



COMPARING NETWORKS ACROSS SPACE AND TIME, SIZE AND SPECIES

*Katherine Faust**
John Skvoretz†

We describe and illustrate methodology for comparing networks from diverse settings. Our empirical base consists of 42 networks from four kinds of species (humans, nonhuman primates, nonprimate mammals, and birds) and covering distinct types of relations such as influence, grooming, and agonistic encounters. The general problem is to determine whether networks are similarly structured despite their surface differences. The methodology we propose is generally applicable to the characterization and comparison of network-level social structures across multiple settings, such as different organizations, communities, or social groups, and to the examination of sources of variability in network structure. We first fit a p^ model (Wasserman and Pattison 1996) to each network to obtain estimates for effects of six structural properties on the probability of the graph. We then calculate predicted tie probabilities for each network, using both its own parameter estimates and the estimates from every other network in the collection. Comparison is based on the similarity between sets of predicted tie probabilities. We then use correspondence analysis to represent the similarities among all 42 networks and interpret the resulting configuration using information about the species and relations involved. Results*

*University of California, Irvine

†University of South Carolina

We acknowledge the helpful comments of the editor and anonymous reviewers. For their encouragement and suggestions on the research, we thank H. Russell Bernard, Linton Freeman, and A. Kimball Romney. We thank Tracy Burkett and Douglas Nigh for making their data available to us.

show that similarities among the networks are due more to the kind of relation than to the kind of animal.

1. INTRODUCTION

Much of social network analysis examines a single network at a time. Commonly the analyses comprise case studies of network properties or processes within a single community. For example, dominance relations among chimpanzees are described and the structure of the network analyzed. Or the liking and disliking relations among novices in a monastery are described and the patterning in these networks related to observations about group structure and dynamics. The problem of comparing networks arises more rarely, and when it does the usual context is that of comparing two relations mapped on the same population during the same time period. For example, possible associations between friendship and advice seeking among corporate managers may be studied by comparing the two relations.

In this paper we expand the scope of comparison by describing a general way in which two, three, . . . , many networks can be compared at the same time even though they differ widely in size, type of relation, species of the units, and time and space of the observations. The general question concerns determining whether the networks are similarly structured despite their surface differences. The method we propose and illustrate allows us not only to compare two networks at a time but to look at the overall patterning of similarities among a large collection of networks from diverse settings. Our empirical base consists of 42 different networks from four kinds of species (humans, non-human primates, nonprimate mammals, and birds), varying in size from 7 to 103 units, and covering distinct types of relations such as influence, grooming, and agonistic encounters. Although we illustrate the methodology on a collection of relatively exotic networks, it can be easily applied to a wide range of more familiar substantive situations, such as comparing advice networks among managers in different firms, friendships among schoolchildren in different classrooms, referrals between service agencies in various communities, and so on.

Six of these networks are diagrammed in Figure 1. The diversity in our collection is apparent from the figures. All are directed graphs. In some, the original data refer to counts. We dichotomize these data, regarding any nonzero count as indicating the presence of a tie. Network 1(a) derives from the observation of agonistic encounters between red deer: A

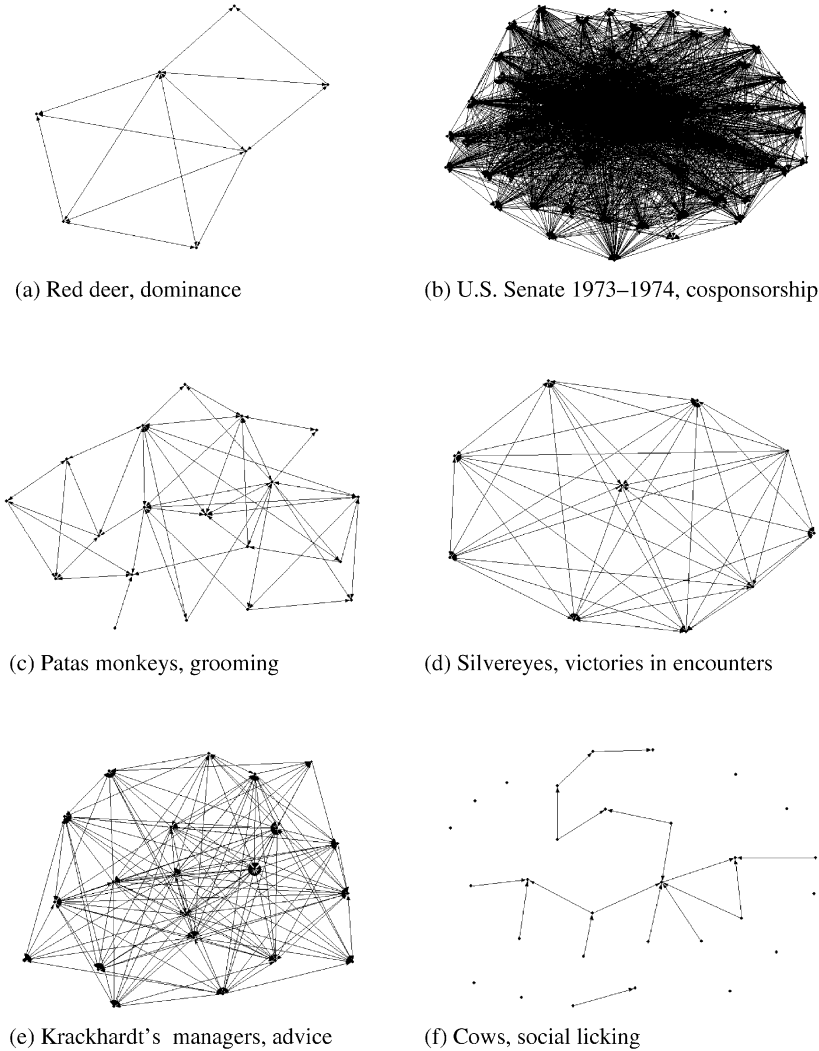


FIGURE 1. Graphs of six networks.

tie exists from animal i to animal j if the first defeated the second in an encounter (Appleby 1983). Network 1(b) diagrams the cosponsorship ties among U.S. senators in the Ninety-third Congress (1973–1974): A tie exists from senator i to senator j if the first cosponsored at least one bill introduced by the second (Burkett 1997). Network 1(c) graphs grooming relations among patas monkeys: The presence of a tie from monkey i to

monkey j indicates that the first groomed the second at least once (Kaplan and Zucker 1980). Network 1(d) depicts victories in encounters among birds called silvereyes: The presence of a tie from i to j indicates that silvereye i was victorious in at least one encounter with j (Kikkawa 1980). Network 1(e) graphs the advice relations among a group of high-tech managers: There is a tie from manager i to manager j if i reports going to j for advice (Krackhardt 1987). Finally, network 1(f) diagrams social licking among cows: There is a tie from cow i to cow j if the first licks the second (Reinhardt and Reinhardt 1981).

Our problem of the comparison of networks can now be posed rather dramatically: Is the network of cosponsorship among senators structurally more similar to the network of social licking among cows, the network of grooming among monkeys, or the network of advice among managers? Or, are the networks of victory in encounters among silvereyes and of dominance among red deer similarly structured? Such questions are substantively interesting and theoretically provocative, but they cannot be addressed systematically without general methods for the comparison of networks. Such methods would enable us to answer certain questions: What structural features are similar or different among networks of different kinds of organisms or different kinds of relations? Which kinds of networks tend to be similarly structured and which tend to be different? The present work contributes to research on these deeper issues.

In the next section we review the relatively sparse literature on the comparison of networks. We then outline the formal background for our approach. We use the p^* modeling framework to build and estimate models for the probability of a graph as a function of its structural properties. The estimates from these models, in turn, form the basis from which the similarity or dissimilarity of pairs of networks is calculated. Correspondence analysis provides a way of representing the similarities among all networks under consideration. We then interpret the resulting configuration using information about the networks and their structural properties. We apply this strategy to 42 networks and discuss the results.

2. COMPARING NETWORKS

The vast majority of social network studies are case studies of individual communities. Nevertheless, comparison of networks can, and does, proceed along several lines. The most straightforward case is the comparison of two networks over the same set of actors. For instance, two different

relations could be measured on the same set of actors or the same relation could be measured on one set of actors at two time points. Methodology for comparison of two relations measured on the same set of actors dates to the early years of social network analysis (Katz and Powell 1953) and has been elaborated by Hubert and colleagues in a matrix permutation context (Hubert and Baker 1978; Baker and Hubert 1981). Moreover, statistical models for multiple relations are well developed (Wasserman 1987; Pattison and Wasserman 1999). There are also models for longitudinal networks, where the same relation is measured on the same set of actors at two (or more) points in time (Wasserman and Iacobucci 1988; Snijders 1996; Snijders and VanDuijn 1997).

