Journal of Classification 11:121-149 (1994)

METRIC INFERENCE FOR SOCIAL NETWORKS

David Banks
Carnegie Mellon University

Kathleen Carley
Carnegie Mellon University

Abstract: Using a natural metric on the space of networks, we define a probability measure for network-valued random variables. This measure is indexed by two parameters, which are interpretable as a location parameter and a dispersion parameter. From this structure, one can develop maximum likelihood estimates, hypothesis tests and confidence regions, all in the context of independent and identically distributed networks. The value of this perspective is illustrated through application to portions of the friendship cognitive social structure data gathered by Krackhardt (1987).

Keywords: Random networks; Random graphs; Digraphs;

1. Introduction

Some researchers in social networks are beginning to move from the analysis of single networks to the analysis of random samples of networks. In looking at multiple networks, issues arise such as the extent to which these networks are similar, and the estimation of the central or consensus network. For example, imagine that one is interested in locating the shared or common perception of the informal communication network in an organization. This shared perception could be located by asking each individual in the

---

We thank Ove Frank, David Krackhardt, the editor and the referees for their constructive comments and suggestions.

Authors' Addresses: David Banks, Department of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213, USA and Kathleen Carley, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

organization to identify all pairs of colleagues whom the respondent believes interacts with whom, and then locating the common perceived ties. Following this procedure, the researcher would locate a set of networks, one generated by each individual in the organization. Such methods have been used by Krackhardt and Porter (1985) and Krackhardt (1987) to generate cognitive social networks. Given this set of networks, we want to locate a common central network and assess the variation among the respondents' perceptions. In this paper, we use a metric to induce a tractable family of probability measures on the set of networks. Researchers can then use standard statistical methods to estimate measures of location and variation.

Our procedures are designed for networks in which vertices (nodes) are labeled and the edges (links) may or may not be directed. The term graph is usually reserved to networks in which all edges are undistinguished (unweighted and undirected), but in some research communities the term graph is used even when edges are distinguished. In this paper we use the term network and graph interchangeably; our application targets a problem in social network theory, but our methods derive from graph theoretic considerations, leading to a collision of terminology.

Network-valued random variables arise in many settings. Social network research is one arena, and early formulations of such problems were undertaken by Moreno (1934), Festinger (1949), Katz (1947, 1953) and Katz and Powell (1955). Statistical investigation of random networks remains active; Knoke and Kuklinski (1982) provide an elementary introduction to some topics in this area, and Fienberg, Meyer, and Wasserman (1985) survey more technical literature in the statistical analysis of such data. Frank (1989) also reviews the area and points out the need for multiparametric models for random networks. Our work describes such a model, using a parameterization that resembles more standard statistical applications.

To frame this paper in the context of previous work, we develop a statistical model and the associated analysis for a sample of n random variables defined over the set of all networks on m labeled vertices. Our attention is upon estimation and hypothesis tests regarding the central network and the dispersion of the data; these quantities are analogous to the mean vector and covariance matrix of a sample of random vectors.

We will use a metric to induce a tractable family of probability measures on networks. Topics related to this technique have been considered in three distinct literatures: mathematics, sociology, and statistics. The mathematical work has focused attention on only three probability models (cf. Bollobás 1985, Ch. 2; Palmer 1985, Ch. 2). These models are useful in developing existence proofs and identifying certain kinds of asymptotic behavior, but they are insufficiently rich for statistical modeling — in particular, they lack parameters that determine the center and dispersion of the

measure. Regarding metrization in this literature, Margush (1982) derives optimality properties for a metric on the set of tree-valued graphs, but this result was not implemented in a probabilistic context, and did not lead to formal inference.

The sociological literature tends to be either very abstract or very applied. At the theoretical end, there is a large literature on metrics for posets, beginning with a seminal paper by Boorman and Olivier (1973). Since some posets correspond to certain classes of graphs or hypergraphs (such as n-trees and phylogenetic trees), this work bears on the metrization of graph-valued random variables. Barthélemy, Leclerc, and Monjardet (1986) offer a general review of these methods, and Day (1986) gives a succinct survey of recent literature. In particular, Jardine and Sibson (1971, Ch. 5) discuss metrics on sets of relations, and Hubert and Arabie (1985) consider methods for comparing partitions of sets; both examine the symmetric difference metric that is of central interest in this paper. Although the ostensible goal of all these efforts is to allow practitioners to use data to make inferences in problems of comparison and consensus among classifications, the work in this area has not been developed in terms of estimates, hypotheses, and test statistics.

The applied literature in sociology is driven by people who are attempting to understand particular datasets or types of datasets. For example, Carley (1984, 1986) examined changes in students' cognitive maps (networks of concepts) over time as they went about the process of selecting a new tutor for their living group (hereafter, this will be referred to as the Carley Tutor dataset). To compare these conceptual networks and to locate the network that was common to all, Carley employed the notion of "facts" (see also Carley 1988). A fact is defined as two vertices and the edge between them. Carley then defined the similarity of two networks according to their intersection — the network formed of all facts that the two separate networks have in common. The common (or central) network was estimated by the set of facts found in the intersection of a specified percentage[1] of the sample networks.

Statisticians have addressed random networks in very standard situations. Holland and Leinhardt (1981), Fienberg, Meyer, and Wasserman (1985), Frank and Strauss (1986), Wasserman (1987), Wong (1987), and Strauss and Ikeda (1989) have all developed probability models and analyses for social network data, usually in the context of the loglinear model. These

1. When this percentage is 50%, the method yields the maximum likelihood estimate for the central network under the probability model we propose.

studies have typically focused on only a single network, rather than a sample of networks, with emphasis on accurate modeling of individual interactions (however, the methodology could be extended to i.i.d. samples). Log-linear models are often, though they need not be, elaborate models with many parameters and can obscure inference on the two parameters (center and dispersion) targeted by the procedures we describe in Section 2. In contrast, Bloemena (1964), Capobianco (1970), and Frank (1971, 1988) discuss methods that explicitly address samples of random graphs. Those authors' perspective matches that of our paper, and it is surprising that more work in this area has not been done. Section 2 relates our approach to the most pertinent models.

Using the methods developed in Section 2, Section 3 reexamines portions of the data collected by Krackhardt (1987) and available through UCINET (MacEvoy and Freeman 1988). These methods discover several interesting features of the data, and we illustrate a range of hypothesis tests and confidence regions. Section 4 draws a few conclusions about the theory and the application.

2. Methodology

We want to make statistical inferences about network-valued random variables. Our strategy is to define a natural metric on the set of networks and use it to induce an interpretable family of probability measures. This strategy enables us to appropriate an arsenal of statistical techniques whose application is almost automatic.

Let G_m denote the set of all networks on m distinct vertices which have undistinguished (unweighted) edges and no loops. (Loops are edges that connect a vertex to itself; our applications exclude that situation, but the method we describe could be extended very generally.) Such networks are commonly called graphs. Elements of G_m include the edgeless graph, the complete graph, and all intermediate possibilities.

Let $\mathbb{R}$ denote the real numbers. Recall that a function $d: G_m \times G_m \to \mathbb{R}$ metrizes G_m if and only if for all networks $g_1, g_2, g_3 \in G_m$,

1. $d(g_1,g_2) = 0$ iff $g_1 = g_2$
2. $d(g_1,g_2) = d(g_2,g_1)$
3. $d(g_1,g_2) \leq d(g_1,g_3) + d(g_3,g_2)$.

The function d is called a metric, and (G_m,d) is called a metric space.

One can define many possible metrics on the set of networks with m vertices. In a particular application, the metric should reflect a sense of distance that honors the context of the data. However, the symmetric difference metric on sets, also known as the Kemeny metric (1959), is broadly

applicable and enables more tractable analyses. In this paper, that is the only metric used.

Heuristically, the symmetric difference metric simply counts the number of discrepant edges between two networks on the same vertex set. A computationally convenient specialization of this metric to our application uses the fact that any network $g \in G_m$ is uniquely characterized by its adjacency matrix $G = [I_{ij}(g)]$, where

$$I_{i,j}(g) = \begin{cases} 1 & \text{if an edge links vertices } i \text{ and } j \text{ in } g \\ 0 & \text{else.} \end{cases} \tag{1}$$

All $m \times m$ matrices whose entries are elements of $\{0,1\}$ can be considered adjacency matrices. Excluding matrices that have ones on the diagonal, there is a one-to-one correspondence between symmetric adjacency matrices and the networks in G_m. (Nonzero diagonal entries in G indicate loops. Digraphs have a similar representation, except that superdiagonal and subdiagonal entries indicate direction. Both cases are discussed at the end of this section.)

