Computer Science and Artificial Intelligence Laboratory

Technical Report

MIT-CSAIL-TR-2010-033 July 27, 2010

Characteristics of Small Social Networks

Whitman Richards and Owen Macindoe

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Characteristics of Small Social Networks

Whitman Richards, Owen Macindoe
CSAIL, Massachusetts Institute of Technology, 32-364, Cambridge, MA. 02139
{wrichards, owenm}@mit.edu

Abstract

Two dozen networks are analyzed using three parameters that attempt to capture important properties of social networks: leadership ***L***, member bonding ***B***, and diversity of expertise ***D***. The first two of these parameters have antecedents, the third is new. A key part of the analysis is to examine networks at multiple scales by dissecting the entire network into its *n* subgraphs of a given radius of two edge steps about each of the n nodes. This scale-based analysis reveals constraints on what we have dubbed "cognitive" networks, as contrasted with biological or physical networks. Specifically, "cognitive" networks appear to maximize bonding and diversity over a range of leadership dominance. Asymptotic relations between the bonding and diversity measures are also found when small, nearly complete subgraphs are aggregated to form larger networks. This aggregation probably underlies changes in a regularity among the ***LBD*** parameters; this regularity is a U-shaped function of networks size, *n*, which is minimal for networks around 80 or so nodes.

1.0 Overview

One of the ways science advances is to discover non-trivial relations among measurements of an entity, which then give insights into the nature of that entity. Classic examples include: (a) the PV/T relation between pressure, P, temperature, T, and the volume, V, of a gas, (b) the Strahler number, Ni/(Ni+1) which characterizes the branching structure of rivers or trees, or (c) supply-demand relations. In the study of social networks, these kinds of discoveries of parametric relations are rare (but see Barabasi & Albert 2002, Watts & Strogatz 1998, Newman, 2003). Using a multi-scale analysis, we report three new properties that are characteristic of one important class of social networks.

Although our specific measurement parameters might be questioned, they were chosen to have relevance to the building of "cognitive" social networks, as contrasted with networks such as the spread of infectious diseases or telecommunications infrastructure. Two of the three parameters chosen are closely allied with previous measures. These choices, however, are not the key issue. Rather, can these parameters lead to the discovery of new properties characteristic of a class of social networks? Will these discoveries of new relationships lead to insights about the structure of social networks?

We proceed first by presenting the ***LBD*** parameters, using the evolution of a half-dozen small "start-up" groups of 10 to 30 individuals to gain the basic insights (Richards & Wormald, 2009). Then about two dozen small networks (50 – 500 nodes) are examined. Part of this latter analysis entails breaking down the network into smaller components. These components resemble the small groups. The next step is to aggregate these components to form a larger network. This analysis shows that there are two distinct processes involved in the network formation, with each process having quite different relations among parameters. Finally, we make the conjecture that as networks are increased further in size to 1000 or 10^6 nodes, a third, stochastic process comes into play. Hence, network analysis must be conducted at multiple scales.

2.0 The Representation

Let G_n be an unlabeled graph with *n* vertices. Each vertex v_i of G_n corresponds to a group member (i) and each undirected edge e(i,j) indicates a symmetrical relationship between two group members. Each individual will have at least one relation to another. We assume that G_n will be a connected graph. This characterization is very simple, and can easily be augmented using directed edges, for example. If G_n is not stationary, but is evolving, we so indicate by G_n^+ .

Unfortunately, even with our simple characterization, the number of different graphical forms explodes rapidly as the number of vertices increases. A group of only 8 individuals has over 10,000 different possible graphs for reciprocal relationships; for 12 individuals, there are over 100 billion; for 16 members the number explodes to $O[10^{23}]$. Hence detailed representations such as pictures of graphs, matrices, or edge lists quickly become impractical for network classification. Consequently, a common strategy is to focus on a few key parameters that capture regularities or which define classes of graphical forms. Over the past decade or so, popular choices have included degree distributions, edge probabilities, characteristic path lengths, clique number, diameter, chromatic numbers, or spectral coefficients – plus others more esoteric (Read & Wilson, 1998, Newman, 2003). In the area of social networks, such choices have led to distinctions such as random graphs (Bollabas, 2001), scale-free or multiscale graphs (Barabasi, 2002; Kasturirangan, 1999), small world graphs (Watts & Strogatz, 1998), peer-to-peer graphs (Bourassa & Holt, 2003) and p* graphs

(Anderson et al. , 1999.) Almost all of these parameterizations are applied to characterize large-scale graphs (>>1000 nodes) and are of limited value for small group studies (< 100 nodes.) One notable exception are motifs that appear as induced subgraphs in large networks (Milo et al, 2002; Stoica & Prieur, 2009; Wolfram, 2002) or the studies of subgraph cascades (Leskovec et al, 2007; Watts, 2002). Our proposal follows these leads, identifying three types of subgraphs that capture important characteristics of small group and network formation.

The main proposal is that the evolution of a network or group G_n^+ -> G^+_{n+1} entails the interplay of leadership, team building, and diversity in expertise. These attributes are particularly relevant in start-ups. First, the start-up needs the initial visionary leadership, plus a few other founding members to create the seed. But at some point, additional expertise is needed, such as legal aid, venture capital financing, or special talents for product development, etc. (Page 2007). All together, bonding among these members creates a successful start-up that functions smoothly as a highly motivated team. In sum, the team has leadership, there is a range of different talents, and critically, close alignments are present among team members. Each of these three factors can be associated with different types of subgraphs, which in turn can be used to parameterize the group structure. Accordingly, we define the following three parameters:

2.1 Leadership ***L***: This parameter is a measure of the extent to which the edge connectivity in a network is dominated by one (or more) vertices Hence the parameter ***L*** is based on vertex degree, d_i, with vertice(s) having the highest degree d_{max} taken as the leader(s). Consider the "star" subgraph S_k , with one dominant vertex and k-1 vertices all of degree one. If $G_n = S_k$, the leadership index for S_k is defined to be "1". Following Freeman (1978), the leadership score for any graph is then given by:

$$L = \sum_{i=1}^{n} (d_{\max} - d_i) / ((n-2)(n-1)) \qquad 1.$$

where d_i is the degree of vertex v_i . This relation sums the difference in the degree of a vertex with respect to the maximum degree in G_n, and normalizes this sum by the maximum possible sum, $\sum_{i=1}^{n} ((n-1)-1) = (n-2)(n-1)$ which is derived from the case when $G_n = S_n$.