Another type of comparison, replication, arises when the same relation is measured on two (or more) different sets of actors. Researchers are usually concerned with whether the networks exhibit similar structural properties or relationships or whether nonnetwork properties of the groups are associated with network-level properties. Examples include both classics, such as Laumann and Pappi's study of elite networks in the communities of Altneustadt and Towertown (Laumann and Pappi 1976) and Hallinan's (1974) studies of sentiment structures in school groups, and more recent studies such as Shrader, Lincoln, and Hoffman's (1989) study of networks in 36 agencies, Johnson and Boster's study of winter-over research teams at the South Pole (Johnson, Boster, and Palinkas n.d.), the National Longitudinal Study of Adolescent Health replication of friendship networks across schools (Bearman, Jones, and Udry 1997), and Rindfuss and Entwisle's studies of networks of kinship and social and economic relations in 51 villages in Nang Rong District, Thailand (Rindfuss et al. 2000). Until recently, methodology for the comparison of replicated networks was primarily descriptive. For example, Breiger and Pattison's comparison of elite structures in two communities used joint homomorphic reduction of the semigroup algebras in the two communities (Breiger and Pattison 1978). Recently, however, Anderson et al. (1999) and Martin (1999) describe statistical approaches that evaluate whether a common set of parameter estimates provides adequate fit to two or more networks.

A fourth type of comparison arises when data on roughly similar relations are available from different settings with different sets of actors. Unlike the situation just described, which is "pure" replication, relations in this case are only roughly comparable. The classic series of studies by Davis (1979) and by Holland and Leinhardt (1978) using the sociometric data bank of several hundred sociomatrices is a case in point. The studies asked whether sociometric data from diverse sources tended to exhibit

greater than chance tendencies for transitivity, balance, or clustering. They calculated standard transitivity statistics on each network and then examined the distribution of the scores. Another example is studies of informant accuracy in different settings (for example, Bernard and Killworth 1977; Bernard, Killworth, and Sailer 1980), where observational data were collected in different ways depending on the setting (e.g., monitoring radio transmissions among ham radio operators, or observing interactions in the office or fraternity). Similarly, Freeman (1992) compiled examples of observations of interactions among people in seven different communities to explore the question of which of two alternative grouping models was more consistent with the observed interactions. In these examples of comparison, interest centers on whether hypothesized structural patterns or relationships are found across a range of roughly similar settings.

Common to the examples cited is the fact that the comparisons involve communities of identical actor types, usually humans. Only rarely have comparisons been made between networks of different kinds of organisms—for example, different animal species (Sade and Dow 1994). A notable exception is Maryanski's (1987) comparison of weak and strong ties in gorilla and chimpanzee social networks.

A more abstract and methodologically more challenging type of comparison arises when networks not only have different actor sets but also vary greatly in size, have substantively different relations, and include actors that are different kinds of organisms. The methods we propose address this problem. Our overarching question is whether pairs or sets of networks are similarly structured despite being based on substantively different relations measured on quite different kinds of organisms. An important contrast between our approach and previous methods for network comparison is that it measures directly the similarity between pairs of networks rather than simply determining whether (or to what degree) each exhibits specific structural tendencies. That is, our method provides an index, akin to a correlation coefficient, that quantifies the degree of similarity between two networks. An additional contrast with previous methods derives from the number of networks our method compares simultaneously. In the most straightforward case of comparison—two networks over the same set of actors—several measures of association can be calculated and evaluated—for example, using matrix permutation tests (Hubert and Baker 1978; Baker and Hubert 1981) or estimating multiplexity parameters in statistical models for multiple relations (Wasserman 1987; Pattison and Wasserman 1999). Extending comparisons to more than two networks requires calculating similarities between all pairs of networks

and analyzing these simultaneously. In the method we propose, the idea is to represent the similarities among all networks in the set being compared via scaling or clustering techniques to depict graphically the similarity space of networks.

In overview, our approach consists of four steps: (1) We characterize each network in terms of a set of structural properties using a statistical model for the probability of the graph; (2) we measure the similarity between networks based on parameter estimates for the structural properties in the models as they predict network tie probabilities; (3) we represent the similarities among the networks using a spatial model; and (4) we interpret the resulting spatial configuration using information about the networks. Each of these steps involves decisions about possible alternative approaches, which we discuss as they arise in our description below and consider in detail in the discussion.

3. FORMAL BACKGROUND

Our aim is to assess whether two, three, . . . , many networks are similarly structured despite their surface differences. We argue that two networks are similarly structured to the extent that they exhibit the same structural tendencies, to the same degree. Obviously there are numerous structural tendencies that could be used to characterize networks, and selection centers on which properties are argued to be theoretically important for characterizing the networks at hand. Some widely used properties provide the basis for our comparison: mutuality, transitivity, cyclical triples, and star configurations (in-stars, out-stars, and mixed stars), as illustrated below. To be precise, we focus on predicting the probability of a network from a profile of structural properties of the network. Two networks are similarly structured to the extent that the probabilities of both networks depend on the same set of structural properties, used as predictor variables in the models, and on each property to the same degree.

We draw upon recent developments in the statistical modeling of networks—in particular, the development of models known as p^* models (Anderson et al. 1999; Crouch et al. 1998; Pattison and Wasserman 1999; Wasserman and Pattison 1996; Robins, Pattison, and Wasserman 1999). Statistical models for networks were long based on the assumption of dyadic independence. Dyadic independence means that the presence or absence of a tie in the ij dyad is independent of the presence or absence of a tie in any other dyad. It is widely recognized that this assumption clearly oversimplifies matters. As one example of the inappropriateness of the

dyadic independence assumption, triadic effects such as the presence of an ij tie being significantly more likely if there are several others k who have ties to i and to j abound in real social networks. Modeling these effects is beyond the capability of statistical models that assume dyadic independence. The new statistical approaches, the p^* family of models, explicitly model nonindependence among dyads by including parameters for structural features that capture hypothesized dependencies among ties.

In the p^* framework, the probability of a digraph G is expressed as a log-linear function of a vector of parameters θ , an associated vector of digraph statistics $x(G)$, and a normalizing constant $Z(\theta)$:

$$P(G) = \frac{\exp(\theta'x(G))}{Z(\theta)}. \quad (1)$$

The normalizing constant ensures that the probabilities sum to unity over all digraphs. The θ parameters express how various “explanatory” properties of the digraph affect the probability of its occurrence. Different models use different profiles of digraph properties. Our models use a profile of six structural properties: mutuality, out 2-stars, in 2-stars, mixed 2-stars, transitive triples, and cyclical triples, as diagrammed in Figure 2.

Taken together, these effects constitute a Markov graph model (Frank and Strauss 1986) in which the probability of the ij tie depends only on other ties in which i and j might be involved but not on ties that involve neither i nor j . This model includes substantively interesting dyadic and triadic effects and provides a base to which higher order network properties (such as subgrouping or graph connectivity) might later be added. These effects are assumed to be homogeneous. The homogeneity assumption means that a particular structural property has the same effect regardless of the specific individual nodes involved. Obviously when comparing networks of different individuals the homogeneity assumption is desirable. Under the assumption of homogeneity, then, our model stipulates that the probability of a graph is a log-linear function of the number of mutual dyads, the number of out 2-stars, the number of in 2-stars, etc. If the resulting parameter estimate for a specific property is large and positive, then graphs with that property have large probabilities. For example, if the mutuality property has a positive coefficient, then a graph with many mutual dyads has a higher probability than a graph with few mutual dyads. Or, if the cyclical triple property has a negative coefficient, then a graph with many cyclical triples has a lower probability than a graph with few cyclical triples. Thus, the resulting parameter estimates associated with

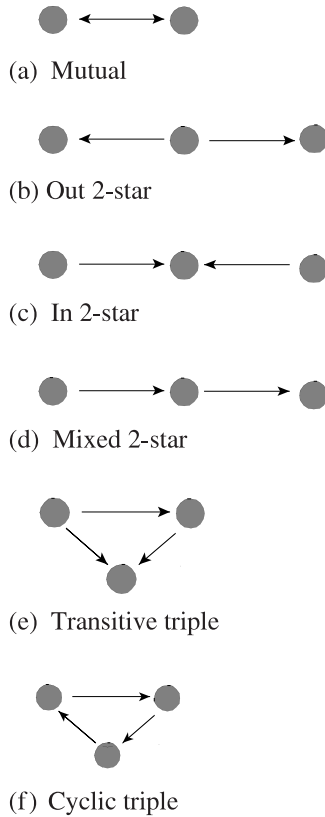


FIGURE 2. Network properties included in the p^* models.

the structural properties capture the importance of their respective properties for characterizing the network under study.

Conceptually, the models are easy to understand. The real difficulty comes in trying to estimate the effect coefficients. Consider the estimation problem. Suppose we assigned a set of values to the effect coefficients. Then for each digraph realization over the set of all digraphs for a particular node set of size g , we could calculate the numerator of equation (1). Summing the numerator over all realizations yields the normalizing constant in the denominator of equation (1). The probability of a particular digraph realization is then given by the ratio of its numerator to the normalizing constant. One particular realization is the observed digraph. The estimation problem can be thought of as finding an assign-

ment of values to effect coefficients that maximizes the probability of the observed digraph. Conceptually, of course, these are the maximum-likelihood estimates of the effect parameters. One could imagine estimation by a numerical search procedure through an orthogonal space of parameter values. But the number of combinations to be searched and the number of digraph realizations to be calculated on each pass are so huge for even relatively small networks that such a procedure is simply not practicable. Clearly, direct analysis via the solution of simultaneous differential equations for values that maximize equation (1) is equally out of the question.

