Spatial arrangements; Graphs; Perceptions; Experiment*

1. Introduction and problem statement

It is commonly assumed that graphs communicate important characteristics of network data. In fact, the beginning of the field of social networks can be found in Moreno's use of the sociogram to communicate the power of network ideas (Moreno, 1953). As Harary points out, ''It is very useful to have diagrams of graphs available for the accumulation of data leading to conjectures'' (Harary, 1969, p. 213). When actors in networks are presented as nodes in a graph and the relationships between them are displayed as lines connecting the nodes, the graph conveys all the information contained in the adjacency matrix and so the two are 'informationally equivalent' (Larkin and Simon, 1987). However, the graph can also use Euclidean spatial relations to highlight the relationships among actors. Thus, drawing network data in a graph may influence individuals' perceptions of social network attributes completely determined by the structure of the network, such as the existence of subgroups and the relative centrality of actors.

[1] Also at 5000 Forbes Avenue, Pittsburgh, PA 15213, USA.

In this paper we explore the influence of the spatial arrangement of networks on individuals' perceptions of common social network measures, in particular 'prominence' and 'bridging', which are two facets of centrality, and grouping. Prominence deals with a node's popularity and most closely relates to the degree of the node. Importance as a bridge deals with a node's strategic positioning between groups, and most closely relates to betweenness centrality (Freeman, 1978). By spatial arrangement we mean a mapping from the nodes of the graph to $\mathscr{R} \times \mathscr{R}$ (the real plane).

A large body of work exists on computer programs to draw graphs according to certain 'aesthetics' (see Di Battista et al., 1994, for a survey of this work). However, almost all of this work considers aesthetics that attempt to improve graph readability from a very general point of view without considering specific applications, and uses general aesthetics such as the regular spacing of nodes and minimization of edge crossings.

Only a few studies explicitly question the aesthetics in use and there has been very little work on analyzing how well they actually improve the information conveyed in a graph's spatial arrangement. Ding and Mateti (1990) consider the subjective factors that go into the drawing of a diagram intended to explain data structures in computer programs. They produce their factors by examining the pictures that appear in a number of text books in that field. Batini et al. (1985) make an experimental study of the aesthetics used in entity relationship diagrams from the field of software engineering. We are not aware of any studies that consider the information one might want to convey in a social network spatial arrangement.

In this paper we present an empirical study of how the spatial positioning of nodes in the graph influences individuals' perceptions of network characteristics. We build on previous work on graph drawing and apply it specifically to drawing graphs for social networks. We find that individuals' perceptions of graphs may depend on both the structure of the network itself, that is the pattern of ties between nodes, and the spatial arrangement of nodes.

In the rest of this section we describe the graph theoretical and Euclidean features of a spatial arrangement that might influence the values an individual assigns to measures of prominence, bridging, and grouping while looking at a graph. In the subsequent sections we describe our experimental design and the results of the study.

Our experiment was conducted using five different spatial arrangements of one particular network having twelve actors and 48 ties. This framework allows us to hold structural relationships among nodes constant while varying their spatial relationships. The arrangements vary in the proximity of nodes to each other and the positioning of nodes toward the center or periphery of the graph.

The factors that influence perception of graphs can be divided into two groups: those that are concerned with spatial properties of the graph layout and those that are concerned with graph theoretical properties of the network itself. In this study we consider how Euclidean spatial factors affect the viewer's perception of the graph when structural features are held constant. In particular we consider the influence of:

1. proximity to the center of spatial arrangement on perception of prominence;
2. positioning between clusters of nodes in spatial arrangement on perception of bridging;

3. spatial clustering of groups of nodes in spatial arrangement on perception of grouping.

In Section 3 we explain in more detail how the various measures were calculated and show our results based on ordinary least squares (OLS) and analysis of errors in reporting.