Let G_1, G_2 be the adjacency matrices of $g_1, g_2 \in G_m$. Define the symmetric difference metric on networks by

$$d(g_1,g_2) = \frac{1}{2} \operatorname{tr}[(G_1 - G_2)^2] \tag{2}$$

where $\operatorname{tr}[\cdot]$ denotes the trace of a matrix; i.e., the sum of the diagonal elements. This function counts the number of edge discrepancies between g_1 and g_2.

The metric d is the Hamming metric (1950, 1980), used in information theory. The networks in G_m can be viewed as the vertices of an r-dimensional hypercube, where $r = \left\lceil \frac{m}{2} \right\rceil$. From this perspective, the distance between networks is just the Hamming distance between the sequences of zeroes and ones that identify corresponding hypercube vertices. A particular vertex with a given binary sequence indicates the network with edges determined by the ones and non-edges by the zeroes in the sequence. It follows that G_m contains 2^r networks.

Given the metric, we can mimic Mallows's method (1957) for setting probabilities on the set of permutations. For networks, this approach yields the probability measure $H(g^*, \sigma)$, defined by

$$P_{(g^*,\sigma)}[g] = c(\sigma) e^{-\sigma d(g,g^*)} \quad \forall\, g \in G_m . \tag{3}$$

where $g^* \in G_m$ is the central network (or mode of the distribution), σ is a

dispersion parameter and $c(\sigma)$ is a normalizing constant that ensures the probability of the sample space is unity. Heuristically, g^* is a central value parameter (analogous to the mean of a distribution), and σ is a scale parameter (analogous to the inverse of the standard deviation). These parameters index the family of probability measures $\{H(g^*, \sigma)\}$. When $\sigma = 0$, all networks are equiprobable.

There are many variations of the model in (3) that might be considered; e.g., one could square the distance, or use a mixture of such models, or allow the probability to change as a function of the distance other than the exponential. Each of these may be appropriate in a given context, but closed form expressions for key estimates are not generally available. We will develop the analysis for the most tractable model and note that the only barrier to a wider range of plausible models is the availability of computational resources.

Proposition 1: *The normalizing constant does not depend on* g^*.

Proof: Necessarily, the normalizing constant satisfies

$$\frac{1}{c(\sigma)} = \sum_{g \in G_m} e^{-\sigma d(g, g^*)} \tag{4}$$

$$= \sum_{k=0}^{r} \begin{bmatrix} r \\ k \end{bmatrix} e^{-\sigma k}$$

$$= (1 + e^{-\sigma})^r .$$

The penultimate step follows by counting the number of networks in G_m that are exactly k edge-changes distant from an arbitrary central network g^*; the last step is obtained from the binomial theorem. This result holds even when $\sigma = 0$. ■

Other measures for random graphs (networks) have been previously proposed. Mathematicians use three basic families (cf. Bollobás 1985, Ch. 2), but these models are insufficiently rich for some statistical applications. From the perspective of this paper, their chief deficiency is the lack of a unique central graph and a dispersion parameter to control the degree of probability concentration about the central graph. The family of measures $\{H(g^*, \sigma)\}$ automatically avoids these limitations, as would any similar family developed from a different metrization than the one defined by (2). Although this metric seems most appropriate for the applications described in this paper, we note that different metrics, corresponding to alternative topologies or senses of nearness, enable very flexible generalizations of the analysis undertaken here.

More directly pertinent are the probability models discussed by Frank and Strauss (1986) and by Holland and Leinhardt (1981). The former is developed in Strauss and Ikeda (1990), and the latter is extended in Wasserman (1987), Wasserman and Anderson (1987), Wasserman and Galaskiewicz (1984), Iacobucci and Wasserman (1986), and Fienberg, Meyer and Wasserman (1985). Wang and Wong (1987) and Wong (1987) have also done relevant work on Holland and Leinhardt's model. In principle, these models could be used to analyze multiple networks, although there has not yet been any effort in that direction.

Frank and Strauss (1986) used the Hammersley-Clifford theorem to show that all probability models for undirected random graphs can be written in the form

$$P_D[g] = c \exp\left[\sum_{A \subseteq g} \alpha_A\right] \quad \forall\, g \in G_m \tag{5}$$

where c is a normalizing constant and α_A is a nonzero constant iff A is a clique of the nonrandom dependence graph D. In this context, the nodes of D are all possible edges on m vertices; a clique in D is a subset of the vertex set of D that is either a singleton set or has the property that all pairs of elements are connected by edges in D. The D determines the dependence structure between random edges; if D connects a specific pair of edges, then those edges in g are conditionally dependent given the other edges in g. Frank and Strauss proceed to define a class of random models called Markov graphs; these correspond to a special sparse structure in the dependence graph D (i.e., the adjacency matrix of D has ones in a few particular locations and zeroes everywhere else).

In this setting, the Holland and Leinhardt p_1 model is a submodel of the version of (5) that describes directed graphs. The p_1 model and subsequent developments of it correspond to assuming the dependency graph D to be edgeless (i.e., the adjacency matrix consists of zeroes), while imposing a special structure on the α_A terms. This result follows from the stochastic independence of different dyads. Work in this area has focused on building log-linear models that describe the probability that a particular edge occurs, as a function of characteristics of the vertex and its neighbors.

From this perspective, the model described in this paper also entails an edgeless dependence graph. Equating the representation in (5) with the form given in (3), one finds that the probability that an edge or a non-edge in a random network g agrees with the corresponding edge or non-edge in g^* is $(1 + e^{-\sigma})^{-1}$, and this is independent of the presence or absence of all other edges. Thus the model we discuss is especially appropriate whenever the data can be regarded as deriving from a true network g^* that is measured with independent, edgewise error.

Proposition 2: *Under the model in (3), the dependence graph D is edgeless.*

Proof: Without loss of generality, assume that g^* is the edgeless graph; also, let g be a graph with $k + 2$ edges, two of which are e_1 and e_2. We show that the joint probability of edges e_1 and e_2, conditional on the other edges in g, is not the product of the marginal probabilities of e_1 and e_2, each conditional on the other edges in g. Clearly,

$$P[e_1 \text{ and } e_2 \mid \text{the remaining } k \text{ edges}] = \frac{e^{-(k+2)\sigma}}{e^{-k\sigma} + 2e^{-(k+1)\sigma} + e^{-(k+2)\sigma}}$$
$$= \frac{e^{-2\sigma}}{(1 + e^{-\sigma})^2},$$

and

$$P[e_1 \mid \text{the remaining } k \text{ edges}]\, P[e_2 \mid \text{the remaining } k \text{ edges}]$$
$$= \left[\frac{e^{-(k+1)\sigma} + e^{-(k+2)\sigma}}{e^{-k\sigma} + 2e^{-(k+1)\sigma} + e^{-(k+2)\sigma}}\right]^2$$
$$= \frac{e^{-2\sigma}}{(1 + e^{-\sigma})^2}.$$

These quantities are equal; all edges are thus conditionally independent. By definition, D is edgeless. ■

Proceeding now to data analysis, suppose one observes $g_1, \ldots, g_n$, a random sample of networks with distribution $H(g^*, \sigma)$, where g^* and σ are unknown. In this framework, we find that we can develop many of the usual tools of statistical inquiry.

Estimation

Following the conventional strategy for maximum likelihood estimation, we find the log-likelihood function of the sample as

$$L(g^*, \sigma) = n \log c(\sigma) - \sigma \sum_{i=1}^{n} d(g_i, g^*). \tag{6}$$

We seek the estimates $(\hat{g}^*, \hat{\sigma})$ that maximize this quantity.

For any value of σ, it is clear that (6) is maximized in g^* by[2]

$$\hat{g}^* = \text{argmin}_{g^* \in G_n} \sum_{i=1}^{n} d(g_i, g^*). \tag{7}$$

The quantity $\sum d(g_i, g^*)$ is called the remoteness function, and the solutions of (7) are called medians (cf. Barthélemy and Monjardet 1981, 1988). Similar estimators were derived in the context of trees by Margush (1982), Margush and McMorris (1981), Barthélemy and McMorris (1986), and McMorris (1990). These works did not focus upon statistical inference, and only McMorris uses an explicit probability model; that model describes consensus in voters' rankings, and extends work by Young (1988) and Condorcet (1785).

For the metric considered in this paper, the median can be found by majority rule; i.e., an edge is in $\hat{g}^*$ if and only if it is present in more than 50% of the observed networks (the usual non-uniqueness problem may arise when n is even).[3] This result is standard in minimizing the remoteness function under this metric, with early work going back to Condorcet (1785). For other metrics, the solution is more complicated and can become computationally intensive. Barthélemy and Monjardet (1988) give a careful survey of the issues in this area, together with new results for specific applications. In many cases, steepest descent search of the network space yields a practical solution for alternative metrics.