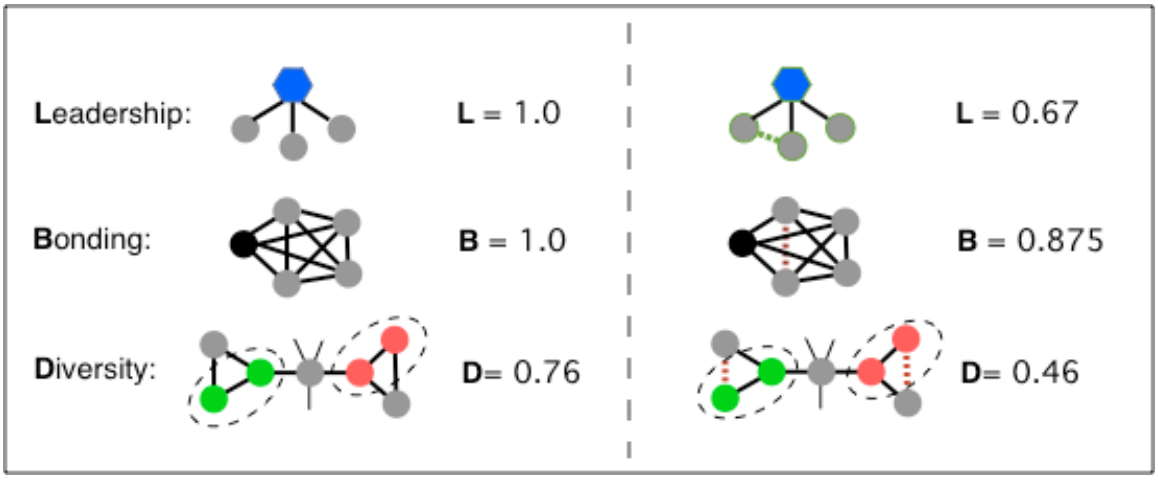


Fig 1: Subgraphs that capture the key properties of groups and their evolution. Left: Graphs with maximal or near-maximal values of the ***L, B, D*** parameters. Right: Graphs revised with an edge addition (green) or deletion (red), with new ***L, B, D*** values.

An obvious weakness of the leadership index is that for graphs with directed edges, such as hierarchical trees T_n, the dominant vertex may not be the vertex with the maximum degree. Such situations are common in military or business organizations. In these cases, the leadership measure could be revised to count the total of the in-degree for each vertex, accumulated over the subtrees.

2.2 Bonding ***B***: As team building increases, group members find similar interests. These alignments are new bonds, which are represented as new edges between vertices in G_n^+. With increasing connectivity, if vertex v_k is joined to both vertices v_i and v_j, then in social networks the likelihood increases that v_i and v_j are also joined. (In other words, if the friend(s) of your friend(s) is also your friend(s), then you belong to a tightly bonded clique.) A popular measure for this clustering is the number of triangles about each vertex, normalized by the maximum achievable by a graph with the same number of (directed) paths of length 2:

$$B = 6 * (\#\, triangles) \,/\, (\#\, paths\_length\_two) \qquad 2.$$

with the factor "6" needed to correct for the number of (directed) paths associated with any triangle (Newman, 2003.) Note that if G_n is the fully connected graph K_n, then bonding ***B*** is maximal with value "1", whereas for the "star" graph S_n or for any tree T_n, the bonding will be zero. Hence when

L is one, ***B*** will be zero, and vice versa. This interdependence among these two parameters, and also the third parameter ***D*** described below, suggests that a useful dimensionality reduction in the representation is possible, as will be seen shortly.

Other definitions of bonding could be used. For example, the number of triangles could be calculated (and normalized) locally about each vertex and summed over all vertices. (This is the clustering coefficient used by Watts & Strogatz, 1998) A table by Newman (2003) compares the values of three versions of "clustering coefficients", showing high correlations over most of their ranges. For our study, the definition (2) typically used in describing social networks appears to be the most sensitive to the different types of observed small group evolution effects.

2.3 Diversity (or heterogeneity) ***D***: The "bow-tie" illustrates our diversity measure (Fig 1.) The minimum unit is the dipole K_2 consisting of two connected members forming a partnership or "mini-team". Diversity emerges when two such dipoles are separated by a minimum of two edge steps, which we call disjoint dipoles. Diversity increases with an increasing number of such pairs of disjoint dipoles. Note that this measure is related to Granovetter's (1973) weak ties, and is high when centrality indices are high (Newman, 2005). To obtain the diversity measure, the number of pairs of disjoint dipoles in G_n with $n \geq 4$ is normalized using the count of the number of induced squares in the complete bipartite graph $K_{F[n/2],\ C[n/2]}$ as follows:

$$\boldsymbol{D} = Sqrt[(\#\ disjoint\_dipoles) / (\frac{1}{2} * \frac{n}{2}(\frac{n}{2} - 1))^2] \qquad 3.$$

It is clear that each pair of independent dipoles in a graph corresponds precisely to an induced square (i.e. a closed path of length 4 with no diagonals) in the complement of the graph. This number is maximized, over all graphs on *n* vertices when *n* is even by the complete bipartite graph (Schelp & Thomason, 1998). (Note that when *n* is odd, the Floor(F) and Ceiling(C) of *n*/2 apply.) The square root in the definition is used to bring the measure into a more appropriate range for comparison between graphs that are of the density we will be considering, but does not change the maximum possible value of ***D***, namely "1". (In the results to follow, we computed the denominator in expression (3) rounded to the nearest even

integer. This procedure was a convenient approximation for cases when *n* was odd.)

Again, like the ***L*** and ***B*** parameterizations, other measures related to our diversity measure have been proposed. For example, Caldarelli et al (2004) also count 4-cycles, but in the graph G_n, not its complement. More relevant, are centrality measures. When centrality is high for sparse graphs, so will diversity be high. Although the motivation may differ, the intent is to unveil hidden, higher-order properties of complex networks, including aspects of "weak, distal ties".

2.4 The ***L, B, D*** simplex: The interdependence of the ***L, B, D*** parameterizations have already been noted (see also Fig 1.) Without excessive loss of information, we can project the ***L, B, D*** values onto their <1,1,1> plane as follows:

$$\begin{aligned} l &= L/(L + B + D) \\ b &= B/(L + B + D) \\ d &= D/(L + B + D). \end{aligned} \qquad 4.$$

Figure 2 illustrates. Here the simplex has been divided into nine parts, with the interior triangle roughly corresponding to some common types of graphs. The interior triangles near each vertex correspond to regions of dominance of one parameter. For example, if $l >> b, d$, then the region abuts the $l = 1$ point and includes variations on star-like subgraphs $\mathbf{S_n}$. Similarly, near $b = 1$, we find the complete graphs $\mathbf{K}_n$, and near $d = 1$ are rings $\mathbf{R_n}$ (i.e. graph cycles) or "umbels" $\mathbf{U_n}$. The latter are extreme cases of sparse graphs where clusters of small complete graphs are linked through one central vertex.

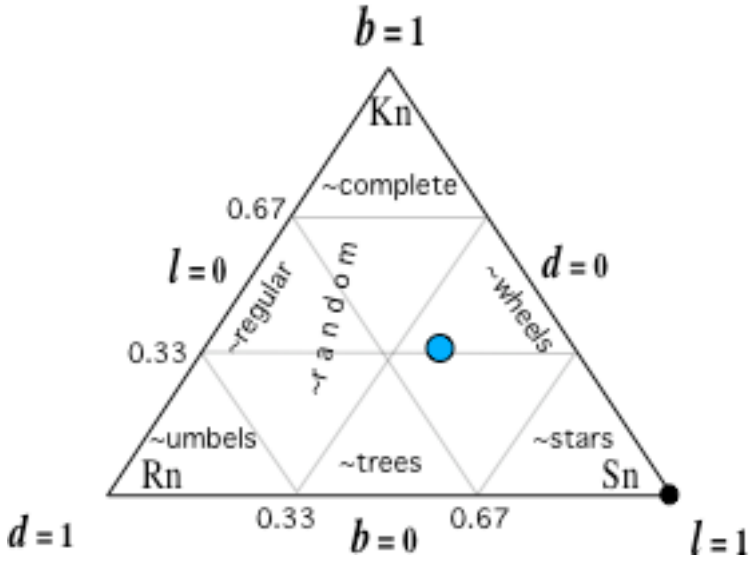


Fig 2: Regions of some familiar graphs are indicated on the projection of ***L, B, D*** onto the 1, 1, 1 plane. Kn = complete graph; Sn = star graph; Rn = ring. The blue circle indicates the terminal equilibrium location of evolving small groups.