2. Experimental design

2.1. The study

Eighty graduate students who had just completed a course in organizational theory (a course that emphasized the importance of understanding networks in organizations) volunteered to be subjects in the experiment. The subjects were given a questionnaire containing three of the five graph spatial arrangements and told the following:

> The following three graphs are modeled after networks of communications observed in three different merger and acquisition teams of an investment banking firm. A connection between two team members means that they discuss work-related matters with each other. If no line exists between two team members then they never discuss work with each other.

All nodes were labeled with first names in the arrangements presented to the subjects. By providing a context, the investment banking firm and member names, we attempt to focus the subjects' attention on the social aspect of communication networks. In every arrangement, each node was mapped to a new name. For analysis and discussion purposes, we have relabeled the nodes of interest with letters from A to E. For prominence and bridging, A and B have the highest scores, followed by C, and D and E have the lowest scores. The five spatial arrangements are shown in Fig. 1. It is important to remember that all five spatial arrangements display identical networks.

We asked respondents two questions about the same five focal nodes in each graph: how 'prominent' was each of five particular players in the graph; and how important a 'bridging' role did each of these five particular players occupy in the graph. We also asked them to report the number of subgroups in the graph. Subjects rated the prominence and bridging of the same five nodes (with different first name labels) for each spatial arrangement by circling a number from 1 to 7 on a Likert scale that went from not prominent (or not important as a bridge) to most prominent (or most important as a bridge). The question about prominence was worded as follows: "Some individuals have a more prominent role in their team than other individuals. Please rate the following people according to how prominent within their team they appear to you by circling the appropriate number next to each name." The question about importance as a bridge was worded in the following way: "Some individuals are important because they form a bridge between subgroups. Please rate the following people according to how important they appear to be as bridges between subgroups by circling the appropriate number next to each name." Finally, we created the questionnaire using a format that allowed one spatial arrangement and questions about that spatial arrangement to be viewed at the same time, without viewing any other spatial arrangement.