To find $\hat{\sigma}$, differentiate (6) with respect to σ, set this to zero, and solve for σ. We obtain the solution

$$\hat{\sigma} = -\ln \frac{(rn)^{-1} \Sigma_{i=1}^{n} d(g_i, \hat{g}^*)}{1 - (rn)^{-1} \Sigma_{i=1}^{n} d(g_i, \hat{g}^*)}. \tag{8}$$

Thus the estimate of σ can be found analytically, once a majority rule calculation has developed the estimate $\hat{g}^*$.

Goodness-of-Fit

The easiest method for assessing the validity of the family $\{H(g^*, \sigma)\}$ in a given application depends upon the chi-squared test pioneered by Pearson (1900). Unfortunately, the parameter g^* is discrete, and thus conventional asymptotic theory, in which parameters take values over a vector

2. The argmin function returns the value of the index set which minimizes the argument.

3. Krackhardt (1987) refers to the central network, when it is derived from socio-cognitive structures, as the consensus structure. Carley (1984) refers to it as the majority intersection.

space, does not apply. This deficiency prevents the use of alternative methods for assessing goodness-of-fit, such as a likelihood ratio test, but does not preclude a conservative version of the chi-squared test. The procedure for employing the chi-squared test consists of the following five steps:

1. Use the data to find maximum likelihood estimates $\hat{g}^*$ and $\hat{\sigma}$.
2. Partition G_m into a relatively small number of regions $R_1, \ldots, R_p$, such that the expected numbers of observations in each region under the fitted null model are approximately equal. Denote the expected value of region R_i by E_i.
3. Count O_i, the number of sample networks observed in region R_i.
4. Calculate the test statistic

$$X^2 = \sum_{i=1}^{p} \frac{(O_i - E_i)^2}{E_i} . \tag{9}$$

5. Compare the test statistic to the value in a chi-squared table having $p - 2$ degrees of freedom and some appropriate α level, such as .05. If the test statistic is larger, then one has evidence that the model is incorrect.

We employ this procedure in Section 3.

We emphasize that the underlying assumptions for this test are not fully met. In particular, it is unclear how the degrees of freedom described in Step 5 above should be adjusted to account for the estimation of the central network by $\hat{g}^*$. If that parameter lay in a linear space, then we would subtract a degree of freedom for each dimension of that space; however, a network-valued parameter does not lie in a linear space, and thus classical theory does not apply (cf. Cramér 1961, Ch. 30.3). We suggest not subtracting any degrees of freedom for estimating the central network; then the degrees of freedom for the null distribution will be large, and the test will be conservative.

In general, assessing goodness-of-fit for social network data is very difficult. The discreteness of the parameter space prevented Frank and Strauss (1986) from examining the fit of their Markov graph model, and similar problems can arise in evaluating the p_1 model. Holland and Leinhardt (1981) applied an ad hoc test based on triad counts but found that their example dataset exhibited very bad fit. Subsequent authors have largely employed tests of simple versions of the p_1 model against more complicated p_1 versions, but this practice does not bear on the appropriateness of the basic model. Fienberg, Meyer, and Wasserman (1985) identify this problem as a key area for future research.

Confidence Regions

Confidence regions on the parameters are fundamentally important, but there is no immediately available theoretical prescription for determining these. Instead, we propose the use of either the parametric bootstrap or the nonparametric bootstrap; these are easily implemented computer-intensive procedures which automatically extend to the treatment of more complicated functions of the parameters. The parametric bootstrap makes heavy use of the model for the data. The nonparametric bootstrap is more generally applicable but tends to give confidence regions with larger volumes.

The parametric bootstrap, described in Efron (1982), uses the original sample to assess the standard errors in one's estimates. Operationally, one proceeds as follows:

1. Use the sample of random networks $g_1, \ldots, g_n$ to find the maximum likelihood estimates $\hat{g}$ and $\hat{\sigma}$. Draw a random sample of size n from the distribution with these parameter values. Denote the new random resample by $h_1, \ldots, h_n$.
2. Using the resample $h_1, \ldots, h_n$, apply the maximum likelihood procedure to estimate the center and dispersion by $(\hat{g}^*, \hat{\sigma}^*)$. This estimate takes values in $G_m \times \mathbb{R}^+$, where we take $\mathbb{R}^+$ to denote the non-negative numbers.
3. Repeat the first two steps B times, for B about 1000. (Efron and Tibshirani 1986 discuss heuristics for determining B.) Record the estimates for each repetition; denote the estimate corresponding to the i-th resample by $(\hat{g}_i^*, \hat{\sigma}_i^*)$.
4. Find the smallest volume in $G_m \times \mathbb{R}^+$ that contains $100(1-\alpha)\%$ of the B resampled estimates to get the bootstrap approximation to the desired confidence region.

The more common nonparametric percentile bootstrap proceeds exactly as above, except that in the first step, $h_1, \ldots, h_n$ is drawn at random with replacement from $g_1, \ldots, g_n$. A more thorough discussion of bootstrap methods, properties, and performance appears in Banks (1989).

The previous steps raise some theoretical points. The first concerns the determination of the minimum volume region $V = V_g \times V_\sigma$, defined by the Cartesian product of a set $V_g \subset G_m$ and $V_\sigma \subset \mathbb{R}^+$. The generalized volume of any such product is ab, where a is the cardinality of V_g and $b = \int_{V_\sigma} dx$. The volume must be minimized over regions V that attain the desired confidence; i.e.,

$$100(1-\alpha)B \le \sum_{i=1}^{B} \sum_{g \in G_m} I_V[(\hat{\hat{g}}_i^*, \hat{\sigma}_i)]\, d\sigma \tag{10}$$

where $I_V[\cdot]$ is an indicator function equal to zero unless the argument lies in V, in which case it takes the value one.

This joint minimization can be extremely difficult. Often it is easier to find a region that satisfies (10), and thus is an approximate $100(1-\alpha)\%$ confidence region, but which does not quite attain the minimum possible volume. A natural strategy is to require that the region have the form $\{g: d(g,\hat{g}^*) \le \delta\}$; if this volume is still too large, one can ask that the region contain $\hat{g}^*$ and all networks between $\hat{g}^*$ and a set of specified networks, where g is defined to lie between $\hat{g}^*$ and g^{**} if $d(\hat{g}^*, g^{**}) = d(\hat{g}^*, g) + d(g, g^{**})$. (This equality is equivalent to $\hat{g}^* \cap g^{**} \subset g \subset \hat{g}^* \cup g^{**}$, where the network is identified with its edge set.) More elaborate methods are also available that sometimes employ more complicated bootstrap methodologies.

The second theoretical issue is that the primary interest may not be (g^*, σ), but rather some functional $\theta(g^*, \sigma)$. For example, for a confidence region on g^*, we use the functional $\theta(g^*, \sigma) = g^*$. More generally, we might require the functional to pick out a subnetwork of g^* corresponding to certain vertices of particular interest, or evaluate the average vertex degree, or capture some other feature of the central network and the dispersion. Such generalizations are immediate, since the maximum likelihood estimate of the functional is just the functional of the maximum likelihood estimates (cf. Lehmann 1983, p. 112). Therefore we can proceed as indicated through Step 4 and then find the image of the points (g, σ) in the confidence region under the mapping induced by θ. Again, with greater effort we may obtain regions of slightly smaller volume.

Hypothesis Testing

If the union of the regions specified in the null and alternative hypothesis is the entire space, then the duality between confidence regions and hypothesis tests enables direct use of the previous bootstrap method. Specifically, suppose we want to test

$$H_o: (g^*, \sigma) \in \Theta_0 \text{ versus } H_A: (g^*, \sigma) \notin \Theta_0 \tag{11}$$

where $\Theta_0 \subset G_m \times \mathbb{R}^+$. Then we pick the desired form of the confidence region (either the difficult minimum volume region or one of the more tractable approximations), gather the sample, and set the bootstrap confidence region of size α. If the intersection of the bootstrap confidence region with

Θ_0 is empty, then we can reject the null hypothesis at level α. The same strategy generalizes immediately to functionals of (g^*,σ).

Powerful alternatives to the approximate bootstrap test are available in two serendipitously common applications. We note that these tests have meaningful interpretations even when the the postulated model in (3) fails to hold. In the first case, we want to test

$$H_o: \sigma = 0 \text{ versus } H_A: g^* = g_0 \text{ and } \sigma > 0. \tag{12}$$

The null hypothesis corresponds to uniform measure over the set of graphs (and thus no graph is the unique central graph); the alternative specifies non-uniform probabilities and a specific central graph. Under the null hypothesis, all networks are equally likely, and so the probability that a given observation is k units distant from g_0 has binomial distribution Bin(r, .5). For an independent sample $g_1, \ldots, g_n$, the reproductive property of the binomial distribution ensures that (under the null hypothesis) the sum of the distances has binomial distribution Bin(nr, .5).