The literature proposes a way of out of this impasse. The estimation approach, suggested by Strauss and Ikeda (1990) and elaborated by Wasserman and Pattison (1996), uses equation (1) to express the probability of tie, conditional on the rest of the digraph:

$$P(x_{ij} = 1 | G^{-ij}) = \frac{P(G^+)}{P(G^+) + P(G^-)}, \quad (2)$$

where G^{-ij} is the digraph including all adjacencies except the i, j^{th} one. The digraph G^+ is defined by the adjacency matrix plus $x_{ij} = 1$ while G^- is defined as the adjacency matrix plus $x_{ij} = 0$. This equation expresses the probability that $x_{ij} = 1$ conditional on the rest of the graph. Note that equation (2) does not depend on the normalizing constant because upon rewriting we get

$$P(x_{ij} = 1 | G^{-ij}) = \frac{\exp(\theta'x(G^+))}{\exp(\theta'x(G^+)) + \exp(\theta'x(G^-))}. \quad (3)$$

The conditional odds of the presence of a tie from i to j versus its absence is expressed by

$$\frac{P(x_{ij} = 1 | G^{-ij})}{P(x_{ij} = 0 | G^{-ij})} = \frac{\exp(\theta'x(G^+))}{\exp(\theta'x(G^-))}. \quad (4)$$

From equation (4), we derive the log of the odds or *logit* model:

$$\text{logit } P(x_{ij} = 1 | G^{-ij}) = \theta' [x(G^+) - x(G^-)]. \quad (5)$$

The quantity in brackets on the right side is a vector of differences in the profile of structural properties (which are assumed in equation (1) to affect the probability of the digraph) when x_{ij} changes from 1 to 0. Finally, we

can derive an equation for the probability that $x_{ij} = 1$, conditional on the rest of the digraph, from equation (5):

$$P(x_{ij} = 1 | G^{-ij}) = \frac{\exp(\theta'(x(G^+) - x(G^-)))}{1 + \exp(\theta'(x(G^+) - x(G^-)))}. \quad (6)$$

The estimation method proposed by Strauss and Ikeda (1990) forms a pseudolikelihood function for the graph in terms of the conditional probabilities for x_{ij} as follows:

$$PL(\theta) = \prod_{ij} P(x_{ij} = 1 | G^{-ij})^{x_{ij}} P(x_{ij} = 0 | G^{-ij})^{1-x_{ij}} \quad (7)$$

Strauss and Ikeda prove that equation 7 can be maximized using maximum-likelihood estimation of the logistic regression, equation (5), assuming the x_{ij} 's are independent observations. Thus the p^* family of models can be estimated, albeit approximately, using logistic regression routines in standard statistical packages. However, since the logits are not independent, the model is not a true logistic regression model and statistics from the estimation must be used with caution. Because goodness-of-fit statistics are pseudolikelihood ratio statistics, it is questionable whether the usual chi-square distributions apply, and standard errors have only "nominal" significance (see Crouch and Wasserman 1998). These reservations have little or no importance in our use of the p^* framework. We are not concerned with exactly how good a fit a particular model has to a particular network. Nor are we concerned with identifying just those coefficients that are statistically "significant." Instead we use estimates from the model in conjunction with the calculated changes in graph statistics to calculate an estimated probability for each ij tie in the network.

For each of our data sets, we estimate a p^* model that expresses the probability of a tie being present (conditional on the rest of the graph) as a function of the six structural properties diagrammed in Figure 2. Fitting the p^* model results in estimates of θ 's for the effects of each of the graph properties hypothesized to affect the likelihood of a tie. These estimates express the importance of the properties for the probability of the graph, but they can also be used (via equation (6)) to calculate the probabilities of the individual ties in the network. We use all parameter estimates to calculate predicted probabilities regardless of their level of statistical significance.

With these considerations in hand, we may return to the question of whether two networks are similarly structured. Consider two networks,

A and B, in which the θ 's from the p^* model (equation (6)) are similar in direction and magnitude. We would argue that these two networks are similarly structured in that the same structural tendencies are important, and important to the same degree, in predicting tie probabilities in both networks. In such a case we should be able to predict the tie probabilities in one network not only from its own parameter estimates but also from the parameter estimates of its "twin." On the other hand, this would not be the case if the p^* models for two networks resulted in quite different estimates of the θ 's.

An important general principle for comparison is that the magnitudes of the effects should be independent of scale differences in the explanatory variables in the models. The networks we compare vary widely in size and density, leading to distributional differences in the explanatory variables—the change statistics $x(G^+) - x(G^-)$. Thus, for comparison, the effects should be expressed as standardized logistic regression coefficients. Two networks are similarly structured if network structural properties have the same impact, net of distributional differences in the explanatory variables; that is, if the impact is the same in standardized terms.

Comparison can now proceed at different levels. First, we could directly compare the standardized parameter estimates from models for different networks. Alternatively, we could use sets of parameter estimates to get predicted tie probabilities for the networks and then compare these predicted probabilities. We use the second mode of comparison for three reasons: (1) We are interested in the collection of structural effects that characterize the network rather than individual parameter comparisons; (2) we are fundamentally interested in the structure of the network as manifested in the tie probabilities predicted by the network structural effects; and (3) resemblance between networks based on predictions from the parameter estimates may be asymmetric; parameter estimates from network A may predict ties in network B better than parameter estimates from B predict ties in A.

The task of comparing networks proceeds by using the standardized parameter estimates for one data set to predict tie probabilities for every other data set in the collection, in a pair-wise fashion. Predictions are made using equation (6) but entering the standardized parameter estimates from one network and the standardized change statistics ($x(G^+) - x(G^-)$) from the network that is being predicted. We do this for each pair of networks. The result is a set of predicted tie probabilities for each network, one based on its own p^* parameter estimates and the rest based on

the estimates from the other networks. The next step assesses the relative similarity between one set of parameter estimates and another set of estimates via their predicted tie probabilities. We now turn to a description of this step in the comparison process.

4. DATA AND METHODOLOGY OF COMPARISON

Table 1 lists the 42 data sets we use to illustrate our methodology of comparison. The networks range in size from 7 red deer stags to 104 U.S. senators. The ties composing the networks also vary from grooming relations and advice seeking to victories in agonistic encounters. Each of the networks that we compare is represented by a 0,1 adjacency matrix (created by dichotomizing all nonzero entries equal 1 if the original relation was valued). More details about each of the data sets can be found in the Appendix.

The strategy of comparison consists of four steps: (1) for each data set, we estimate a p^* model that expresses the conditional probability of a tie as a function of six structural factors: mutuality, out 2-stars, in 2-stars, mixed 2-stars, transitive triples, and cyclical triples (since we use standardized estimates there is no intercept); (2) we use these standardized parameter estimates and the standardized change scores in these structural factors to calculate the predicted probability of a tie in each i, j pair in each data set using as coefficients the parameter estimates from its own model and from each of the remaining 41 models. Thus for each data set, we have 42 sets of predicted probabilities, one from each set of parameter estimates including the set of estimates from the focal data set itself. The third step calculates a (dis)similarity score between the predicted probabilities from the estimates on the focal data set and each of the other 41 sets of predicted probabilities. The fourth and final step uses correspondence analysis to represent the proximities among all of the networks, using as input the 42 by 42 matrix of (dis)similarity scores. To illustrate the methodology, we can follow through the steps for the six networks diagrammed in Figure 1.

For these six networks, the results of the p^* model estimation are displayed in Table 2. The estimates vary considerably and many of the coefficients are not statistically significant at the $p < .05$ level. However, all estimates are retained in the prediction equation regardless of their nominal statistical significance. In one case, "cows, licking," the full model cannot be estimated due to multicollinearity among the predictor variables. In that case and three others like it, we use the model that is estima-

TABLE 1
List of Networks

Label	Network	Relation	p* parameter profile						N
			mut	trans	cycle	ostar	istar	mstar	
1. s93	U.S. Senate, 1973–74	Influence/cosponsorship		+	–	+	+		103
2. s94	U.S. Senate, 1975–76	Influence/cosponsorship		+		+	+	–	101
3. s95	U.S. Senate, 1977–78	Influence/cosponsorship	+	+		+		–	104
4. s96	U.S. Senate, 1979–80	Influence/cosponsorship	+	+		+		–	101
5. s97	U.S. Senate, 1981–82	Influence/cosponsorship	+	+		+		–	101
6. s98	U.S. Senate, 1983–84	Influence/cosponsorship	+	+		–	–	–	101
7. s99	U.S. Senate, 1985–86	Influence/cosponsorship	+	+		–	–	–	102
8. s100	U.S. Senate, 1987–88	Influence/cosponsorship	+	+			–	–	101
9. s101	U.S. Senate, 1989–90	Influence/cosponsorship	+	+		–	–	–	102
10. krack	Krackhardt’s managers	Advice	+			+	+	–	21
11. sampin	Sampson’s Monastery	Influence, positive	+	+			+		18
12. sampnin	Sampson’s Monastery	Influence, negative	+	–		–		–	18
13. sampnpr	Sampson’s Monastery	Blame	+	–	+	+	+	–	18
14. samppr	Sampson’s Monastery	Praise		+		–			18
15. ua02	Athanassiou & Nigh TMT 2	Advice				+			12
16. ue02	Athanassiou & Nigh TMT 2	Worked together	+						12
17. ua06	Athanassiou & Nigh TMT 6	Advice						+	11
18. ue06	Athanassiou & Nigh TMT 6	Worked together			–	+			11
19. chimp1	Chimpanzees	Pant grunt calls	–		+			–	9
20. chimp2	Chimpanzees	Agonistic					–	–	9
21. chimp3	Chimpanzees	Grooming	+						9
22. macaca	<i>Macaca mulatta</i>	Grooming	+	+				–	16