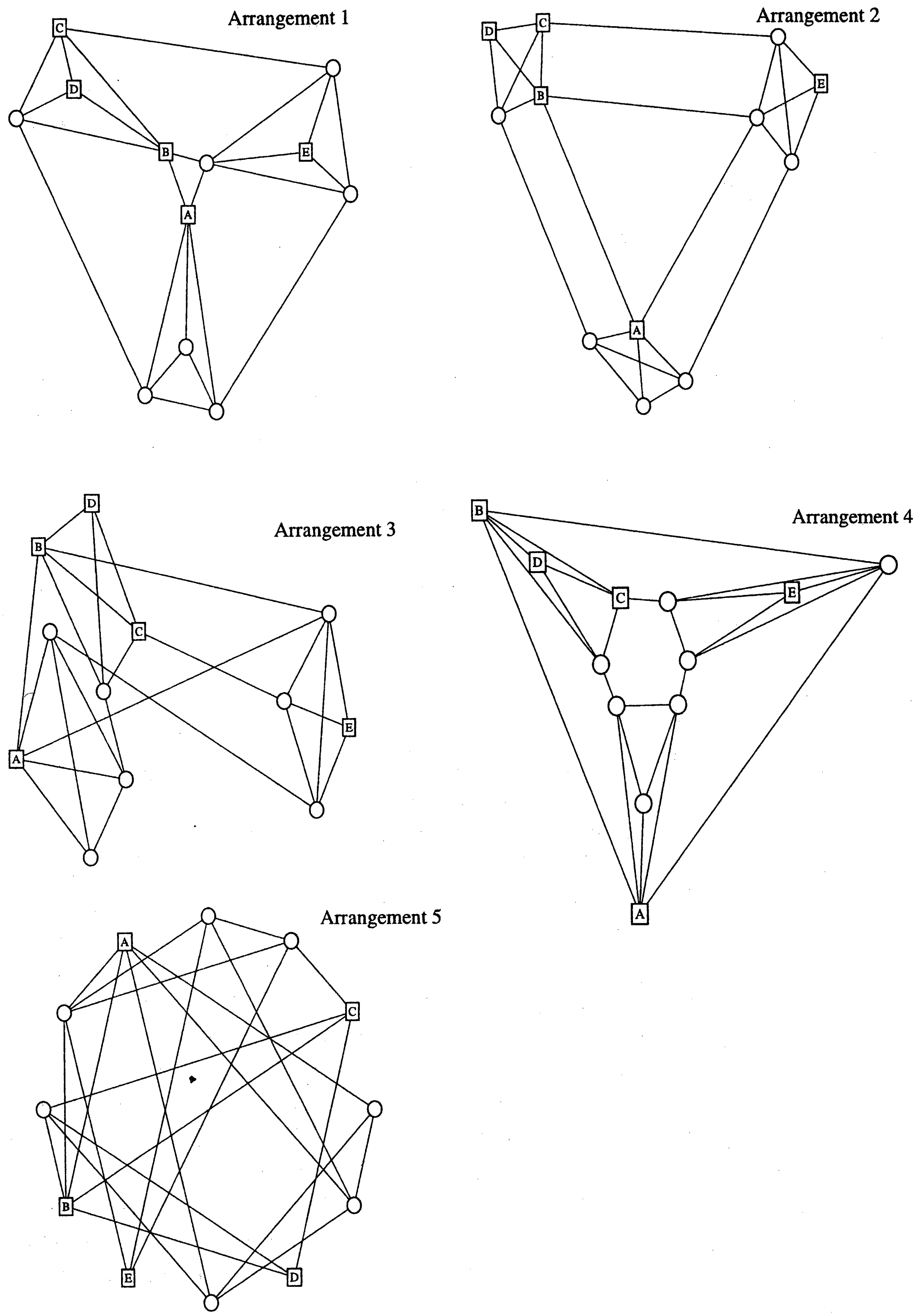

Fig. 1. The five different spatial arrangements of the same labeled graph.

We used an incomplete balanced Latin square design to control for order in the presentation of the graphs by displaying each spatial arrangement first, second, and third exactly once. There were five versions of the questionnaire, each version containing only three of the five drawings. Table 1 depicts which questionnaire version (Q1 to Q5)

Table 1
Order of appearance of spatial arrangements in five questionnaires

	A1	A2	A3	A4	A5	*N*
Q1	1		2		3	14
Q2		2	1	3		18
Q3			3	2	1	16
Q4	3	1			2	20
Q5	2	3		1		12

contained which spatial arrangements (A1 to A5) and the order in which those spatial arrangements occurred. For example, the second row (labeled 'Q2' in Table 1) indicates that Questionnaire 2 presented the respondent with Arrangement 3 (A3) first, followed by Arrangement 2, followed by Arrangement 4. By design, each spatial arrangement is preceded by every other except one exactly once. Order effects are discusses in Section 3.1.

2.2. The network

The network itself is symmetric, with twelve actors and 48 directed ties. We used a small graph so that the subjects would not be overwhelmed by the amount of information presented to them. The overall density is 36%. Table 2 shows values for the two measures of centrality for nodes A through E. Degree centrality measures the number of nodes that are adjacent to the focal node. Betweenness centrality measures the number of times a node is on the shortest path between two other nodes (Freeman, 1978). In our experiment, degree centrality and betweenness centrality are strongly correlated ($r = 0.93$). While degree centrality and betweenness centrality remain conceptually distinct, in this case they are not very different empirically. Nodes A and B are automorphically equivalent as are nodes D and E (see Borgatti and Everett, 1992, for a discussion of automorphic equivalence). The network has four cliques.

Fig. 1 shows the five different spatial arrangements in which the network was presented. Keeping the aesthetics standards of regular spacing of nodes and minimization of edge crossings in mind, we arranged the graph in ways that would be interpretable by the respondents. Arrangement 5, the circle, was not drawn with the

Table 2
Centrality measures

	Degree centrality	Freeman betweenness centrality
A	5	8.67
B	5	8.67
C	4	4.67
D	3	0.00
E	3	0.00

standards in mind. It is included because of its general acceptance as a means for presenting social network data. In his handbook on social network analysis Scott says:

> various *ad hoc* extensions to the idea of the sociogram have been used as researchers have sought to complement their mathematical measures with some kind of diagrammatic representations. One common technique has been to construct the sociogram around the circumference of a circle so that the pattern of lines becomes more visible... (Scott, 1991, p. 148).