Let the test statistic be $s=\Sigma_{i=1}^{n} d(g_i,g_0)$. It is straightforward to show that this selection gives a likelihood ratio test under our model, but the procedure is applicable far more generally. Since s directly measures the degree of concentration of the sample about the network g_0 specified in the alternative, then a test based on s should work well (though not optimally) for any model for the alternative in which g_0 is the true mode and the probability mass function is monotone with respect to distance from g_0. The significance probability p of such a test is

$$p = \sum_{j=0}^{s} \binom{nr}{j} 2^{-nr}. \tag{13}$$

This test depends only upon the null distribution and thus provides meaningful information even when our model is not posited in the analysis. Small values of p indicate that the sample is highly improbable when the null hypothesis is true.

The second common application is to test

$$H_o: \sigma = 0 \text{ versus } H_A: \sigma > 0, \tag{14}$$

which is equivalent to testing whether all networks are equally likely. Although this can be handled by the bootstrap procedure, a probably superior approximate test, avoiding pitfalls pointed out by Fisher and Hall (1990), is based upon a U-statistic (cf. Randles and Wolfe 1979, Ch. 3). A U-statistic is the average of dependent quantities; for a sample $g_1, \ldots, g_n$, define

$$U(g_1,\ldots,g_n) = \binom{n}{2}^{-1} \sum_{i<j} d(g_i,g_j). \tag{15}$$

Although one can show that the terms in the sum are pairwise independent, the triangle inequality implies that the entire collection is not mutually independent. Under the null hypothesis, standard U-statistic derivations show that this statistic is asymptotically normal with mean $r/2$ and variance $r/2n(n-1)$. Under the alternative, the asymptotic distribution is shifted to the left, thus enabling the usual machinery of hypothesis testing; we reject the null hypothesis if and only if

$$\frac{U(g_1,\ldots,g_n) - \frac{r}{2}}{\sqrt{r/2n(n-1)}} < z_\alpha \tag{16}$$

where z_α is the appropriate critical value from a standard normal table for a level α test.

Extensions to Directed Networks and Networks with Loops

The methods and results described so far extend immediately to the case of directed networks (commonly, digraphs) with loops permitted or not. Specifically:

1. If the data have directed edges but loops are excluded, then replace r by $m(m-1)$ throughout the preceding discussion. Let G_m^+ denote the set of loopless directed networks on m vertices, and use the metric
$$d^+(g_1,g_2) = \text{tr}\,[(G_1 - G_2)^T (G_1 - G_2)]. \tag{17}$$
2. If the data have directed edges and loops are allowed, then replace r by m^2 and use the metric d^+.
3. If the data have undirected edges and loops are allowed, then replace r by $(m^2 + m)/2$ and use the metric defined in
$$d^{++}(g_1,g_2) = \tfrac{1}{2}\,\text{tr}[(G_1 - G_2)^2] + \text{tr}[(\text{Diag}(G_1 - G_2))^2] \tag{18}$$
where Diag[·] denotes the diagonal matrix whose diagonal values agree with those of the matrix argument.

It is easy to verify that these variants of the original metric are metrics upon the corresponding set of networks; the key is to note that each variant simply counts the number of discrepancies between any two elements of the set. All the theoretical properties developed previously apply to these cases, and the proofs are entirely straightforward modifications of the ones already stated.

Section 3 examines one dataset that consists of directed networks with loops prohibited, to illustrate implementational details in applying and generalizing the methodology we have discussed.

3. Example — Krackhardt's Friendship Cognitive-Social Structure Data

The example we consider is drawn from the data set described in Krackhardt (1987), who collected what he refers to as cognitive social structure data for 21 managers in Silicon Systems, a high-tech computer consulting firm. The cognitive social structure data was obtained as part of a larger study on the effect of the perception of network structure on individual behavior. Several different network measures were collected including both friendship and advice based cognitive social structures.

To illustrate our methodology we will focus on only the friendship cognitive social structure. These data are of the form F_{ijk} where i is the sender of the friendship tie, j the receiver, and k the perceiver. These data consist of 21 networks, one for each of the 21 managers, such that each network is that manager's perception of who interacts with whom among the 21 managers. These networks include the manager's perceptions about his or her own friendship relations and the friendship relations among the other managers. The random network thus consists of nodes that are managers, and directed edges indicating perceived friendship relations. The 21 digraphs are available as krackfr.dat in UCINET (MacEvoy and Freeman 1988), a respository of social science data.

Step 1: Locate the central network.

Using all 21 digraphs, the majority rule algorithm finds that the maximum likelihood estimate of the central network for friendship has the adjacency matrix shown in Figure 1. This central network represents the "social cognition"; the perceptions about friendships among group members that are shared by members of the group. For these data, it happens that the estimate is unique. The central friendship network is sparse, many individuals (43 percent) appear to have no friends within the organization (1,6,7,10,11,13,15,16,10), and of the eleven friendship ties only six represent symmetric relationships (2-21,4-12,5-19). As to this latter point, Carley (1991) has argued that asymmetric ties are a natural result of interactions being based on relative similarity; they are part of the underlying structure and should not be viewed as errors in the data collection process.

The sum of the distances from each sample network to the estimated central network, $\hat{g}^*$, is 713, and the estimate from (8) is $\hat{\sigma} = 2.431$. Notice that we have used $r = m(m-1)$, or 420, since the networks are directed graphs without loops.

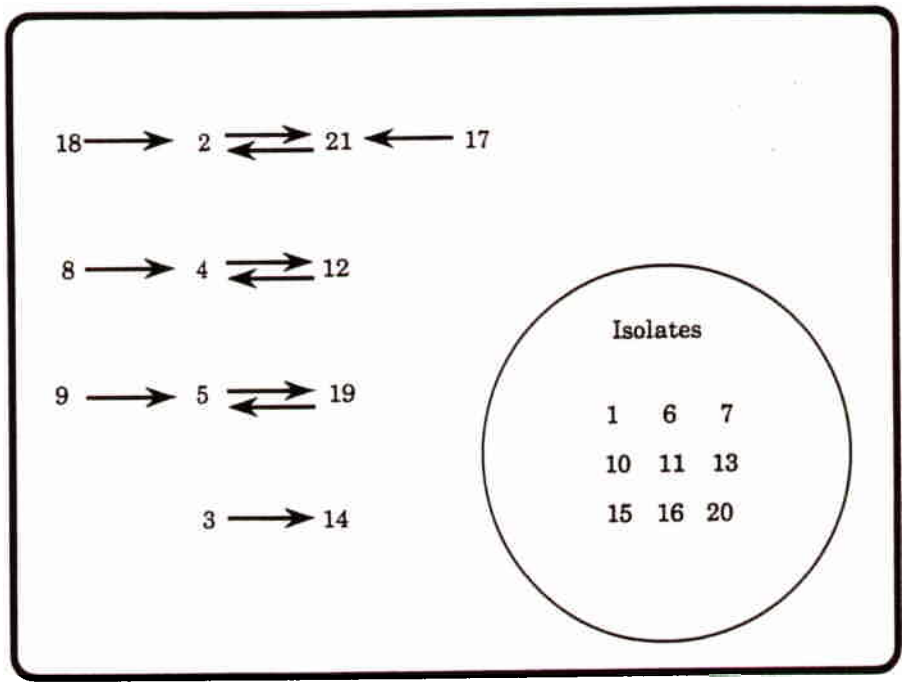

Figure 1. Central Network for Friendship Socio-Cognitive Structure using all 21 Matrices.

Step 2: Perform a goodness-of-fit test.

There are many ways we could partition the space of networks, but a natural division is based upon distance from the estimated central network. We form four subsets (or cells), consisting of networks that are 0 to 29 units away from $\hat{g}^*$, those that are 30 to 33 units away, those that are 34 to 37 units away, and all those that are further. We chose to make four subsets since: (a) fewer than three would preclude any degrees of freedom for a conservative chi-squared goodness-of-fit test, (b) the larger the number of cells, the greater the ability of a residual analysis of discrepancies between observed and expected cell counts to detect patterns of ill-fit; e.g., increasingly poor fit as one moves away from the central graph; and (c) more than four reduces the expected numbers of networks in each subset below the levels needed to obtain an approximate chi-squared distribution when the model is correct. Conditional on the number of subsets, we chose the cutpoints on the distance scale to obtain expected proportions that were as nearly equal as possible; *ceteris paribus*, this choice tends to maximize the power while improving the accuracy of the chi-squared approximation.