23. macaqa	<i>Macaca sylvanus</i>	<i>i</i> carries baby away from <i>j</i>	+		+				8
24. macaqb	<i>Macaca sylvanus</i>	<i>i</i> leaves baby w/ <i>j</i>			-			-	8
25. macaqc	<i>Macaca sylvanus</i>	<i>i</i> w/baby approaches <i>j</i>							8
26. macaqd	<i>Macaca sylvanus</i>	<i>j</i> w/baby approached by <i>i</i>	-	+	+	-	-	-	8
27. macaqu	<i>Macaca artaides</i>	Aggression	-						14
28. patasf	Patas monkeys, female	Fight	-						18
29. patasg	Patas monkeys	Groom	+		+	+			19
30. vervet1a	Vervet monkeys, juveniles	Aggressive/submissive							14
31. vervet1m	Vervet monkeys, juveniles	Aggressive/submissive		-		+	+		14
32. vervet2a	Vervet monkeys, juveniles	Aggressive/submissive	-		-				11
33. vervet2m	Vervet monkeys, juveniles	Aggressive/submissive						-	11
34. cowg	Cows, <i>bos indicus</i>	Grazing preference	+				+		29
35. cowl	Cows, <i>bos indicus</i>	Licking	+	+			+		29
36. hyenaf	Hyaena, female, <i>crocuta crocuta</i>	Dominance		+	-	+			25
37. hyenam	Hyaena, male, <i>crocuta crocuta</i>	Dominance				+			13
38. ponies	Highland ponies	Threat			-				17
39. reddeer	Red deer stags, <i>cervus elaphus L.</i>	Dominance	+						7
40. silver	Silvereyes, <i>zosterops lateralis</i>	Victory in encounter							10
41. sparrow	Harris' sparrows	Dominance	-		-				26
42. tits	Willow tits	Dominance		+					8

TABLE 2
p* Parameter Estimates for Six Networks

	Red Deer	Senate 93 rd	Patas	Silvereyes	Managers	Cows, Lick
Intercept	4.291	-5.773*	-2.308*	3.801	-3.676*	-4.179*
Mutual	6.520*	-0.074	1.813*	1.057	1.714*	2.032*
O-Star	-0.795	0.061*	0.152*	-0.347	0.260*	0.115
I-Star	-8.482	0.066*	-0.045	-0.741	0.249*	0.488*
M-Star	2.689	0.002	-0.131	-0.443	-0.152*	-0.458
Trans	2.085	0.007*	0.092	0.450	0.039	1.402*
Cycle	-5.973	-0.017*	0.647*	0.180	0.102	...

*Significant at $p < .05$

ble and contains the greatest number of original structural factors. The omitted one(s) are set equal to zero. Inspection of the parameter estimates reveals several similarities and differences. For instance, mutuality has a positive effect in all six networks but “Senate 93rd”. Transitivity has a positive effect in all six networks—that is, completing a triple transitively tends to be an important property in all of them, although the size of the effect varies considerably from one network to another.

Table 3 presents six sets of predicted tie probabilities for a portion of the red deer network, one made by its own parameters and the others by the estimates from the other five networks. In general, if another network has a structure similar to the “red deer” network, then its model should provide predicted probabilities that are close to the probabilities predicted from the “red deer” model itself. That is, we would expect the average difference between predictions to be small.

To calculate the dissimilarity between the predicted probabilities, we use the Euclidean distance function:

$$d(t, y) = \sqrt{\frac{\sum_{i,j} (p_t(i, j) - p_y(i, j))^2}{g_t(g_t - 1)}}, \quad (8)$$

where $p_t(i, j)$ is the standardized tie probability for pair (i, j) predicted from the target network t 's model, $p_y(i, j)$ is the standardized tie probability for pair (i, j) predicted from network y 's model, and g_t is the number of nodes in network t . Results for our six illustrative networks are given

TABLE 3
 Predicted Tie Probabilities for Red Deer Dominance

<i>i</i>	<i>j</i>	Obs x_{ij}	Model Providing Predictions					Cows, Lick
			Red Deer	Senate 93rd	Patas	Silvereyes	Managers	
1	2	1	0.99097	0.60750	0.80128	0.78903	0.75413	0.50732
1	3	1	0.96682	0.60753	0.69805	0.66990	0.82355	0.59876
1	4	1	0.99978	0.32027	0.66999	0.66996	0.60650	0.40379
1	5	1	0.99535	0.61614	0.74646	0.74403	0.76769	0.53816
1	6	1	0.98674	0.32624	0.38869	0.38818	0.37927	0.3266
1	7	1	0.87031	0.63108	0.40235	0.41386	0.58531	0.48569
2	1	1	0.99264	0.80810	0.70941	0.74697	0.78741	0.63561
2	3	1	0.37954	0.52914	0.56319	0.63587	0.53177	0.47484
2	4	1	0.99802	0.24093	0.66128	0.68044	0.57629	0.42665
2	5	1	0.99847	0.54663	0.65779	0.72511	0.70561	0.56147
2	6	0	0.59946	0.62525	0.43461	0.47461	0.66197	0.54392
2	7	1	0.69987	0.54721	0.40736	0.51333	0.56865	0.53659
3	1	1	0.60305	0.55773	0.56403	0.44505	0.45961	0.47982
3	2	0	0.69020	0.65131	0.65062	0.64060	0.66261	0.59907
3	4	0	0.39391	0.35465	0.44473	0.47845	0.24874	0.35147
3	5	1	0.89568	0.27196	0.64230	0.70525	0.44261	0.48479
3	6	0	0.35590	0.36566	0.32775	0.41485	0.34780	0.46709
3	7	0	0.13754	0.68110	0.35404	0.52912	0.56595	0.65571
...

in Table 4. We find, for instance, that the model that best predicts the “red deer” target (other than the “red deer” model itself) is the model for encounters between silvereyes. The model of social licking among cows best predicts, as a target, cosponsorship among U.S. senators in the Ninety-

TABLE 4
 Distances Between Networks

Target Network	Model Providing Predictions					
	Red Deer	Senate 93rd	Patas	Silvereyes	Managers	Cows, Lick
Red Deer	0	0.07082	0.05098	0.04568	0.05477	0.06412
Senate 93rd	0.00505	0	0.00232	0.00229	0.00117	0.00128
Patas	0.01969	0.01186	0	0.00714	0.00680	0.00706
Silvereyes	0.03438	0.02214	0.00724	0	0.01861	0.01610
Managers	0.02386	0.00580	0.00997	0.00977	0	0.00616
Cows, Lick	0.01678	0.00565	0.00427	0.00662	0.00533	0

Third Congress. However, that network of cosponsorship ties is best predicted by the model for advice seeking between managers, among the five alternatives. Note that the matrix of distances is not symmetric. In fact, there is no reason to expect symmetry—the (i, j) cell expresses the distance between the predicted probabilities of target network i 's ties from network j 's model, while the (j, i) cell expresses the distance between the predicted probabilities of target network j 's ties from network i 's model. The final step takes the full matrix version of Table 4 (transformed to similarities) and scales it using correspondence analysis. These results are reported and interpreted in the next section.

5. REPRESENTING SIMILARITIES AMONG NETWORKS: CORRESPONDENCE ANALYSIS

We use correspondence analysis to represent the similarities among the networks. Correspondence analysis (Greenacre 1984; Weller & Romney, 1990) is a data analytic technique for studying two-way arrays such as contingency tables or similarity matrices. It is one of several closely related scaling approaches, also including dual scaling (Nishisato 1994), homogeneity analysis (Gifi, 1990), and optimal scaling. It aims to represent proximity data in a low-dimensional space using scores for categories of the variables. These scores then serve as coordinates in graphical displays in which points represent the categories of the variables and the distance between points represents the similarity between their respective entities. We use as input the matrix of distances between networks, appropriately transformed into similarities by subtracting each value from a large positive number. As Carroll, Kumbasar, and Romney (1997) show, this is equivalent to multidimensional scaling of the original distances. The advantage of correspondence analysis is that it can be used to analyze nonsymmetric matrices, such as the distances between the target networks and the networks providing the model predictions. In our application, two networks will be close in space if the predictions provided by their models are similar, in the sense that they similarly predict other networks in the collection.