3. Results

3.1. Order effects

Because each subject provided responses for three spatial arrangements, we compare both within and between subjects. We were concerned that the within subject comparisons may be biased if subjects were learning about interpreting graphs as they moved from the first to the third graph evaluation task. Poulton (1982) discusses the strategy transfer hypothesis describing cases in which subjects learn a strategy in one condition and transfer it to another condition. In exploring the effect of order, we found that there were no significant overall mean effects. To test for order effects, we ran two regressions, one with prominence as the dependent variable and one with bridging as the dependent variable. In each regression we controlled for individual respondents, the node, and the layout. When we included order of appearance, order had no significant effect on the independent variable for either prominence or bridging.

However, when each spatial arrangement was considered separately, Arrangement 3 showed some difference depending on order, as seen in Table 3. This suggests that Arrangement 3 may be particularly difficult to interpret and therefore subject to different interpretations depending on the layout after which it appears. None of the other spatial arrangements showed any significant difference in prominence scores when separated by order of appearance.

3.2. Predicting judgment

The results of our analysis of individuals' reports of each node's prominence and importance as a bridge and of the overall number of groups in the network suggest that spatial relationships between nodes influence viewers' perceptions of the graphs. The

Table 3
Analysis of variance for mean centered prominence scores for Arrangement 3 (the only arrangement showing significant order effects for prominence)

	First	Second	Third	*F* value
A	1.58	1.01	0.57	2.76
B	1.77	0.70	0.31	4.91 **
C	0.33	−0.22	0.51	1.98
D	−2.29	−0.53	−0.56	8.68 ***
E	−2.23	−0.61	−0.49	7.66 ***

** $p < 0.01$; *** $p < 0.001$.

Table 4
Individual mean centered scores for prominence (standard deviation in parentheses)

	Arrangement 1	Arrangement 2	Arrangement 3	Arrangement 4	Arrangement 5
A	0.69 (1.79)	1.46 (1.18)	1.11 (1.24)	1.08 (1.60)	0.36 (1.88)
B	0.64 (1.86)	1.30 (1.39)	1.00 (1.45)	1.06 (1.59)	0.29 (1.76)
C	−0.30 (1.30)	−0.13 (0.92)	0.24 (1.00)	0.26 (1.42)	−0.42 (1.60)
D	−0.12 (1.39)	−0.96 (1.29)	−1.15 (1.54)	−0.24 (1.47)	−0.93 (1.32)
E	−0.16 (1.67)	−1.58 (1.31)	−1.18 (1.55)	−0.40 (1.48)	−1.02 (1.64)

mean centered average prominence and importance as a bridge for each node in every arrangement shows that while the ordering of prominence and bridging does not change across spatial arrangements, the relative values assigned to prominence and bridging change. We use OLS to compare Arrangements 1 through 4 to Arrangement 5, the circle. When nodes are arranged in a circle, they are all equally placed from the center and equally placed between other nodes. There is no variation in the spatial information provided in the circle.

For grouping, we compare the distribution of groups across all five layouts and find some differences in the shapes of the distributions across the five layouts. The following sections give details of the analysis for prominence, importance as a bridge, and grouping.

3.2.1. Prominence

Before reporting the prominence data, we converted the prominence reported on a seven-point Likert scale to a mean centered prominence score. This controls for individuals' tendencies to rate high or low in general. We did this by subtracting each respondent's average prominence score across all three spatial arrangements from the prominence score assigned to the node by the respondent. The average mean centered prominence scores for each node in every spatial arrangement are reported in Table 4. Overall, the ordering of the value for prominence is stable. Ordering changes only once, in Arrangement 1, where node C's mean centered average prominence is less than node D's and node E's.