Under the null model with the fitted parameter values, the expected proportions in the four cells are .215, .262, .264, .258, respectively. Thus the expected numbers of sample networks in each cell are 4.52, 5.51, 5.55, and 5.41. The corresponding observed numbers are 10, 0, 3, and 8, so standard calculation shows that the chi-squared goodness-of-fit test statistic is 14.57, with a significance probability less than .001 (regardless of whether the conservative test with $p - 2 = 2$ degrees of freedom, or the less conservative test with $p - 3 = 1$ degrees of freedom, is used). A residual analysis of cell discrepancies shows that a model with both a greater peak and a heavier-tail is required. In other words, most of the sample "clusters" tightly around the central network, but a significant fraction of the observations are very distant. We will return to this point later.

Step 3: Locate the confidence region.

The next step in the analysis is to place a confidence region upon the central network and/or the dispersion parameter. Since the model does not fit, we use the nonparametric strategy and illustrate the case in which the parameter of interest is $\theta(g^*, \sigma) = g^*$, as mentioned at the end of the discussion of confidence regions in Section 2. The implementation is based on $B = 1000$ bootstrap resamples and exactly conforms to the four-step bootstrap algorithm described previously (sampling, of course, from $g_1, \ldots, g_n$ to obtain the nonparametric bootstrap rather than the parametric bootstrap). Since we are dealing with networks, the confidence region is a set of networks within a given distance from the estimated central network. The distances for the 90%, 95% and 99% confidence regions are 13, 15, and 19, respectively. In Figure 2, the central network and the networks for those six individuals (2,3,8,9,18, and 20) whose networks lie within the 99% confidence region are displayed. The identification number of the manager whose network is being displayed is printed within an oval. To aid the reader visually, those portions of the managers' networks (nodes and edges) that are part of the central network are emboldened in Figure 2. The networks of these individuals are of particular interest socially, as their networks could arguably equal the unknown central network. The perception of the friendship relations held by these individuals is the closest to the social cognition. In this sense, they can be thought of as the keepers of the cultural truth.

An Aside: Contrasting Social and Individual Perception

It is interesting to compare this shared or social perception with the individuals' perception. To do so, we can compare the central network with the individually perceived network which contains an edge (i.e., friendship relation) from manager i to manager j just in the case that i perceives that i

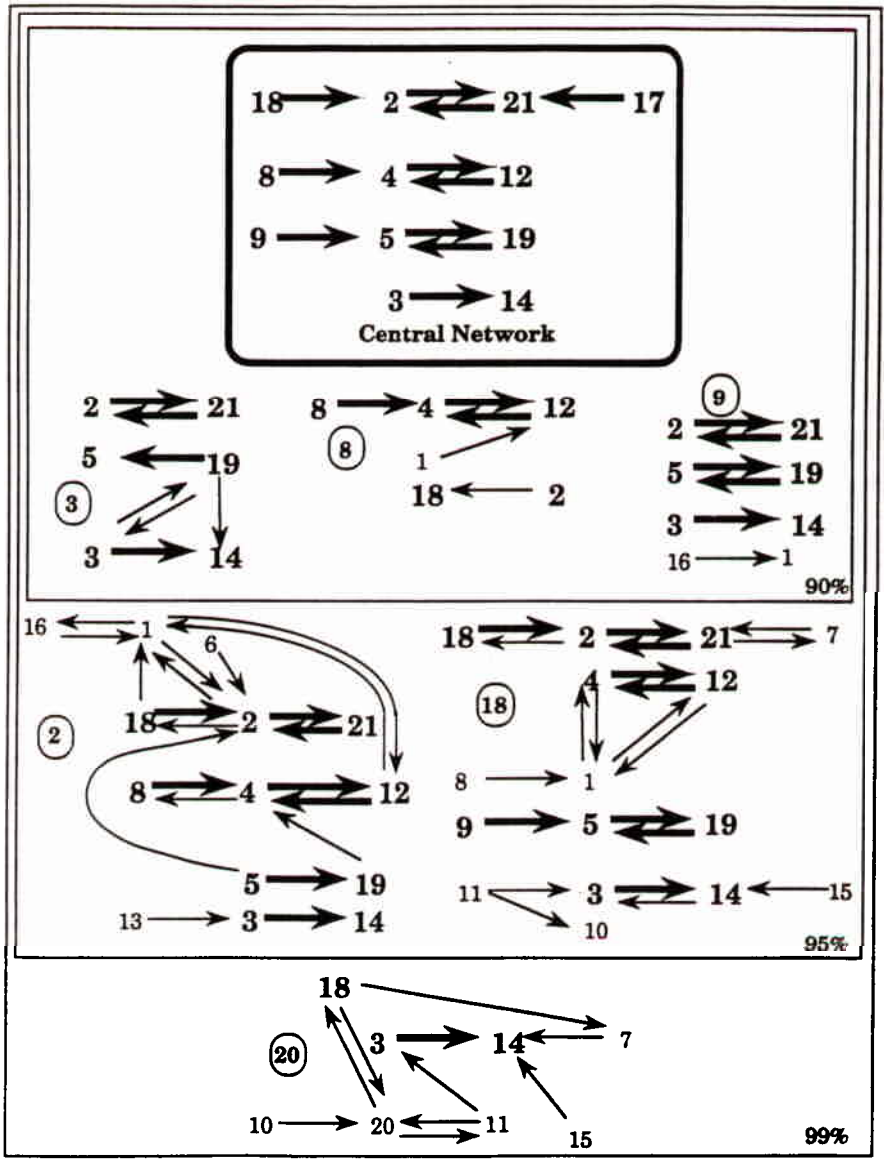

Figure 2. Graph Confidence Region.

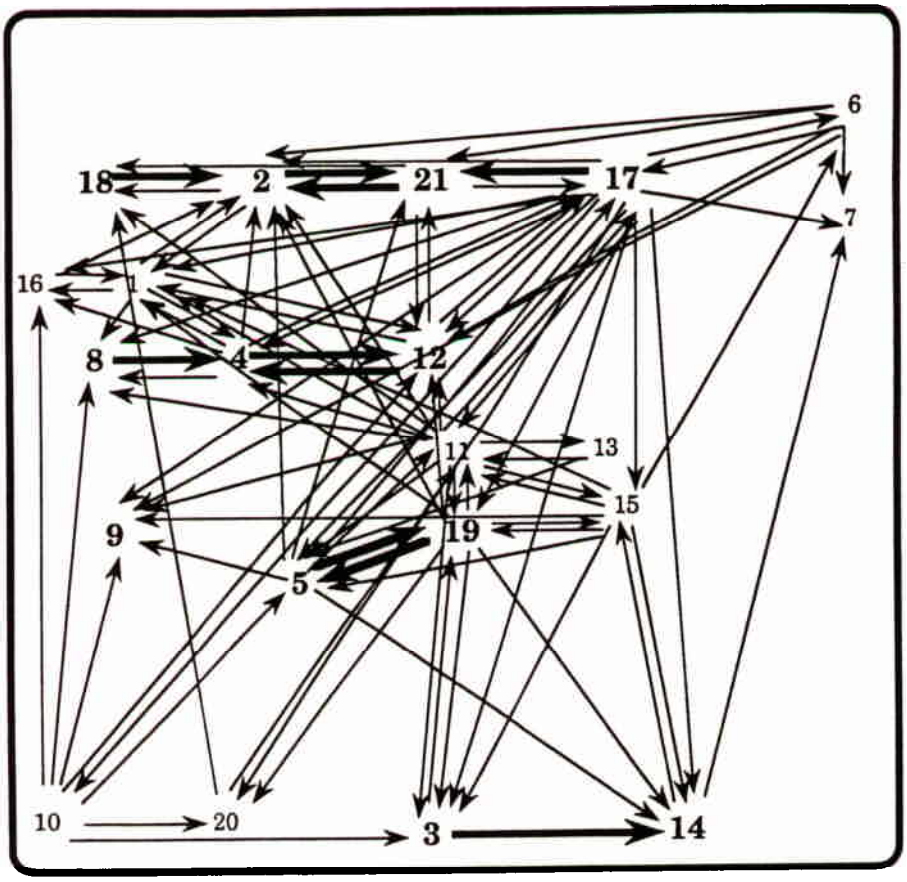

Figure 3. Individually Perceived Friendship Network.

sends a friendship tie to *j*. The resultant individually perceived friendship network is shown in Figure 3.

Contrasting Figures 1 and 3 provides insight into the difference between the friendship network as generally perceived by members of the group and the friendship network as defined by the individual. We see that the shared perception of who is friends with whom does not necessarily equal the set of dyads that consider themselves friends. The distance between the central network and the individually perceived network is 93. This distance is largely attributable to the fact that the individually perceived network is more dense (contains more edges) and contains all but one of the edges in the central network. This missing edge is that from Manager 9 to Manager 5. It is only in this one case that the society at large sees a friendship between two individuals where neither of the individuals involved considers the other as a

friend. For the most part, individuals claim friendships that are not perceived by the group at large. The fact that the individually perceived friendship network is more dense than the central network is to be expected. The larger the number of individuals who must agree for an edge to be included in a graph, the smaller the number of edges that occur. In this case, for the network in Figure 3, only one person must agree for an edge to occur, but for the network in Figure 1, at least 11 individuals must agree for an edge to occur.