Also shown on the plot by a blue circle is the approximate equilibrium location for small groups. Note that this location is on the leadership side of the ***d, l*** bisector through ***b*** = 1, roughly on the partition ***l*** > ***b,d***. and well to the right of the region of small Erdös-Rényi random graphs. (The random graph area illustrated is for 20 vertex graphs of varying probabilities; as *n* increases, the region moves toward ***l*** = 0.)

3.0 Small Group Representation (and Evolution)

Small groups are defined here as a collection of less than 100 individuals or agents. As mentioned, they are typically formed by one or two individuals, who then enlist other colleagues for support and expertise. Using the ***LBD*** representation, the observed evolution of a typical small group is shown in Fig 3. These are averaged results for six small groups, described in more detail elsewhere (Richards & Wormald, 2009).

The left panel of Fig 3 shows the ***L, B , D*** values for several stages of the group development indicated at the top of the figure. (The numbers at the bottom show the group size.) Although these curves represent a simplification over the actual graph structures, the dynamics is obviously still complex

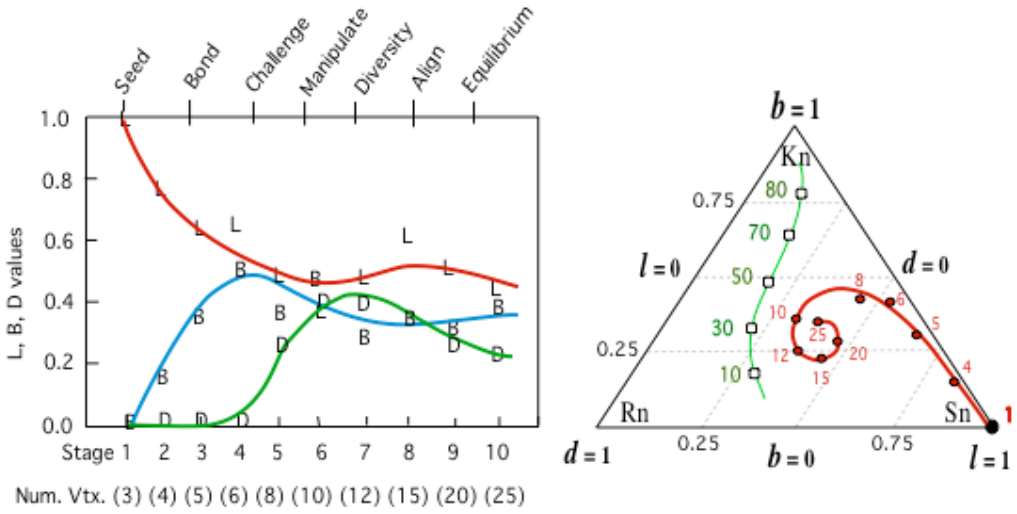


Fig 3. An example of how ***L, B, D*** and ***l, b, d*** vary in the evolution of a small group. In the left panel the size of the group is shown on the bottom line. The red spiral on the right panel shows typical small group evolution, averaged over six small groups. The green curve shows the positions of 20 vertex random graphs with edge probabilities as indicated.

(Dorogovtsev & Mendes, 2003). As a further simplification, we show these same data replotted on the ***l, b, d*** simplex in the right panel of Fig 3. Now we see a regularity in the evolution in the form of a counter-clockwise red spiral. Elsewhere, we have shown that this regularity cannot be modeled using a probabilistic evolution for recruiting and bonding of new members. Such probabilistic models yield a smooth *clockwise* arc from ***l*** = 1 upward toward ***b*** = 1, joining the green curve, which indicates loci for Erdös-Rényi random graphs.

An important point emerges from the small group study: the evolution is complex, but seems to move from ***l*** = 1 in a counter-clockwise spiral toward an equilibrium point where ***b*** ~= ***d*** < ***l***. This "final" location is along the ***b*** = ***d*** divider of the simplex. This result is a characteristic that also appears in subgraph components of larger networks.

4.0 Small Networks

We consider networks having roughly 50 to 500 nodes to be small. Table 1 lists the collection analyzed here, along with some of their parameters. (See Appendix 1 for more details.) Fig 4 presents a encapsulated summary showing the range of network positions in the ***lbd*** simplex. In the table, as well as in the

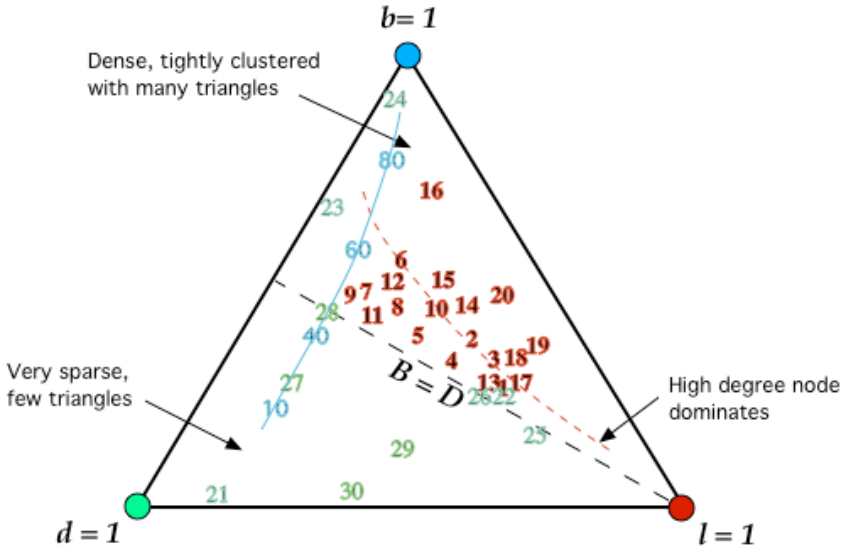


Figure 4. Illustration of the variability of ***LBD*** indices for different Social networks. For clarity, the raw ***LBD*** values have been projected onto the 111 plane (ie normalized by their sum.) The red numbers refer to "cognitve" networks; the green numbers are pseudo-trees or artificial constructions. The blue line indicates the range of Erdös-Rényi random graphs with edge probabilities as indicsated (n= 100.)

simplex, we have distinguished between the "cognitive" or more interactive social networks (red numbers in Fig. 4) and those which might be considered "passive" or constructed (green numbers) – such as infectious diseases or more clearly, networks of trees without graph cycles etc (or with very few three-cycles). Networks generated by preferential attachment with one link only between a new and old node would be in this latter class.