Correspondence analysis is accomplished through a singular value decomposition of an appropriately scaled matrix. Entries in the input matrix are divided by the square root of the product of the row and column marginal totals, prior to singular value decomposition. Let \mathbf{F} be a rectangular

matrix of positive entries. \mathbf{R} and \mathbf{C} are diagonal matrices with entries equal to the row and column totals of \mathbf{F} , respectively. Correspondence analysis consists of a singular value decomposition of the matrix $\mathbf{R}^{1/2}\mathbf{F}\mathbf{C}^{1/2}$

$$\mathbf{R}^{1/2}\mathbf{F}\mathbf{C}^{1/2} = \mathbf{U}\mathbf{D}\mathbf{V}, \quad (9)$$

where \mathbf{D} is a diagonal matrix of singular values, and \mathbf{U} and \mathbf{V} are row and column vectors, respectively. For visual displays, \mathbf{U} and \mathbf{V} are rescaled. We use principal coordinates, where, on each dimension, the weighted mean is equal to 0 and the weighted variance is equal to the singular value squared. In the following graphs we present the column scores from correspondence analysis of the matrix of similarities among the networks. Column scores show similarities among networks in terms of the predictions they make for other networks. Row scores would show similarities among the targets being predicted. We should note that for our analyses using the row scores leads to essentially the same results and conclusions as those presented here.

To interpret the correspondence analysis configuration, we employ information about the networks and about the species and relations that are involved. There are four kinds of species: human, nonhuman primate, mammal, and bird. Relations are first categorized by how they were collected: observation or report by respondent. Obviously this is confounded with the type of animal since only humans provided reports of their ties to others. We then categorize the relation as either positive or negative. Grooming, advice seeking, cosponsorship, and working together are considered positive, whereas dominance, agonistic encounters, and blaming are negative. This leads to four types: observed positive, observed negative, reported positive, or reported negative.¹ We also use information about the structural tendencies exhibited by each network, including the extent and direction of each of the structural properties included in the p^* models, based on the nominal significance of the coefficients (θ 's) from the p^* model: positive, none, or negative. We use a cutoff value of a .05 significance level only as a heuristic to determine whether the tendency is positive or negative.

¹We also tried a four-category coding for the kind of relation: groom, agonistic, influence, and other. The conclusions from that analysis are similar to the ones reported here for the four group categorization.

6. RESULTS OF THE CORRESPONDENCE ANALYSIS

Let us turn now to the full set of 42 networks. Figure 3 presents the first two dimensions of column scores from the correspondence analysis. The first three dimensions of the correspondence analysis accounted for 24.7, 12.9, and 11.1 percent of the variance, respectively. The column scores plotted in Figure 3 pertain to the model providing the predictions. Networks that are in close proximity in this figure are similar in the extent to which their p^* parameter estimates predict other networks in the set. Looking at Figure 3, we see in the center toward the top a grouping of networks including cosponsorship in all of the U.S. senates except the Ninety-third and Ninety-fourth (labeled s95 through s101), dominance among willow tits (tits), advice among Krackhardt's high-tech managers (krack), and dominance among male hyenas (hyenam). On the far right of the figure, we see aggressive/submissive relations among juvenile vervet monkeys (vervet1a and vervet2a, vervet2m), dominance among sparrows (spar-

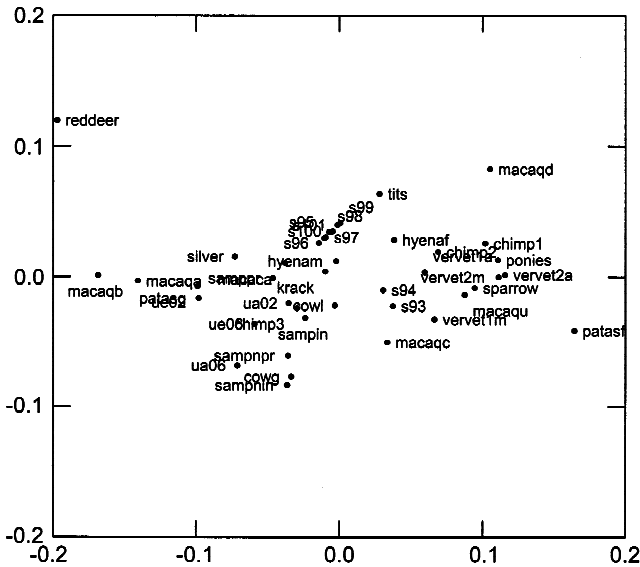


FIGURE 3. Correspondence analysis of similarities between networks from p^* model parameters, column scores.

row), threats among highland ponies (ponies), pant-grunt calls between chimpanzees (chimp1), and patas monkeys fighting (patasf).

In Figure 3 networks that are close to one another tend to exhibit similar structural properties. How can we interpret the overall spatial patterning in this figure? First, we use information about the structural features of the networks themselves, as seen in the directions and magnitudes of their p^* parameter estimates. For each network, we code it as positive, negative, or none on each of the structural features based on the direction and nominal significance of the estimated coefficient for that property, as described above. Table 1 reports these codings for each network as the p^* parameter profile. For example, we can see that the Ninety-third Senate has positive tendencies for transitive triples, out-stars, and in-stars and a negative tendency for cyclic triples. We then draw confidence ellipses around the networks with each property on the correspondence analysis configuration.² The results for mutuality, transitivity, and cycles are presented in Figures 4 through 6.

We examine the extent to which networks with specific structural tendencies occupy distinct regions of the correspondence space using an analysis of variance with the dimension scores as the dependent variables and the three category classifications of structural tendencies as factors, using the procedure described in Kumbasar, Romney, and Batchelder (1994) and Romney, Batchelder, and Brazill (1995). An analysis of variance comparing column dimension scores along the first three dimensions between three categories of structural properties gives the proportion reduction in error (PRE) in dimension scores due to the categorical grouping variables, as measured by the correlation ratio squared, η^2 . Table 5 presents these statistics for the first three dimensions of the correspondence analysis. From these results it is clear that the first dimension distinguishes networks in which mutuality is an important property from those in which it is not, or in which there is a tendency away from mutuality ($\eta^2 = 0.43$). Transitivity is an important contrast along the second dimension ($\eta^2 = 0.27$).

We use the same procedure to study whether similarities among networks are patterned by animal type (human, nonhuman primate, non-primate mammal, or bird) or by relation type (observed positive, observed negative, reported positive, reported negative). The confidence ellipses

²The confidence ellipse is centered on the means of the dimension 1 and dimension 2 coordinates. Its orientation is determined by the covariance of the two variables. We present 68.27 percent confidence ellipses.

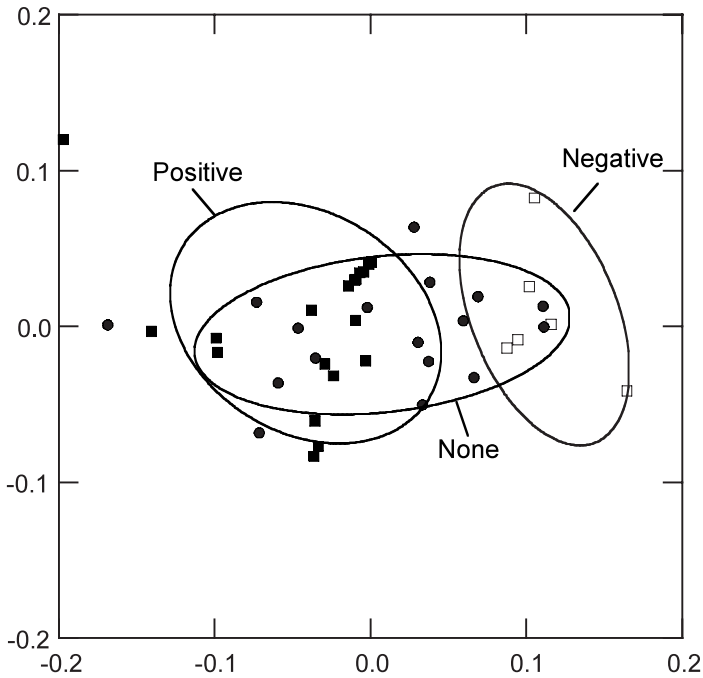


FIGURE 4. Confidence ellipses for mutuality overlaid on correspondence analysis of similarities between networks from p^* model parameters.

for animal type and for relation type, overlaid on the correspondence analysis configurations, are in Figures 7 and 8. Results in Table 5 show that the kind of animal is not an important distinction along any of the first three dimensions of the correspondence analysis. Whether the relation is observed or reported is important along both of the first two dimensions ($\eta^2 = 0.10$ and $\eta^2 = 0.20$), and whether the relation is positive or negative is an important distinction along the first and third dimensions ($\eta^2 = 0.12$ and $\eta^2 = 0.13$). Relation type, coded into four categories, is an important aspect of the second dimension ($\eta^2 = 0.26$). Overall the type of relation appears to be more important than the type of animal in distinguishing among the networks.