That the two networks differ invalidates neither the procedure for locating the central network nor the use of the central network as a measure of the underlying structure. The two networks are different types of perceptions of the friendship network. Individuals vary in their willingness to ascribe such relations as friendship. The number of links in each manager's network and the number of claims of friendship by each manager are listed in Table 1. Note that Manager 9 is quite exclusive, perceives little friendship among others and does not claim friendship to any other managers; however, Manager 11 is quite inclusive, perceiving both a large number of friendship ties among others and claiming thirteen friendships to other managers (Figure 3). Since the central network (Figure 1) contains an edge only if the majority agree to it, the central network can be thought of as filtering these individual differences in perception. Relationships strong enough to be observed by many, despite the opinion of the participants (such as that between Managers 9 and 5) appear in the central network, but relationships that are observed only by the participants are dropped. For example, Manager 17 claims 18 friendships (Figure 3), only one of which, that to 21 (Figure 1), is observed by the majority of the other organizational members. The central network is thus a socially robust representation of the underlying friendship network.

In Table 1 the distance of each manager's network from the central network is displayed, with the managers ordered by distance. The average distance from the central network is 33.95 with a standard deviation of 19.88. Five managers (1,5,19,11,7) stand out as almost or more than one standard deviation further from the central network than the average manager.

The responses of these managers appear different in kind from their colleagues'. Actor 7 was the chief executive officer. Referring to Table 1, we see that Manager 7 is the furthest from the central network and that Manager 7 perceives more friendship ties than any other manager. Four other managers (1, 5, 11, and 19) also give networks with very large numbers of connections. This kind of systematic structure is commonly found in social network analysis (cf. Bernard and Killworth 1977; Bernard, Killworth and Sailer 1984; Romney and Faust 1982).

We now reapply our analysis to the 16 "normal" cases, excluding the five distinctly different high-connectivity cases. Let us refer to the 16 "normal" cases as insiders and the five high-connectivity cases as outsiders. First

Table 1.

Characteristics of Managers' Networks

Manager ID	Distance from Central Graph	Number of Links	Number of Friendship Ties Sent
7	71	78	0
11	66	71	13
19	65	74	9
5	64	69	7
1	51	60	5
14	41	48	2
10	40	45	7
21	39	50	4
15	37	34	8
17	34	41	18
13	34	33	2
4	27	36	6
6	24	29	6
16	22	25	2
12	20	27	4
20	19	10	2
2	16	21	3
18	16	21	1
3	10	7	2
8	10	5	1
9	7	6	0

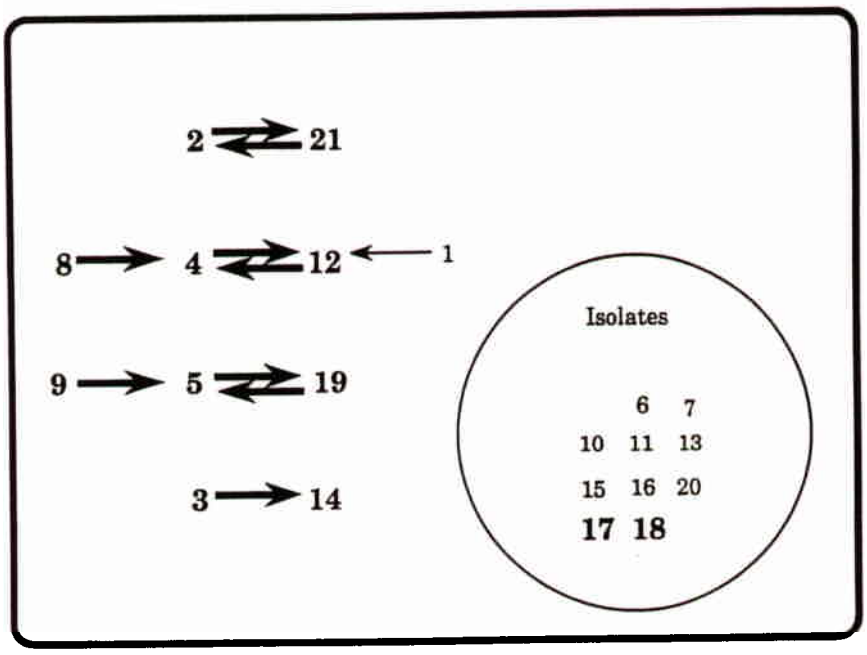

Figure 4. Central Network for Insider's View of Entire Organization.

we find the central network for friendship ties between all 21 managers, using only the digraphs of the sixteen insiders. This network can be thought of as the insiders' view of the organization. Then we consider the central network for the friendship ties between just the insiders using just the insiders' digraphs. This network can be thought of as the insiders' view of the insiders.

Insiders' view of the organization.

We recompute the central network after leaving out the digraphs for the five individuals whose digraphs are furthest from the central network. Doing so produces the estimate shown in Figure 4. Using the majority rule algorithm, we find that the maximum likelihood estimate of the central network $\hat{g}^*$ is not unique; there are four solutions that maximize the likelihood. One solution has all of the friendship ties shown in Figure 4, another adds the tie 8-4, a third adds 9-5, and the fourth has both 8-4 and 9-5. Contrasting Figures 1 and 4 we see that they differ only in that the friendship ties 17-21 and 18-2 occur in Figure 1 and not Figure 4, and tie 1-12 occurs only in Figure 4. The ties 17-21 and 18-2 are perceived by less than 45 percent of the insiders (six and seven respectively) but are known by all five outsiders. These ties can be thought of as serving to anchor the outsiders' perception of the friendship network to the insiders' perception. Indeed, the central graph formed from just the five outsiders' digraphs includes all the friendship ties in Figure 1 and several others besides. In contrast, the tie 1-12 is perceived by a majority of the insiders (9) but only by one outsider (Manager 1). And the central network for the insiders (Figure 4) contains only this tie in addition to those in the central graph for the entire organization (Figure 1).

For this reduced dataset we once again determine whether the models fits. In this case $\sigma = 2.792$ and the sum of the distances from the central network is 388. (Note that the sum of the distances from each sample network to the estimated central network does not depend upon which of the four location estimates is chosen.) Once again, we partition the space of networks according to distance from the central network. The first subset contains networks that are 0 to 21 units away from the central network, the second contains those between 22 and 24 units distant, the third those between 25 and 27 units distant, and all others are in the fourth subset. Under the null model with the fitted parameter values, the expected proportions in the four cells are .289, .243, .224 and .244, respectively. The corresponding expected numbers of sample networks in each subset are 4.63, 3.89, 3.58, and 3.90, and the observed numbers are 8, 1, 1, and 6. Standard calculation shows that the chi-squared goodness-of-fit test statistic is 7.609, with a significance probability less than .025 with $p - 2 = 2$ degrees of freedom (and .010 with $p - 3 = 1$ degrees of freedom). Without the outsiders, the fit of the model is improved,

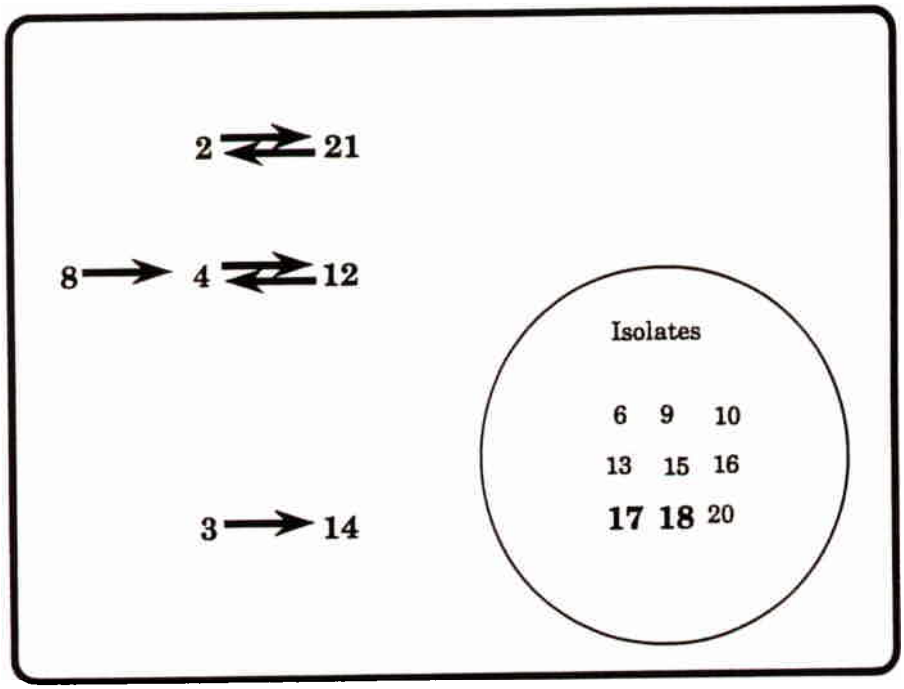

Figure 5. Central Network for Insider's View of Insiders.

although the data still suggest a large peak and heavy tails. Using the non-parametric bootstrap procedure to determine the confidence region on the central network reveals that the distances for the 90%, 95%, and 99% confidence region are, respectively, 7, 8, and 12. This finding is a much tighter confidence region than was observed for the full data set. Only three of the individual's networks fall within the region prescribed by this confidence interval, Managers 3, 8, and 9.