Note first that all our "social" networks have ***lbd*** values that fall in the upper half of the simplex, above the line ***b*** = ***d***. Below this dividing line are various forms of "pseudo"-trees, with varying degrees of complete graphs as components. (Strict trees will be located on the ***b*** = 0 edge of the simplex.) By construction, we can create networks that occupy arbitrary positions below the ***b*** = ***d*** divider – for example by adding "buds" of complete graphs to trees, thus moving the network position for a strict tree upward from the ***b*** = 0 boundary. However in the class of "interactive" social networks we have studied, this lower half of the simplex below ***b*** = ***d*** appears unoccupied. This implies that for social networks, ***B >= D***, as was seen for the small group equilibrium point.

5.0 Constraints on L, B, D relations:

Social networks aside, obviously our ***L, B, D*** parameters are not independent. For example, ***L*** is maximal for the "star" like tree that has no triangles (3 cycles.) Hence ***B***=0. But as edges are added to the star graph, triangles are formed, and

Num	Graph	Vertices/Nodes	*v*	*e*	Edges	Ref.
1	Seed+	Start-up	6	7	Collaboration	34
2	SmallGroup	Start-up (avg)	25	42	Collaboration	34
3	LosAlamos	Scientists	30	78	Collaboration	31
4	Karate	Club Members	34	78	Friendship	42
5	Enron	Employee email	37	50	Email exchange	9
6	LesMiserable	Characters*	40	105	SceneCoappearance	18
7	HIV	Core Group	40	56	Friendship	32
8	Bright	Words	54	175	Free associations	31
9	Dolphins	Dolphins	62	159	Time in proximity	22
10	Enron	Employee email	79	147	Email exchange	9
11	PolVotes	Senate 2009	99	356	Same Votes	19
12	PolBooks	Books	105	441	Purchased together	20
13	Adj-Noun	Adjectives & Nouns	112	425	Co-occurence in DavidCopperfield	27
14	SantaFe	Scientists	116	174	Collaboration	13
15	Enron	Enron Employees	143	623	Email exchange	9
16	JJATT	Terrorists	263	998	Known associate	3
17	C.Elegans	Neurons	297	2148	Neural connection	40
18	Linux2001	Kernel mailing list members	302	749	Email exchange	14
19	Linux2008	ditto	447	2122	ditto	14
20	PolBlogs	Poltical Blogs	1490	16715	Blog Hyperlinks	1
		“Non-Cognitive”				
21	BinaryTree	Binary tree	127	126		
22	InfxDisease	HIV spread	250	266	transmission	32
23	Football	College football tournament	115	613	Match played	13
24	SmallWorld	Ring seed	1000	10^3		
25	Barabasi_2	Multi-scale (2 attach)	500	982	PrefAttach	5
26	Barabasi_5	Multi-scale (5Attach)	500	2422	PrefAttach	5
27	Erdös-Rényi	Random graph (eP=0.1)	100	524		
28	Erdös-Rényi	Random graph (eP=0.4)	100	2083		
29	BuddedTree	Tree + Triangle buds	109	110		
30	RandomTree		100	99		

Table 1. Networks analyzed. See Appendix 1 for more details. (* Two co-appearances were required to eliminate random co-occurrences in scene.)

the ***lbd*** position will move upward toward ***B***=1 to the location of complete graphs with all vertices of the same degree and ***L*** = 0. Similarly, if edges are removed from the complete graph (where ***B***=1) such that all vertices continue to have equal degree, then the revised graphs will lie along the ***L*** = 0 boundary and eventually the graph will morph to a ring located at ***D*** = 1, with both ***B*** and ***L*** = 0. What we seek, then, are relations among the ***LBD*** parameters that are not a trivial consequence of the innate dependencies among the parameters.

Specifically, the role of the ***b=d*** divider in the Simplex should be a property of the networks, not a trivial consequence of innate dependencies among the parameters. Similarly, the mean positions of the set of "social" networks in Fig 4 seem to be somewhat regular. To a crude approximation (red dashed line) this locus is $b = 0.75(1 - l)^{(4/3)}$. This result is clearly a potential network property, and will be discussed in a later section. To understand this approximation, however, we need to proceed first by breaking up the network into small subgraphs, noting the position of these subgraphs in the ***lbd*** Simplex, and then showing the effects of linking subgraphs into a larger network.

5.1. Fine structure of a network

Consider the small graph in Fig 5 centered about the circled node. This graph exemplifies a subgraph of a much larger network. One node and its edges have been highlighted in green. Note that all vertices lie at most one-edge step from the center node. Hence our subgraph has radius 1. Rather than computing ***LBD*** values for the full network, we can gather additional information by the distribution of ***lbd*** values for all *n* subgraphs. Although these data can be gathered for subgraphs of various radii up to the diameter of the full graph, we

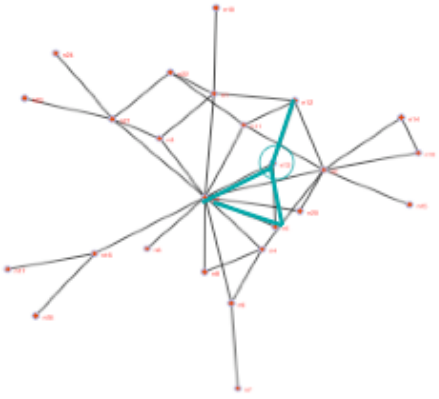


Fig 5. Example of an induced subgraph of radius 1 about a (circled) node.

limit the subgraphs to those with radius 2. (This choice is the smallest needed to yield good estimates of the parameter ***D***.) For each full graph, we now proceed to plot the ***lbd*** values of its *n* subgraphs, obtaining a scatter plot on the Simplex. These locations indicate the fine structure of the network for the chosen radius.

Using a radius of 2 for thesubgraphs, Fig 6 (right) shows this fine-structure scatter for two quite different social networks, Linux-01 and JJAT. (More examples are shown in Appendix 3.) Not surprisingly, the scatter is large. However, two important observations consistent over other social networks are illustrated: (1) the full graph ***lbd*** values typically lie above the radius 2 scatter (* in Fig 4), and (2) the lower boundary of the scatter tends to hug the ***b*** = ***d***

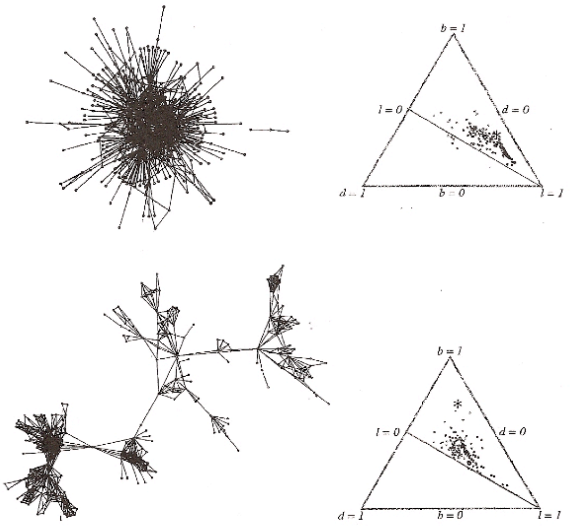


Fig 6. Radius 2 subgraph scatter plots for Linux01 (top) and JJAT Networks (bottom). See Table 1 for network descriptions. Full graph ***lbd*** value is indicated by *. Note that almost all points lie above the ***b = d*** divider.

divider. These observations demonstrate that small subgraphs (or small groups) have a structure that differs from that which characterizes the entire network and which is more typical of small groups. (Note: ***LBD*** calculated for a subgraph will still find the maximum degree node, hence the maximal degree nodes may appear in more than one subgraph.) Thus we have one process characteristic of small groups, and a second that presumably arises from how the small groups are linked to form very large networks.