Further investigation of the associations between the relation type and properties of the networks reveals some interesting relationships for our sample of networks. Observed positive relations (for example, grooming between nonhuman primates and cosponsorship between senators) tend

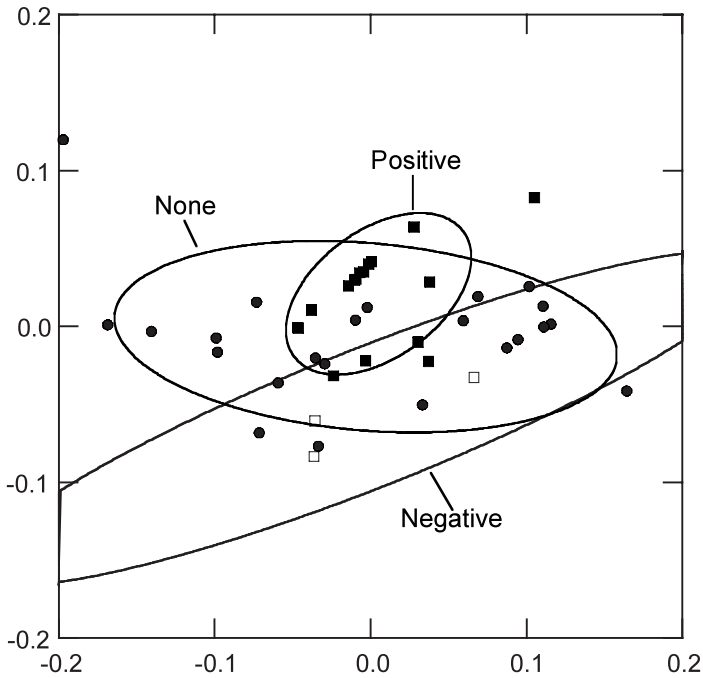


FIGURE 5. Confidence ellipses for transitivity overlaid on correspondence analysis of similarities between networks from p^* model parameters.

to be mutual, as do reported negative relations (blame and negative influence). In general, transitivity is characteristic of observed positive relations, and a tendency away from transitivity is characteristic of reported negative relations. Whether these associations hold in larger samples of networks is a topic for future research.

7. DISCUSSION

We have described a methodology for comparing networks from diverse settings including vastly different species and relational contents. This methodology allows one to assess not only what structural features are important in a given network but also how similar various networks are in terms of these properties. Important features of our approach are the calculation of an index of (dis)similarity between each pair of networks, and

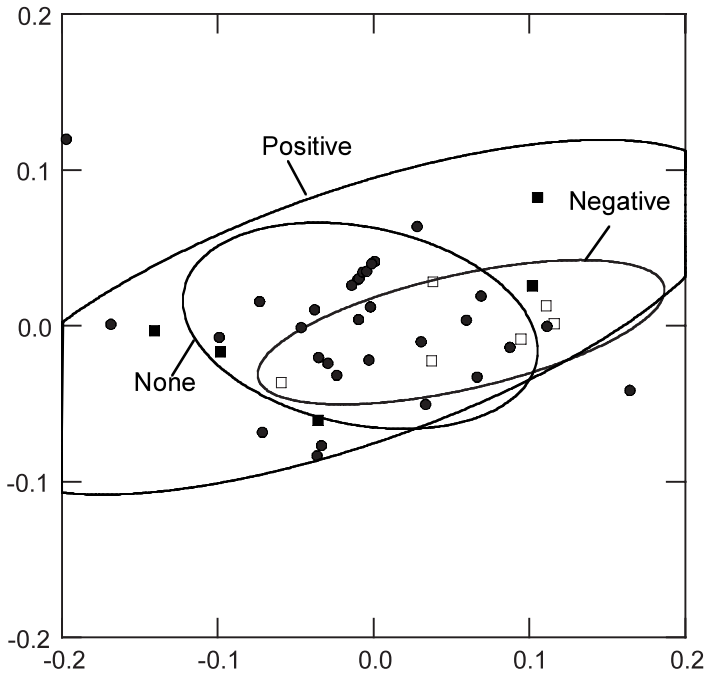


FIGURE 6. Confidence ellipses for cyclic triples overlaid on correspondence analysis of similarities between networks from p^* model parameters.

then the representation of these similarities among the diverse networks using correspondence analysis. Information about characteristics of the networks, including the kinds of actors and types of relations, is then used to interpret this spatial configuration.

In our results it appears that the kind of relation involved rather than the species underlies similarities among the networks. It is the nature of relation that determines the structural features of its network. For example, agonistic relations, whether between red deer or highland ponies, are similarly structured. This leads to the speculation that distinctions among species in network structures are due to differences in the distributions of relations in which they typically engage. This also naturally suggests that greater efforts should be devoted to measuring the typical range of relations for a species. For example, it would be useful to have observational data on different kinds of human interactions (though interviewing chimpanzees about who they go to for advice is probably out of the question).

TABLE 5
 Proportion Reduction in Error Measures (η^2) for Correspondence Analysis Dimensions by Network Structural Properties, Type of Animal, and Type of Relation

Dimension	Mutual ^a	Transitive Triples ^a	Cyclic Triples ^a	Type of Animal ^b	Observed or Reported Relation	Positive or Negative Relation	Type of Relation ^c
1	0.43**	0.00	0.09	0.09	0.10*	0.12*	0.19
2	0.01	0.27**	0.00	0.05	0.20**	0.00	0.26*
3	0.03	0.10	0.13	0.16	0.02	0.13*	0.14

* $p < .05$

** $p < .01$

^aMutual, transitive, and cycle coded: positive, none, negative.

^bHuman, nonhuman primate, non-primate mammal, bird.

^cObserved positive, observed negative, reported positive, reported negative.

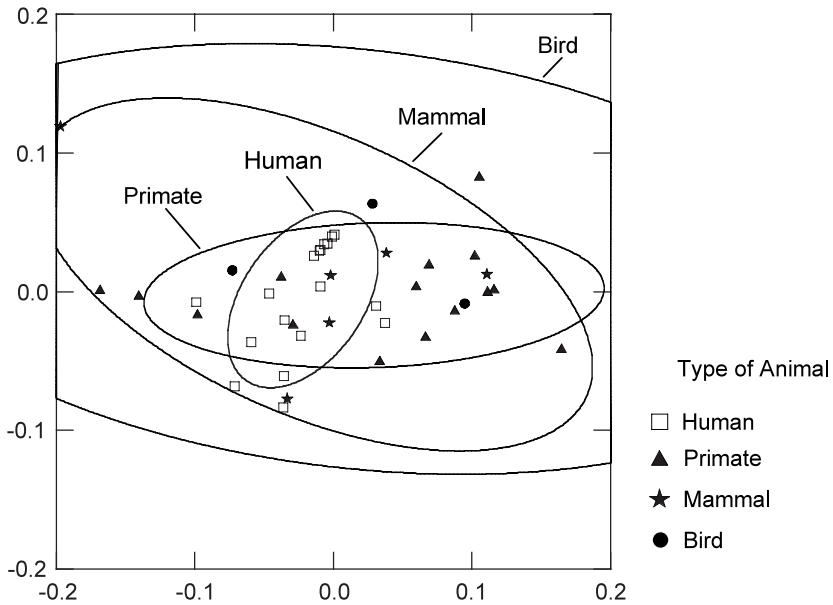


FIGURE 7. Confidence ellipses for type of animal overlaid on correspondence analysis of similarities between networks from p^* model parameters.

Our methodology of comparison consists of four steps: (1) characterizing the structural properties of each network using a statistical model, (2) comparing pairs of networks based on parameter estimates for the effects of these structural properties, (3) representing spatially the similarities among the networks, and (4) interpreting the resulting configuration using information about the networks. At each juncture there are alternative approaches that might be used. Thus it is important to consider the principles on which we base our choices and the robustness of our results in light of decisions about particular alternatives.

First, we use the p^* family of statistical models to estimate the effects of network structural properties on the probability of the graph. In the present analysis, the model includes six relatively local properties (mutuality, out 2-stars, in 2-stars, mixed stars, transitive triples, and cyclic triples). This collection of effects constitutes a Markov graph model but can easily be expanded to include other structural properties. Building models with lower-order effects before adding more complex higher-order effects is standard practice in statistical modeling, and one that we

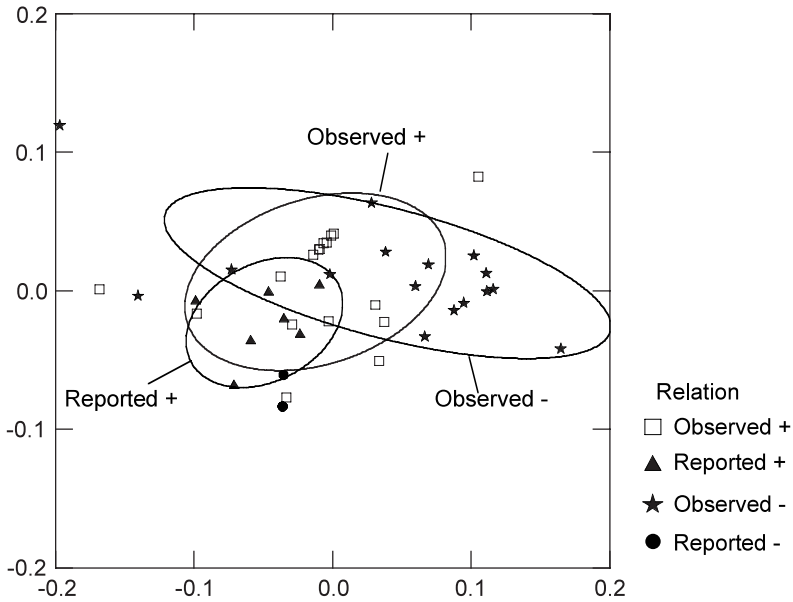


FIGURE 8. Confidence ellipses for type of relation overlaid on correspondence analysis of similarities between networks from p^* model parameters.

follow here. In addition, there are alternatives to the p^* modeling framework that also could be used to estimate effects of network structural properties—for example, Friedkin’s local density model (Friedkin 1998) could be used to estimate tie probabilities.