Insiders' view of insiders.

Now let us consider the insiders' view of just the insiders. To do so, we compute the central network for the 16 digraphs formed by eliminating the five digraphs for the outsiders, and eliminating from each of the remaining sixteen digraphs the rows/columns associated with these five outsiders. The resulting central network is displayed in Figure 5. In this case, the majority rule algorithm finds that the maximum likelihood estimate of the central network is not unique. The two solutions differ in that one has all the ties shown in Figure 5, but the other does not have edge 8-4. In the following illustrations, it is sometimes convenient to assume that the estimate is unique; when that happens, we shall use the first estimate.

The sum of the distances from each sample network to the estimated central network does not depend upon which of the two location estimates is chosen. In either case, the total distance to either $\hat{g}^*$ is 182, so the estimate from (12) is $\hat{\sigma} = 3.001$. Here $m = 16$ and we use $r = m(m-1) = 240$, since we have directed graphs without loops.

Partitioning the space of networks according to distance from the central network, we find that the first subset contains networks whose distance is 9 or less; the second subset has those with distance 10 or 11, the third those with distance 12 or 13, and the fourth consists of all other networks. Under the null model with the fitted parameter values, the expected proportions in the four subsets are .296, .238, .216, and .250, respectively. The expected numbers of sample networks in each cell are 4.73, 3.81, 3.45 and 4.00, and the corresponding observed numbers are 6, 2, 2, and 6. We find that the chi-squared goodness-of-fit test statistic is 2.809, with significance probability of .25 for the conservative test with 2 degrees of freedom, or .1 for the test with 1 degree of freedom. In either case, the model appears to give a reasonable description of the insiders' view of the insiders.

Although we find no basis to reject the model, examination of the deviations in the cells with increasing distance suggests that even for just the insiders' view of the insiders, the exponential function in (3) still gives a model that tends to be too light-tailed and insufficiently peaked for these data. The simplest explanation is that this regularity is spurious, since the test did not reject the model; however, given that the context of the problem, a core/periphery phenomenon may be operating. That is, even among the insiders, people on the fringes of the social and power centers may have a different perspective on the friendship relations. For future work, there is probably value in generalizing the exponential function used in this paper to capture such behavior. An alternative explanation for the deviations is that they partly reflect the non-uniqueness of the estimated central graph.

To gain further insight into this issue, we place a joint confidence region upon the central network and the dispersion parameter. Since the model fits, this approach can be accomplished using either a parametric or a nonparametric bootstrap. The nonparametric bootstrap drew 1000 resamples of size 16 equiprobably from the observed data. For the j-th resample, we calculated (X_j, Y_j), where X_j is the distance $d^+(\hat{g}_j, \hat{g}^*)$ between the resample estimate and the center of the empirical distribution, and Y_j is the estimated spread $\hat{\sigma}_j$ obtained from the j-th resample. This procedure gave 1000 random vectors whose empirical cumulative distribution approximates the bootstrap distribution.

Finding the smallest joint confidence region in this situation is computationally difficult. Instead, we find a normal theory approximation to this

smallest possible confidence region. To do so, we first note that although a Q-Q plot indicates that the Y_j follow a normal distribution reasonably well, the X_j tend to "cluster" near zero and have long tails. A standard normalizing transformation in such cases is to replace X_j by $W_j = 1/(1 + X_j)$ (cf. Weisberg 1980, Ch. 6, for details and alternative transformations). We then find the sample means and covariance matrix:

$$\hat{\mu} = \begin{bmatrix} \overline{W} \\ \overline{Y} \end{bmatrix} = \begin{bmatrix} .4122 \\ 3.0409 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} s_W^2 & s_{WY} \\ s_{WY} & s_Y^2 \end{bmatrix} = \begin{bmatrix} .0565 & .0060 \\ .0060 & .0204 \end{bmatrix}.$$

From multivariate normal theory, we know that if a p-variate vector $\mathbf{V}$ is normally distributed with mean μ and covariance matrix Σ, and if we estimate Σ by $\mathbf{S}$ with a sample of size N, then

$$\frac{N-p+1}{pN}(\mathbf{V}-\hat{\mu})^T \mathbf{S}^{-1}(\mathbf{V}-\hat{\mu}) \sim F_{p,N-p+1}$$

In this application, we have $N = 1000$ and $p = 2$.

As a result, the approximate $100(1-\alpha)\%$ joint confidence region on (g^*, σ) is

$$\left\{(g,s): \frac{999}{2000}\left[\left(\frac{1}{1+d(g,\hat{g}^*)}, s\right)^T - \mu\right]^T \mathbf{S}^{-1}\left[\left(\frac{1}{1+d(g,\hat{g}^*)}, s\right)^T - \mu\right] \le F_{2,999,1-\alpha}\right\}. \quad (19)$$

For a 95% confidence region, we find that $F_{2,999,.95} = 3.00$. The area of this ellipse is .6299.

The parametric bootstrap is very similar in execution. Since it uses information about the structure of the model, it tends to produce confidence regions that have less volume than those produced by the nonparametric bootstrap. We took 1000 resamples from the hypothesized model (3) with the fitted values $\hat{g}^*$ and $\hat{\sigma}$ estimated from the data. For the j-th resample, we calculated X_j and Y_j; to a marked but lesser degree than before, the Q-Q plot supported the use of the normalizing transformation that sends X_j to W_j. We thus recapitulate the previous approximation, which led to the confidence region described in (19); the differences now are that

$$\hat{\mu} = \begin{bmatrix} .9847 \\ 2.8407 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} .0080 & .0009 \\ .0009 & .0041 \end{bmatrix}.$$

We find that the 95% confidence region on (g^*,σ) has area .1066, which is appreciably smaller than the area for the nonparametric bootstrap and emphasizes the advantage of this technique when the model is correct.

Using the nonparametric bootstrap procedure to determine the confidence region on the central graph reveals that the distances for the 90%, 95%, and 99% confidence region are, respectively, 4, 5, and 6. The confidence region is again tighter than those previously located. In this case, five of the individuals' networks fall within the region prescribed by this confidence interval, Managers 2, 3, 8, 9, and 18. When considering only the insiders' view of insiders, we see that only the network of Manager 20 no longer falls within the confidence region. This set of analyses suggests that there is a central perception or a cultural consensus, and that individuals 2,3,8,9, and 18 are representative of this consensual view.

4. Conclusion

The model proposed in this paper enables researchers to frame testable hypotheses about network-valued random variables. The parameterization of the model is attractive in that the parameters correspond to interpretable properties. Using the methods described, large portions of standard statistical theory extend naturally to these applications. These include maximum likelihood estimation, hypothesis testing, confidence regions, and goodness-of-fit tests.

The metric strategy we discuss can be generalized in many ways. One approach is to consider mixtures of models of the kind we discussed to capture multimodality in the data. Also, we could use a different metric designed to capture some contextually pertinent topology, or design a metric space that is restricted to an interesting class of networks, such as trees. In most applications, we would probably want to tailor the analysis to the situation at hand.

To illustrate the potential of our methods, we reexamined a dataset from Krackhardt (1987), finding several interesting features. Although the intention was not to undertake a comprehensive analysis of this very complex data set, our simple tools did discover interpretable structure that complements and supports the conclusions of the original investigators. For Krackhardt's data, after some exploratory analysis, our model gave a good description of the data and led to both a point estimate of the central network and a confidence region upon that parameter. Although not of fundamental interest, similar results were obtained for the dispersion parameter.

It appears that this perspective on social network data has two major advantages for many situations. First, it leads to a simple mathematical description of such data. This description corresponds to a practical parameterization of the problem, using the center of the data and variation around it.

Second, the formulation is tightly connected to conventional statistical inference for data that take values in a Euclidean space. Therefore it is nearly automatic to apply a wide range of standard statistical tools. Whenever the goodness-of-fit test we propose supports the applicability of our model to the data, then the methods developed are nearly ideal.