Claim 1. If a network maximizes both ***B*** and ***D***, then the ***lbd*** scatter for its subgraphs of radius >=2 will lie in the upper half of the Simplex, with the lower

boundary of the scatter often hugging the ***b*** = ***d*** divider (with only a few exceptions.)

Proof Sketch: The approach is to specify the graphical form when ***D*** is maximal, showing that ***B*** must also be maximal for that class of graph. As the maximal graphical form is perturbed, disjoint dipoles will be reduced (decreasing ***D***), and the number of paths of length 2 will increase (decreasing ***B***.) Because ***D*** decreases slightly faster than ***B***, the ratio ***B/D*** is seen to increase.

Preliminaries: Consider the graph G_n and its complement $!\underline{\underline{G}}_n$. Then the number of disjoint dipoles in G_n is the number of 4-cycles (induced squares) in $!\underline{\underline{G}}_n$. (This can be seen easily by noting that the complement of an induced square is a pair of disjoint dipoles.) The G_n with the maximum number of 4-cycles is $!\underline{\underline{G}}_n = K_{n/2,n/2}$, namely the complete bipartite graph. (see proof in Schelp & Thompson, 1998.) Hence the maximum number of disjoint dipoles for any G_n is $({}_{(n/2)}C_2)^2 = ((1/2)(n/2)(n/2-1))^2$, where the Floor and Ceiling is used for odd *n*. Note that this is the denominator of the normalized ***D***.

Consider next the complement of the complete bipartite graph $!\underline{\underline{K}}_{n/2,\ n/2}$. This layout has the maximum number of disjoint dipoles. For any $!\underline{\underline{K}}_{n/2,\ n/2}$, both parts of this (unconnected) graph are complete graphs (see Fig 5.) Hence not only is ***D***=1 but ***B***=1 also, hence ***b*** = ***d***.

Case 1: We require our networks to be connected. Joining the two components of $!\underline{\underline{K}}_{n/2,\ n/2}$ with a single edge *j* will create a connected graph of diameter 3, and will reduce both ***B*** and ***D***. Specifically, a single link will increase the paths of length 2 by 2(n/2 –1) and eliminate $(n/2-1)^2$ disjoint dipoles (see Fig 5.) More generally, for *j* links, the paths of length 2 will increase by $2j(n/2-1)$ and the disjoint dipoles will be reduced by $\sum_{k=1}^{n/2}(\frac{n}{2}-j)^2)$. For this first case with $n > 5$,

new values for ***D*** and ***B*** will be:

$$D = Sqrt[({}_{n/2}C_2{}^2 - j(\tfrac{n}{2}-1)^2 + (j(j-1)/2))/{}_{n/2}C_2{}^2] \qquad 5a.$$

alternately,

$$D^2 = 1 - 16j/n^2 + 32j(j-1)/(n^2(\tfrac{n}{2}-1)^2) \qquad 5b.$$

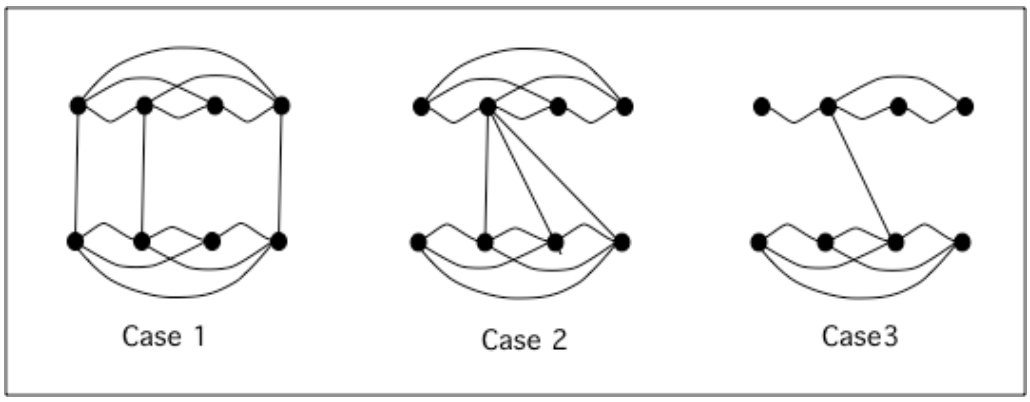


Fig 5. Revisions of complement of complete bipartite graph

similarly,

$$B = 6_{n/2}C_3 / (6_{n/2}C_3 + 2j(\frac{n}{2} - 1)) \qquad 6a.$$

or,

$$1/B^2 = (1 + 2j/(\frac{n}{2}(\frac{n}{2} - 2)))^2 \qquad 6b.$$

Neglecting the fourth-order terms in n, the $(\boldsymbol{D/B})^2$ ratio becomes

$$(D/B)^2 = (1 + 16j/(n(n-2))(1 - 16j/n^2) \qquad 7a.$$

$$\simeq (1 - 16j/n^2)^2 \qquad 7b.$$

which goes to 1 for all $j << n$ as n increases. but is less than 1 otherwise (See Table 2 for small j, n exceptions.) Hence generally, $\boldsymbol{B > D}$ for case 1.

Case 2. In addition to adding links between single nodes in the halves of $!\underline{\underline{K}}_{n/2,n/2}$, several linking edges could emerge from one node and join different nodes in the other half of $!\underline{\underline{K}}_{n/2,n/2}$. Each of these cases will form a triangle, thus increasing the triangle count at the same rate as the path 2 count. ***B*** will be only marginally affected. However, the number of disjoint dipoles will continue to decrease as in case 1. Hence ***B/D*** ~> 1. (See Table 2.)

Case 3. Let one edge link the halves of $!\underline{\underline{K}}_{n/2,\ n/2}$ to create a minimal connected graph. Then proceed to delete edges. Clearly the limiting case is a tree with one edge connecting two "star-like nodes" with (n/2 – 1) vertices of degree one. Because there are no triangles, the ***lbd*** location is on the $\boldsymbol{b} = 0$ boundary.

Intermediate edge deletions will lie between this location at $\boldsymbol{b} = 0$ and the $\boldsymbol{b = d}$ divider. (See "non-cognitive" examples in Fig. 4.)

Case 4. First one vertex is taken to cover both halves of $!\underline{\underline{K}}_{n/2,n/2}$ (i.e the limiting configuration of Case 2.) Then edges are removed equally in both halves of $!\underline{\underline{K}}_{n/2,\,n/2}$. The number of edges removed is either 25% or 50% of *n*.