The second step is to compare networks based on the structural parameters in the models. We base our choice here on the principle that networks of different sizes and of different densities can have similar structures. We view size and density as differences of scale rather than as differences of theoretical significance. This leads us to use standardized regression coefficients and standardized explanatory variables for predicting tie probabilities. Comparison is then based on predicted tie probabilities, using a network’s own parameter estimates and the parameter estimates from other networks. Resemblance between networks is measured using Euclidean distance. Other measures of similarity (such as a correlation coefficient) would also be possible. We have explored other modes of comparison, using predicted probabilities from unstandardized regression coefficients, and using predicted logits rather than

predicted probabilities. In all cases the results and substantive conclusions are substantially similar to those we present here. We have only preliminarily explored another alternative—namely, direct comparison of the parameter estimates themselves. Our preliminary investigation on the current data indicates this comparison would yield the same substantive conclusions.

The third step in our methodology represents spatially the (dis)similarities among the collection of networks. Since the matrix of (dis)similarities is not symmetric we use correspondence analysis rather than other scaling options that require symmetric input data. Finally, we interpret the resulting configuration of similarities among networks by systematically examining which features of the networks are related to the spatial configuration from the correspondence analysis.

This research may be extended in several directions. First, the method can easily be used to compare multiple networks in a wide variety of situations. For example, one could compare friendship networks in multiple schools, communications relations in multiple organizations, or interorganizational transactions in multiple communities. Thus our method can be used to address fundamental questions about variability or similarity in network structure and organization. Importantly, however, our methodology is not restricted to comparing networks where the same relation has been measured in all settings. Second, in future research it will be important to explore two extensions to the models for tie probabilities or strengths. The first extension would handle valued relations. In this paper, we have, perhaps somewhat arbitrarily, dichotomized all relations. The second extension would include additional structural features in the p^* models used to characterize the networks. We have used a limited set of relatively local properties in our models. Certainly graph-level properties, such as network centralization, the diameter of the graph, or the average path length between points could also be included. Theoretically, the addition of these long-range effects may prove quite interesting if it turns out that they have different impacts in the networks of humans as opposed to the networks of other animals.

APPENDIX: LIST OF DATA SOURCES

This appendix lists the 42 networks, describes the relations, gives a reference for the source of the data, and reports the label used in Table 1 and

Figure 2. Where data are published, the table number and page of the source are given. Numbers correspond to numbers listed in Table 1.

- 1–9. *U.S. Senate*. Cosponsorship in nine senates. Records whether senator i cosponsored at least one bill introduced by senator j during that session of the Senate. Data provided by Burkett (1997). Labels: s93, s94, s95, s96, s97, s98, s99, s100, s101.
- 10. *Krackhardt's high-tech managers*. Each manager was asked who they went to for help or advice at work; Krackhardt (1987). Data available in Wasserman and Faust (1994) and in UCINET (Borgatti, Everett, and Freeman 1999). Label: krack.
- 11–14. *Sampson's monastery*. Four relations reported between monks in the monastery: positive influence (Table D15, p. 471), negative influence (Table D15, p. 471), blame (Table D16, p. 472), and praise (Table D16, p. 472); data from Sampson (1968). Data are also available in UCINET (Borgatti, Everett, and Freeman 1999). Labels: sampin, sampnin, sampnpr, samppr.
- 15–18. *Athanassiou and Nigh's top management teams (TMT)*. There are two teams (02 and 05) and two relations: from whom each manager sought advice and how extensively they had worked together; Athanassiou and Nigh (1999). Data provided by the second author. Labels: ua02, ua05 (advice), ue02, ue05 (work with).
- 19–21. *Chimpanzees*. Three relations: pant-grunt calls (Table 9.3, p. 119), initiation of dyadic agonistic confrontations (Table 9.4, p. 119), and initiation of grooming (Table 9.14a, p. 126); data from Nishida and Hosaka (1996). Labels: chimp1, chimp2, and chimp3.
- 22. *Macaca Mulatta*. One relation: grooming (Table 1, p. 274); data are from Sade (1989). Label: macaca.
- 23–26. *Macaques, macaca sylvanus*. Four relations: male carried baby away from another (Table 7a, p. 71), label macaqa; male left another with a baby (Table 7b, p. 71), label macaqb; male carrying a baby approached another male (Table 5a, p. 69), label macaqc; male approached another male who was with a baby (Table 5b, p. 69), label macaqd; data from Deag (1980).
- 27. *Stumptail Macaques (Macaca artaides)*. The relation is aggression (Table 2, p. 247); data are from Dow and de Waal (1989). Label: macaqu.
- 28–29. *Patas monkeys*. Two relations: fighting (Table III, p. 202) and grooming (Table V, p. 205); data from Kaplan and Zucker (1980). Labels: pataf and patag.

- 30–33. *Vervet monkeys (Cercopithecus aethiops sabaesus)*, juveniles from two troops (1 and 2) and two conditions (mother present and mother absent): dyadic aggressive/submissive interactions, both mothers present (Table I, p. 775), labels: *vervet1m* and *vervet2m*; dyadic aggressive/submissive interactions, both mothers absent (Table II, p. 776), labels: *vervet1a* and *vervet2a*; data from Horrocks and Hunte (1983).
- 34–35. *Cows, bos indicus*. Two relations: social licking (Figure 7, p. 130) and social grazing (Figure 4, p. 126); data from Reinhardt and Reinhardt (1981). Labels: *cowl*, *cowg*.
- 36–37. *Hyaena, crocuta crocuta*. Dominance, among females and among males. Dominance among adult females (Table I, p. 1513) and dominance among males (Table V, p. 1519); data from Frank (1986). Labels: *hyenaf*, *hyenam*.
- 38. *Highland ponies*. The relation is threats (Table 2, p. 3); data from Roberts and Browning (1998), originally in Clutton-Brock, Greenwood, and Powell (1976). Label: *ponies*.
- 39. *Red deer stags, Cervus elaphus L.* Winner and loser in encounters (Figure 1(a), p. 601); data from Appleby (1983) and also in Freeman, Freeman, and Romney (1992) and Roberts (1994). Label: *reddeer*.
- 40. *Silvereyes, zosterops lateralis*. One relation, victories in encounters (Table 1, p. 94); data from Kikkawa (1980). Label: *silver*.
- 41. *Sparrows, zonotrichia querula*. One relation: dominance, both attacks and avoidances (Figure 2, p. 19); data from Watt (1986). Label: *sparrow*.
- 42. *Willow tits, parus montanus*. One relation: dominance (Table 1, p. 1492); data from Tufto, Solberg, and Ringgsby (1998). Data originally from Lahti, Koivula, and Orell (1994). Label: *tits*.

REFERENCES

- Appleby, Michael. 1983. "The Probability of Linearity in Hierarchies." *Animal Behaviour* 31:600–608.
- Anderson, Carolyn, J., Stanley Wasserman, and Bradley Crouch. 1999. "A p* Primer: Logit Models for Social Networks." *Social Networks* 21:37–66.
- Athanassiou, Nicholas, and Douglas Nigh. 1999. "The Impact of U.S. Company Internationalization on Top Management Team Advice Networks: A Tacit Knowledge Perspective." *Strategic Management Journal* 20:83–92.
- Baker, Frank, and Lawrence Hubert. 1981. "The Analysis of Social Interaction Data." *Sociological Methods and Research* 9:339–61.
- Bearman, P. S., J. Jones, and J. R. Udry. 1997. "National Longitudinal Study of Adolescent Health: Research Design." Carolina Population Center. Unpublished manuscript.