Next, we estimate the prominence for each node using arrangements as the indepen-

Table 5
Regression prominence by node

Variable	A	B	C	D	E
Inter	4.32 ***	4.48 ***	3.05 ***	2.92 ***	2.80 ***
A1	0.21	0.05	−0.12	0.70 *	0.93 **
A2	0.83 *	0.50	−0.03	−0.46	−0.33
A3	0.69 *	0.42	0.48	−0.16	−0.09
A4	0.50	0.21	0.26	0.72 *	0.72 *
R^2	0.423	0.459 *	0.340	0.587 ***	0.593 ***
N	232	232	229	229	232
R^{2*} (ID control only)	0.391	0.446 *	0.319	0.517 ***	0.520 ***
$R^2 - R^{2*}$	0.031	0.013	0.021	0.070 ***	0.072 ***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

dent variable. The results are shown in Table 5. We include individual identifiers to control for individuals' tendency to rate high or low. And we test for the joint significance of the arrangement variables by comparing R^2 for an equation containing the identifier controls and the arrangement variables (R^2) to the R^2 for an equation containing only the identifier variables (R^{2*}). Nodes D and E are the only nodes for which the arrangement variables are jointly significant. D and E may be most amenable to the influence of arrangement because they are the least structurally prominent.

Node C follows the expected pattern, that is, in Arrangement 1 and Arrangement 2, when C is on the periphery, its coefficients predict values for prominence that are less than the circle. In Arrangements 3 and 4, when C moves toward the center of the graph, its coefficients become positive. Similarly, nodes D and E have positive coefficients for Arrangements 1 and 4 when they are closer to the center of the graph, and negative coefficients for Arrangements 2 and 3 when they are more peripheral. Nodes A and B provide a contradiction to the simple spatial theory we propose. They are not significantly more prominent in Arrangement 1 where they are members of a central cluster. This result contradicts the proposition that positioning in the center of the graph enhances the perceived prominence of nodes. One explanation for this might be that nodes with high structural prominence are not affected by spatial positioning as much as nodes with low structural prominence.

3.2.2. Importance as a bridge

Table 6 shows the mean centered bridging scores for all five nodes and five arrangements. The order of reported bridging is consistent for all nodes in all arrangements: A and B are always greater than C which is always greater than D and E.

Again, we estimate each node's importance as a bridge using arrangements as the independent variable, including individual identifiers and testing for joint significance. The results are shown in Table 7. For bridging, the arrangement variables are jointly significant for nodes A, B and E.

Both A and B have higher bridging scores in Arrangements 1 and 2 when they are positioned between two clusters than they do in Arrangements 3 and 4 when they are not. Node C follows a similar pattern. In Arrangements 1 and 2, C is not the clique member closest to another clique and its reported bridging is lower than in Arrangements 3 and 4, when C is positioned between clusters. D and E are reported to be less important as a bridge in Arrangements 1 through 4 than they are in the circle. However, the coefficients for E are not statistically significant.

Positioning nodes between groups seems to enhance the perception of their impor-

Table 6
Individual mean centered scores for importance as bridge (standard deviation in parentheses)

	Arrangement 1	Arrangement 2	Arrangement 3	Arrangement 4	Arrangement 5
A	1.73 (1.45)	2.02 (0.91)	1.51 (1.29)	1.30 (1.82)	0.68 (1.66)
B	1.76 (1.41)	1.88 (1.25)	1.53 (1.16)	1.47 (1.64)	0.80 (1.60)
C	0.27 (1.24)	0.24 (1.02)	0.53 (1.25)	0.84 (1.37)	−0.47 (1.52)
D	−1.74 (1.33)	−2.06 (0.99)	−1.75 (1.16)	−1.35 (1.67)	−0.86 (1.44)
E	−1.63 (1.48)	−2.10 (0.96)	−1.70 (1.39)	−1.60 (1.42)	−1.09 (1.80)