Table 2 shows some results. Note first that over the situations explored, the ***B/D*** ratios hover near 1, with only the two serious exceptions indicated in bold. (Note these are for small *n*.) In contrast, the ***l,b*** values indicated in brackets can change dramatically. This implies that manipulations of $!\underline{\underline{K}}_{n/2,\,n/2}$ are robust to changes in ***B/D***, but are very sensitive to the parameter ***L***, which can undergo a wide variation depending upon the size *n* and the manipulation. If we had not

Added Links				
	n	1	25% n	50% n
Case 1	8	0.92 [0.08, 0.44]	0.91 [0.06, 0.45]	1.22 [0.00, 0.55]
	16	0.98 [0.03, 0.48]	0.99 [0.02, 0.48]	1.02 [0.00, 0.51]
	32	1.00 [0.02, 0.49]	1.00 [0.01, 0.49]	1.01 [0.00, 0.50]
Case 2	8	ditto	0.96 [0.16, 0.41]	1.10 [0.28, 0.38]
	16	"	1.00 [0.13, 0.43]	1.02 [0.28, 0.39]
	32	"	1.00 [0.12, 0.44]	1.00 [0.22, 0.39]
Dropped Edges				
Case 3	8	one link	**0.79** [0.16, 0.37]	**0.69** [0.24, 0.31]
	16	one link	0.98 [0.06, 0.47]	0.98 [0.08, 0.45]
	32	one link	1.00 [0.01, 0.49]	1.00 [0.02, 0.49]
Case 4	8	covered	1.16 [0.37, 0.34]	1.50 [0.48, 0.31]
	16	covered	1.02 [0.26, 0.37]	1.03 [0.29, 0.36]
	32	covered	1.00 [0.13, 0.43]	1.00 [0.04, 0.42]

Table 2. ***B/D*** ratios and ***l,b*** values (in brackets) for $!\underline{\underline{K}}_{n/2,\,n/2}$ with links added (Case 1 & 2) or edges dropped equally from each *n*/2 complete graph (Case3). Case 4 drops edges as in Case 3, but also assigns one vertex to cover *n*.

keyed the manipulations to $!\underline{\underline{K}}_{n/2,\ n/2}$, then obviously one can expect a variety of ***B/D*** ratios rather than those hovering near 1.

To summarize the above, if social networks strive to maximize both ***B*** and ***D*** on a local basis (ie small groups or induced subgraphs of small diameter), we expect that without edge deletions, the distribution of ***lbd*** values will hug the ***B*** = ***D*** locus, with most of the exceptions lying above that locus as the subgraphs depart more dramatically from the maximal bipartite case. If edge deletions are allowed, the ***lbd*** values will fall below the ***b*** = ***d*** divider – a result not observed for social networks. Hence we have a formal basis for the subgraph locations illustrated in Fig 4.

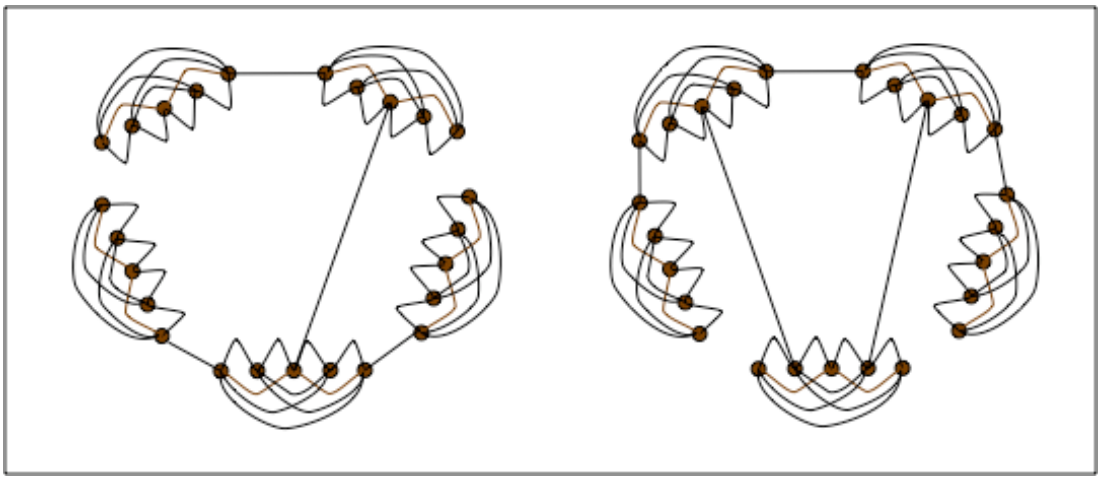


Fig 7. Two examples of loosely linked small complete subgraphs to create a larger network.

6. 0 Aggregating Small Groups.

Consider two complete graphs $K_{n/2}$, that are linked by *j* edges. Clearly as *j* increases to the maximum of $(n/2)^2$ links, the merge will result in the complete graph K_n with ***B*** = 1 and ***D*** = ***L*** = 0. Similarly, if we have *g* complete subgraphs of order $K_{n/g}$, increasing the number of links again will eventually result in ***B*** -> 1. Hence for a network consisting of linked dense subgraphs (see Fig 6), we might expect the full graph ***LBD*** values to move toward ***B*** away from the ***b*** = ***d*** divider.

Claim 2: A network formed by linking a set *g* of complete subgraphs of size (order) *k*, with the number of links *j* = *g*-1 and with at most two links per

subgraph, will result in an ***lbd*** location for the network closer to ***b*** = 1 than the mean of the pairs of linked subgraphs. (Note: the network is a chain of groups.)

Corollary: If the number of links per subgraph $j << k$, then all such trees of complete subgraphs (of order k) will satisfy Claim 1. (Note: Trees that form a "star" with one central hub, for example, are disqualified.)

Proof (asymptotic condition): Let k be the order (i.e. size) of each complete subgraph, g be the number of groups, and $n = kg$ be the order of the network. Then if $j << k << n$, we can ignore the effect of the links j if n is very large (see Proof of Claim 1.) The number of disjoint dipoles will be $({}_kC_2)^2 \; {}_gC_2$ and

$$\boldsymbol{D} = \mathrm{Sqrt}[({}_kC_2)^2 \; {}_gC_2)/ \; (\; (n/4)((n/2)-1) \;)^2 \;] \qquad 8.$$

For the bonding index, ***B***, again if $j << k << n$, then for large n we can ignore the minimal reduction in paths of length two, and ***B*** ~ 1. Hence, solving for (4), we obtain

$$\boldsymbol{B/D} = g/8^{1/2.} \qquad 9.$$

The intuition behind these results is that for weakly linked, dense subgraphs, ***B*** will be minimally affected by the links, but ***D*** will fall dramatically as g increases, due to the power of the denominator in (8).

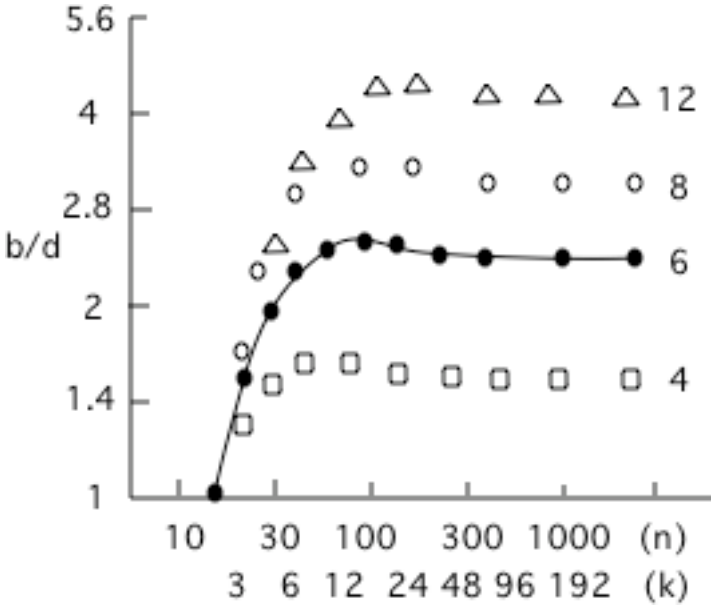


Fig 8. ***B/D*** values for a group g (4, 6, 8, 12) complete graphs of order k, linked in a chain.