- Bernard, H. Russell, and Peter D. Killworth. 1977. "Informant Accuracy in Social Network Data II." *Human Communications Research* 4:3–18.
- Bernard, H. Russell, Peter D. Killworth, and Lee D. Sailer. 1980. "Informant Accuracy in Social Network Data IV: A Comparison of Clique-Level Structure in Behavioral and Cognitive Network Data." *Social Networks* 2:191–218.
- Borgatti, Stephen, Martin Everett, and Linton Freeman. 1999. UCINET 5.0 for Windows. Analytic Technologies.
- Breiger, R. L., and P. Pattison. 1978. "The Joint Role Structure of Two Communities' Elites." *Sociological Methods and Research* 7:213–26.
- Burkett, Tracy. 1997. "Cosponsorship in the United States Senate: A Network Analysis of Senate Communication and Leadership, 1973–1990." Ph.D. dissertation. University of South Carolina.
- Carroll, J. Douglas, Ece Kumbasar, and A. Kimball Romney. 1997. "An Equivalence Relation Between Correspondence Analysis and Classical Multidimensional Scaling for the Recovery of Euclidean Distances." *British Journal of Mathematical and Statistical Psychology* 50:81–92.
- Clutton-Brock, T. H., P. J. Greenwood, and R. P. Powell. 1976. "Ranks and Relationships in Highland Ponies and Highland Cows." *Z. Tierpsychol* 41:202–16.
- Crouch, Bradley, and Stanley Wasserman. 1998. "A Practical Guide to Fitting p^* Social Network Models." *Connections* 21:87–101.
- Davis, James A. 1979. "The Davis/Holland/Leinhardt Studies: An Overview." Pages 51–62 in *Perspectives on Social Network Research*, edited by Paul W. Holland and Samuel Leinhardt. New York: Academic Press.
- Deag, John M. 1980. "Interactions Between Males and Unweaned Barbary Macaques: Testing the Agonistic Buffering Hypothesis." *Behaviour* 75:54–81.
- Dow, Malcolm M., and Frans B. M. de Waal. 1989. "Assignment Methods for the Analysis of Network Subgroup Interactions." *Social Networks* 11:237–55.
- Frank, Ove, and David Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832–42.
- Frank, Laurence G. 1986. "Social Organization of the Spotted Hyaena *Crocuta Crocuta*. II: Dominance and Reproduction." *Animal Behaviour* 34:1510–27.
- Freeman, Linton. 1992. "The Sociological Concept of 'Group': An Empirical Test of Two Models." *American Journal of Sociology* 98:152–66.
- Freeman, Linton C., Sue C. Freeman, and A. Kimball Romney. 1992. "The Implications of Social Structure for Dominance Hierarchies in Red Deer." *Animal Behaviour* 44:239–45.
- Friedkin, Noah. 1998. *A Structural Theory of Social Influence*. Cambridge, UK: Cambridge University Press.
- Gifi, Albert. 1990. *Nonlinear Multivariate Analysis*. New York: Wiley.
- Greenacre, Michael. 1984. *Theory and Applications of Correspondence Analysis*. New York: Academic Press.
- Hallinan, Maureen T. 1974. "A Structural Model of Sentiment Relations." *American Journal of Sociology* 80:364–78.
- Holland, Paul W., and Samuel Leinhardt. 1978. "An Omnibus Test for Social Structure Using Triads." *Sociological Methods and Research* 7:227–56.
- Horrocks, Julia, and Wayne Hunte. 1983. "Maternal Rank and Offspring Rank in Vervet

- Monkeys: An Appraisal of the Mechanisms of Rank Acquisition." *Animal Behaviour* 31:772–82.
- Hubert, Lawrence, and Frank Baker. 1978. "Evaluating the Conformity of Sociometric Measurements." *Psychometrika* 43:31–41.
- Johnson, Jeffrey C., James S. Boster, and Lawrence Palinkas. n.d. "The Evolution of Networks in Extreme and Isolated Environments." Unpublished manuscript.
- Kaplan, J. R., and E. Zucker. 1980. "Social Organization in a Group of Free-ranging Patas Monkeys." *Folia Primatologica* 34:196–213.
- Katz, Leo, and James H. Powell. 1953. "A Proposed Index of the Conformity of One Sociometric Measurement to Another." *Psychometrika* 18:249–56.
- Kikkawa, Jiro. 1980. "Weight Change in Relation to Social Hierarchy in Captive Flocks of Silvereyes, *Zosterops Lateralis*." *Behaviour* 74:92–100.
- Krackhardt, David. 1987. "Cognitive Social Structures." *Social Networks* 9:104–34.
- Kumbasar, Ece, A. Kimball Romney, and William H. Batchelder. 1994. "Systematic Biases in Social Perception." *American Journal of Sociology* 100:477–505.
- Lahti, K., K. Koivula, and M. Orell. 1994. "Is the Social Hierarchy Always Linear in Tits." *Journal of Avian Biology* 25:347–48.
- Laumann, E. O., and F. Pappi. 1976. *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: Academic Press.
- Martin, John. 1999. "A General Permutation-Based QAP Analysis Approach for Dyadic Data from Multiple Groups." *Connections* 22:50–60.
- Maryanski, A. P. 1987. "African Ape Social Structure: Is There Strength in Weak Ties?" *Social Networks* 9:191–215.
- Nishida, Toshisada, and Kasuhiko Hosaka. 1996. "Coalition Strategies Among Male Chimpanzees of the Mahale Mountains, Tanzania." Pp. 114–134 in *Great Ape Societies*, edited by William C. McGrew, Linda F. Marchant, and Toshisada Nishida. New York: Cambridge University Press.
- Nishisato, Shizuhiko. 1994. *Elements of Dual Scaling: An Introduction to Practical Data Analysis*. Hillsdale, NJ: Lawrence Erlbaum.
- Pattison, Philippa, and Stanley Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks: II. Multivariate Relations." *British Journal of Mathematical and Statistical Psychology* 52:169–93.
- Reinhardt, Viktor, and Annie Reinhardt. 1981. "Cohesive Relationships in a Cattle Herd (*Bos Indicus*)." *Behaviour* 76:121–51.
- Rindfuss, Ronald R., Aree Jampaklay, Barbara Entwisle, Yothin Sawangdee, Katherine Faust, and Pramote Prasartkul. 2000. "The Collection and Analysis of Social Network Data in Nang Rong, Thailand." Presented at the IUSSP Conference on Partnership Networks, February, Chiang Mai, Thailand.
- Roberts, John M., Jr. 1994. "Fit of Some Models to Red Deer Dominance Data." *Journal of Quantitative Anthropology* 4:249–58.
- Roberts, John M., Jr., and Bridget A. Browning. 1998. "Proximity and Threats in Highland Ponies." *Social Networks* 20:227–38.
- Robins, Garry, Philippa Pattison, and Stanley Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks. III. Valued Relations." *Psychometrika* 64:371–94.
- Romney, A. Kimball, William H. Batchelder, and Timothy Brazill. 1995. "Scaling

- Semantic Domains." Pp. 267–94 in *Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow Indow's 70th Birthday*, edited by Duncan Luce et al. Hillsdale, NJ: Lawrence Erlbaum.
- Sade, Donald Stone. 1989. "Sociometrics of *Macaca Mulatta* III: N-path Centrality in Grooming Networks." *Social Networks* 11:273–92.
- Sade, Donald Stone, and Malcolm Dow. 1994. "Primate Social Networks." Pp. 152–66 in *Advances in Social Network Analysis*, edited by Stanley Wasserman and Joseph Galaskiewicz. Thousand Oaks: Sage.
- Sampson, S. 1968. "A Novitiate in a Period of Change: An Experimental and Case Study of Social Relationships." Ph.D. dissertation. Cornell University.
- Shrader, Charles B., James R. Lincoln, and Alan N. Hoffman. 1989. "The Network Structures of Organizations: Effects of Task Contingencies and Distributional Form." *Human Relations* 42:43–66.
- Snijders, T. A. B. 1996. "Stochastic Actor-Oriented Dynamic Network Analysis." *Journal of Mathematical Sociology* 21:149–72.
- Snijders, T. A. B., and M. A. J. Van Duijn. 1997. "Simulation for Statistical Inference in Dynamic Network Models." Pp. 493–512 in *Simulating Social Phenomena*, edited by R. Conte, R. Hegselmann, and P. Terna. Berlin: Springer.
- Strauss, David, and Michael Ikeda. 1990. "Pseudolikelihood Estimation for Social Networks." *Journal of the American Statistical Association* 85:204–12.
- Tufto, Jarle, Erling Johan Solberg, and Thor-Harald Ringsby. 1998. "Statistical Models of Transitive and Intransitive Dominance Structures." *Animal Behaviour* 55:1489–98.
- Wasserman, Stanley. 1987. "Conformity of Two Sociometric Relations." *Psychometrika* 52:3–18.
- Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, England: Cambridge University Press.
- Wasserman, Stanley, and Dawn Iacobucci. 1988. "Sequential Social Network Data." *Psychometrika* 53:261–82.
- Wasserman, Stanley, and Philippa Pattison. 1996. "Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and p*." *Psychometrika* 61:401–25.
- Watt, Doris, J. 1986. "Relationship of Plumage Variability, Size, and Sex to Social Dominance." *Animal Behaviour* 34:16–27.
- Weller, Susan C., and A. Kimball Romney. 1990. *Metric Scaling: Correspondence Analysis*. Newbury Park, CA: Sage.