References

BANKS, D. L. (1989), "Bootstrapping II," in *The Encyclopedia of Statistical Science*, Ed., S. Kotz, N. Johnson, and C. Read, New York: Wiley, 17-22.

BARTHÉLEMY, J. P., and MCMORRIS, F. R. (1986)., "The Median Procedure for n-Trees," *Journal of Classification, 3*, 329-334.

BARTHÉLEMY, J. P., and MONJARDET, B. (1981), "The Median Procedure in Cluster Analysis and Social Choice Theory," *Mathematical Social Sciences, 1*, 235-268.

BARTHÉLEMY, J. P., and MONJARDET, B. (1988), "The Median Procedure in Data Analysis: New Results and Open Problems," in *Classification and Related Methods of Data Analysis*, Ed., H. H. Bock, North Holland: Elsevier, 309-316.

BARTHÉLEMY, J. P., LECLERC, B., and MONJARDET, B. (1986), "On the Use of Ordered Sets in Problems of Comparison and Consensus of Classifications," *Journal of Classification, 3*, 187-224.

BERNARD, H. R., and KILLWORTH, P. (1977), "Informant Accuracy in Social Network Data II," *Human Communications Research, 4*, 3-18.

BERNARD, H. R., KILLWORTH, P., and SAILER, L. (1984), "The Problem of Informant Accuracy: The Validity of Retrospective Data," *Annual Review of Anthropology, 13*, 495-517.

BLOEMENA, A. R. (1964), *Sampling from a Graph*, Amsterdam: Mathematisch Centrum.

BOLLOBÁS, B. (1985), *Random Graphs*, London: Academic Press.

BOORMAN, S. A., and OLIVIER, D. (1973), "Metrics on Spaces of Finite Trees," *Journal of Mathematical Psychology, 10*, 26-59.

CAPOBIANCO, M. (1970), "Statistical Inference in Finite Populations Having Structure," *Transactions of the New York Academy of Sciences, 32*, 401-413.

CARLEY, K. M. (1984), *Constructing Consensus*, Unpublished doctoral dissertation, Harvard.

CARLEY, K. (1986), "An Approach for Relating Social Structure to Cognitive Structure," *Journal of Mathematical Sociology, 12*, 137-189.

CARLEY, K. (1988), "Formalizing the Social Expert's Knowledge," *Sociological Methods and Research, 17*, 165-232.

CARLEY, K. (1991), "A Theory of Group Stability," *American Sociological Review, 56*, pp. 331-354.

CONDORCET, M. (1785), *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix*, Paris.

CRAMÉR, H. (1961), *Mathematical Methods of Statistics*, Princeton, N.J.: Princeton University Press.

DAY, W. H. E. (1986), "Foreword: Comparison and Consensus of Classifications," *Journal of Classification, 3*, 183-186.

EFRON, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, SIAM Monograph No. 38, Philadelphia: CBMS-NSF.

EFRON, B., and TIBSHIRANI, R. (1986), "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science, 1*, 54-75.

FESTINGER, L. (1949), "The Analysis of Sociograms Using Matrix Algebra," *Human Relations, 2*, 153-158.

FIENBERG, S. E., MEYER, M. M., and WASSERMAN, S. S. (1985), "Statistical Analysis of Multiple Sociometric Relations," *Journal of the American Statistical Association, 80*, 51-67.

FISHER, N., and HALL, P. (1990), "On Bootstrap Hypothesis Testing," *Australian Journal of Statisticsl, 32*, 177-190.

FRANK, O. (1971), *Statistical Inference in Graphs*, Stockholm: Swedish Research Institute of National Defense.

FRANK, O. (1988), "Random Sampling and Social Networks: A Survey of Various Approaches," *Mathématiques, Informatique and Sciences Humaines, 26*, 19-33.

FRANK, O. (1989), "Random Graph Mixtures," in *Graph Theory and Its Applications: East and West, Proceedings of the First China-USA International Graph Theory Conference*, Eds., M. F. Capobianco, M. Guan, D. F. Hsu and F. Tian, *Annals of the New York Academy of Sciences, 576*, 192-199.

FRANK, O., and STRAUSS, D. (1986), "Markov Graphs," *Journal of the American Statistical Association, 81*, 832-842.

HAMMING, R.W. (1950), "Error Detecting and Error Correcting Codes," *Bell System Technical Journal, 29*, 147-160.

HAMMING, R. W. (1980), *Coding and Information Theory*, Englewood Cliffs, NJ: Prentice-Hall.

HOLLAND, P. W., and LEINHARDT, S. (1981), "An Exponential Family of Probability Distributions for Directed Graphs," *Journal of the American Statistical Association, 76*, 33-65.

HUBERT, L., and ARABIE, P. (1985), "Comparing Partitions," *Journal of Classification, 2*, 193-218.

IACOBUCCI, D., and WASSERMAN, S. (1986), "Statistical Analysis of Discrete Relational Data," *British Journal of Mathematical and Statistical Psychology, 39*, 41-64.

JARDINE, N., and SIBSON, R. (1971), *Mathematical Taxonomy*, Chichester: Wiley.

KATZ, L. (1947), "On the Matric Analysis of Sociometric Data," *Sociometry, 10*, 233-241.

KATZ, L. (1953), "A New Status Index Derived from Sociometric Analysis," *Psychometrika, 18*, 39-43.

KATZ, L., and POWELL, J. H. (1955), "Measurement of the Tendency Towards Reciprocation of Choice," *Sociometry and the Science of Man, 18*, 659-665.

KEMENY, J. G. (1959), "Mathematics Without Numbers," *Daedalus, 88*, 577-591.

KNOKE, D., and KUKLINSKI, J. D. (1982), *Network Analysis*, Beverly Hills: Sage Publications.

KRACKHARDT, D. (1987), "Cognitive Social Structures," *Social Networks, 9*, 109-134.

KRACKHARDT, D., and PORTER, L. W. (1985), "When Friends Leave: A Structural Analysis of the Relationship Between Turnover and Stayer's Attitudes," *Administrative Science Quarterly, 30*, 242-261.

LEHMANN, E. (1983), *Theory of Point Estimation*, New York: Wiley.

MACEVOY, B., and FREEMAN, L. (1988), *UCINET: A Microcomputer Package for Network Analysis*, Mathematical Social Science Group, School of Social Sciences, University of California, Irvine.

MALLOWS, C. (1957), "Non-Null Ranking Models I," *Biometrika, 44*, 114-130.

MARGUSH, T. (1982), "Distances Between Trees," *Discrete Applied Mathematics, 4*, 281-290.

MARGUSH, T., and MCMORRIS, F. R. (1981), "Consensus n-Trees," *Bulletin of Mathematical Biology, 43*, 239-244.

MCMORRIS, F. R. (1990), "The Median Procedure for n-Trees as a Maximum Likelihood Method," *Journal of Classification, 7*, 77-80.

MORENO, J. L. (1934), *Who Shall Survive?*, Washington, D.C.: Nervous and Mental Disease Publishing.

PALMER, E. (1985), *Graphical Evolution: An Introduction to the Theory of Random Graphs*, New York: Wiley.

PEARSON, K. (1900), "On the Criterion That a Given System of Deviations From the Probable In the Case of a Correlated System of Random Variables Is Such That It Can Be Reasonably Supposed to Have Arisen From Random Sampling," *Philosophy Magazine, 50*, 157-172.

RANDLES, R., and WOLFE, D. A. (1979), *Introduction to the Theory of Nonparametric Statistics*, New York: Wiley.

ROMNEY, A. K., and FAUST, K. (1982), "Predicting the Structure of a Communications Network from Recall Data," *Social Networks, 4*, 285-304.

STRAUSS, D., and IKEDA, M. (1990), "Pseudolikelihood Estimation for Social Networks," *Journal of the American Statistical Association, 85*, 204-212.

WANG, Y., and WONG, G. Y. (1987), "Stochastic Blockmodels for Directed Graphs," *Journal of the American Statistical Association, 82*, 8-19.

WASSERMAN, S. (1987), "Conformity of Two Sociometric Relations," *Psychometrika, 52*, 3-18.

WASSERMAN, S., and ANDERSON, C. (1987), "Stochastic a posteriori Blockmodels: Construction and Assessment," *Social Networks, 9*, 1-36.

WASSERMAN, S., and GALASKIEWICZ, J. (1984), "Some Generalizations of p_i: External Constraints, Interactions, and Non-Binary Relations," *Social Networks, 6*, 177-192.

WEISBERG, S. (1980), *Applied Linear Regression*, New York: Wiley.

WONG, G. Y. (1987), "Bayesian Models for Directed Graphs," *Journal of the American Statistical Association, 82*, 140-148.

YOUNG, H. P. (1988), "Condorcet Theory of Voting," *American Political Science Review, 82*, 1231-1244.