Fig 8 shows empirical results for networks having groups of 4, 6, 8, and 12 complete subgraphs, linked to form a chain. The asymptotes are 1.6, 2.3, 3.0, and 4.4 respectively. These results also hold for groups of complete graphs that form trees, with the "star" disqualifier mentioned in the corollary.

In summary, linking or merging operations on dense small groups can typically be expected to increase ***B/D*** for the full network, as compared with the mean scatter of its subgraphs, thus moving the full graph ***lbd*** location upward in the simplex as illustrated in Fig 4.

7. Regularities in *L*, *B*, *D* relations

Referring again to the simplex in Fig. 4, we see a crude relation among social networks, roughly captured by the dashed red line. If indeed all these points in the simplex fell exactly on this red line, then we would infer that there was a non-planar surface in the 3D ***LBD*** space that cuts through the <1,1,1> plane on which the simplex lies. A simple example of this kind of regularity is the green line in the simplex that is the projection of the string of ***LBD*** values for a 100 node Erdös-Rényi random graph as edge probability *p* is varied. If the number of nodes were included as an additional parameter for these random graphs, then there would be a surface in ***LBD*** with the two variables, *n, p*. (See also Appendix 2.) Hence, for our "cognitive" social networks, the rough clustering of points about the red line suggests at best a surface in ***LBD*** space that must be parameterized by at least two variables.

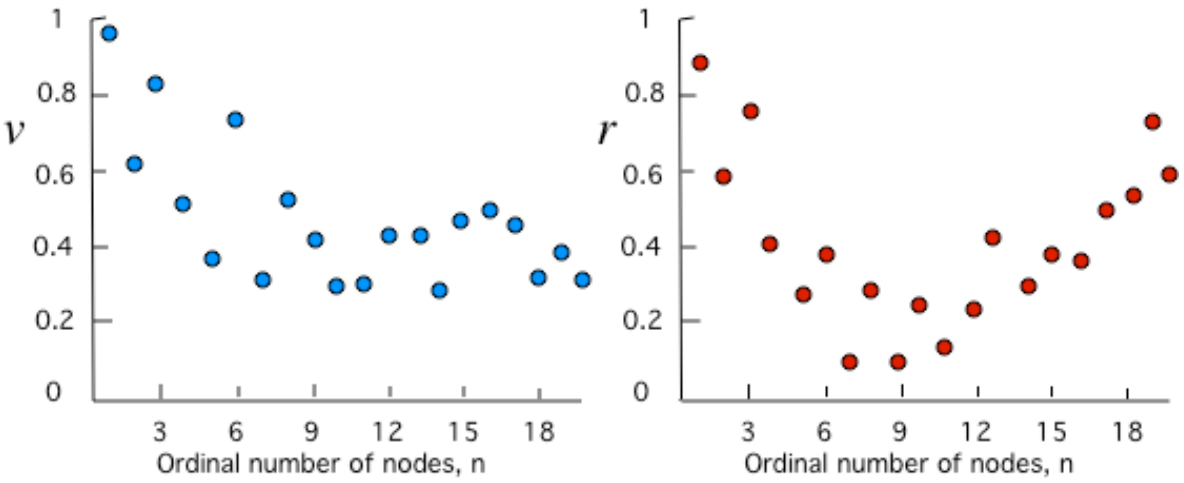


Fig.9 Left: length of **LBD** vector *v* vs. "cognitive" network size. Right: ratio *r* vs. network size. Note the abscissa is an ordinal scale. Position 11 corresponds to n = 100.

A quick check is to examine the lengths of each of the ***LBD*** vectors in Table 3 (Appendix 1). For example, if all of these length vectors are (roughly) equal, then the points for "cognitive" social networks all lie on a spherical surface about ***LBD*** = 0. The left panel of Fig 9 plots these vectors versus the ordinal position of network size (n). Note that the vector lengths decrease to about ordinal position 11, which corresponds n = 100, and then become roughly constant. Hence we infer a regularity associated with "cognitive" social networks that is dependent on n.

Given the importance of the ***B/D*** constraint discussed in sections 5 and 6, plus the fact that this ratio appears to change with ***L*** (i.e. increasing ***B*** decreases ***L***, or vice versa), a potential form for a cognitive social network regularity based on three variables could be:

$$f(L)g(B/D) = h(n) \qquad 10.$$

where now n is included to satisfy the relation seen in Fig. 9 (left). Further support for (10) is seen in the pair-wise angular relations between the ***LBD*** vectors (not shown), which suggests that the decrease in vector length for $n <$ 100 is correlated primarily with the ***L*** scores, which are high for small groups.

In the right panel of Fig 9, we plot one form of (10), namely

$$r = L(B/D)^{1/2} = h(n) \qquad 11.$$

We see that this ratio decreases up to ordinal position 11 (n = 100), and then rises, requiring a non-monotonic function for $h(n)$. This result confirms that small "cognitive" social networks appear to have a character different from larger ones, consistent with empirical studies and proposals by Dunbar (1992).

8. Discussion

Our most striking result is that "cognitive" social networks appear to share parametric properties that are different from non-social networks, such as biological, physiological or physical patterns. This observation rests on our particular choice of parameters (***L,B,D***) to characterize graphs and has not previously been documented. Figures 2 and 4 illustrate. In the first case, all social networks have ***LBD*** values that lie above the ***b=d*** divider. This result

can be modeled by assuming that social networks tend to maximize ***B*** and ***D*** in the presence of ***L***, which leads to ***B*** >= ***D*** as shown by Eqn 7b.

The second property of social networks is seen when the mean ***LBD*** values for subgraphs of radius 2 were compared with the ***LBD*** values for the full network. Then we can infer that $\boldsymbol{B}_* \boldsymbol{>= B_2}$. where $\boldsymbol{B}_*$ and $\boldsymbol{B_2}$. are respectively the values for the full and radius 2 cases. A related result is when small dense subgraphs are weakly linked to form a larger network. Then ***B/D*** increases to an asymptote for large *n*.

Finally, we show that at least two processes underlie "cognitive" social networks, depending upon their size (Fig 9.) For small networks ($n < 100$), the ratio $r = L(B / D)^{1/2}$ decreases and then subsequently rises. For contrast, Appendix 2 shows the behavior of Erdös-Rényi graphs, which have known parameterizations of *n, p*. Unlike Erdös-Rényi graphs, for "cognitive" social networks, we do not yet know how the ratio *r* will behave for very large networks with $n >> 1000$. Our expectation is that a stochastic element would need to be added for haphazard encounters, introducing a process more like that seen for Erdös-Rényi or scale free graphs. To address these issues, we propose modifying the use of the ***LBD*** parameters to emphasize their scale aspects. For example, subgraphs of radius 2 are clearly useful, but the full network ***LBD*** values seem unreasonable, both due to the large number of required "friends" and also to the potentially large number of edge steps among nodes. Social contacts rarely extend more than 4 edge steps, except by haphazard encounters. Hence limiting the maximum radius to 4 about which ***LBD*** is calculated (perhaps with an added random jump) might be a plausible strategy. If the network is homogeneous, then the radius 4 restriction should not be penalizing. This strategy would also tend to equalize the locality of the ***L, B,*** and ***D*** calculations, which are currently dominated by ***B***. The process that differentiates large networks would then be tied principally to the haphazard jumps over nodes that touch a previously unforeseen individual (either via the web, or by a chance encounter.)

Acknowledgements
Supported by AFOSR under a MURI grant A550-05-01-032. Thanks to Omprakash Gnawali for gathering the Linux data. Nick Wormald's expertise was critical to the development of the LBD representation.

References

[1] Adamic, L.A. & N.Glance (2005) The political blogosphere and the 2004 US election. *Proc. WWW-2005 Workshop on sthe Webblogging Ecosystem.*

[2] Anderson, C.J., S. Wasserman, & B Crouch (1999). A p* primer: Logit mdels for social networks. *Social Networks*, 21, 37-66.

[3[Atran, S., S. Bennett, A. Fatica, J. Magouik, D. Noricks, M. Sageman, and D. Wright. (2008) John Jay & ARTIS Transnational Terrorism (JJATT) dataset. http://doitapps.jjay.cuny.edu/jjatt/

[4] Barabasi, A-L (2002) *Linked: The New Science of Networks*. Perseus Press, NY.

[5] Barabasi,A. & R.Albert (1999) Emergence of scaling in random networks, Science 286, 509-511.

[6] Bollabas, B. (2001) Random Graphs, 2nd Ed. Cambridge Univ. Press.

[7] Bourassa, V. & F. Holt (2003) SWAN: Small-world wide area networks. Proceedings of International Conference on Advances in Infrastructure (SSGRR 2003w paper #64), L'Aquila, Italy.

[8] Caldarelli, G; Pastor-Satorras, R. and A. Vespignani (2004) Structure of cycles and local ordering in complex networks. *European Phys. Jrnl* B, 38: 183–186.

[9] Cohen, W. W. (2009) Enron email dataset. http://www.cs.cmu.edu/~enron/

[10] Dorogovtsev, S.N. & J. F. Mendes (2003) *Evolution of Networks*, Oxford Univ. Press.

[11] Dunbar, R. (1992) Neocortical size as a constraint on group size in primates. *Jrl. Human Evolution* 22, 469-493.

[12] Freeman, L. C. (1978) Centrality in social networks: conceptual clarification. *Social Networks*, 1: 215 – 239.

[13] Girvan, M. and M.E.J.Newman (2002).Community structure in social and biological networks. *Proc. Nat. Acad. Sci.* 99, 7821-7826.

[14] Gnawali, O. D. (2009) Linux kernel email communication neworks from Jan 2001 and 2008. Personal communication.

[15] Granovetter, M. (1973) The strength of weak ties. *Amer. Jrl .Sociology* 78, 1360 –1368.

[16] Harary, F. (1969) Graph Theory. *Addison-Wesley,* Reading, MA.

[17] Kasturirangan, R. (1999) Multiple scales in small world graphs. MIT-CSAIL technical reports: //publications.ai.mit.edu/ai-publications/pdf/AIM-1663.pdf

[18] Knuth, D.E. (1993) *The Stanford GraphBase: A Platform for Combinatorial Computing*. Addison-Wesley, Reading Ma.

[19] Kolar, M., L. Song, A. Ahmed, & E.P. Xing (2010) Estimating Time-varying networks. Annals Applied Statistics, 4, 94-123.

[20] Krebs. V. (2003) Books about US politics dataset (unpublished.) http://www.orgnet.com/

[21] Leskovec, J., Singh, A. and J. Kleinberg (2006) Patterns of influence in a recommendation network. Proc. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD).

[22] Lusseau, D. K. Schneider, O.J. Boisseau, P. Haase, E. Slooten, and S.M. Dawson. (2003) The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behav. Ecology and Sociobiology*. 54, 396-405.

[23] Macindoe, O. (2010) Investigating the fine grained structure of networks. MS Thesis. Dept. Elec.Engr. & Comp. Sci. http://

[24] Macindoe, O. and W. Richards (2010) Graph comparison using fine structure analysis. *Proc. IEEE Conf. on Soc. Computing*, #244.

[25] Milo et al (2002) Network Motifs. Science 298, 824-827.

[26] Newman, M. E. J. (2003) The structure and function of complex networks. *SIAM Review* 45: 167 – 256.

[27] Newman, M. E. J. (2006) Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 74

[28] Newman, M.E.J. (2005) A measure of betweennes centrality based on random walk. *Social Networks* 27, 39 – 54.

[29] Page, S.E. (2007) *The Difference: how the power of diversity creates better groups, firms, schools, and societies*. Princeton Univ. Press.

[30] Palmer, E.M. (1985) *Graphical Evolution*, Wiley & Sons.

[31] Palla, G. Derenyi, I. Farkes, I. & T. Viesek (2005) Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814 –818.

[32] Potterat, J.J, L. Phillips-Plummer, S.Q. Muth, R.B. Rothenberg, D.E.Woodhouse, T.S. Maldonado-Long, H.P.Zimmerman, & J.B. Muth (2002) Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs. *Sex Transm. Infect* 78 (suppl 1) 159 –163,

[33] Read, R.C. and R.J. Wilson (1998) *An Atlas of Graphs*. Oxford Press.

[34] Richards, W. & O. Macindoe (2010) Decmposing Social networks. *Proc. IEEE Conf. on Soc. Computing*, #205.

[35] Richards, W. & N. Wormald (2009) Representing small group evolution. *Proc. IEEE Conf. on Soc. Computing*, #232.

[36] Schelp, R. H. & A. Thomason (1998) A remark on the number of complete and empty subgraphs. Combinatorics, Probability & Computing, 7, 217-219.

[37] Stoica, A. and C. Prieur (2009) Structure of neighborhoods in a large social network. *Proc. IEEE Conf. on Soc. Computing*, #225.

[38] Watts, D. (2002) A simple model of global cascades on random networks. *PNAS* 99, 5766 –5771.

[39] Watts,D. & S. Strogatz (1998) Collective Dynamics of small world networks, *Nature* 393 , 440-442;

[40] White, J.G., E. Southgate, J.N.Thompson, and S. Brenner (1986) The structure of the nervous system of the nematode c.elegans. *Phil Trans. Roy. Soc*. 314, 1 – 340.

[41] Wolfram, S (2002) *A New Kind of Science,* pp. 508-515. Wolfram Media, Inc.

[42] Zachary, W.W. (1977) An information flow model for conflict and fission in small groups. *Jr. Anthropological Res*. 33, 452-473.

Appendix 1: Table 3. Additional Details for Networks [$r=L(B/D)^{1/2}$]

Num	Graph	Vertices/Nodes	*eP*	*Dia.*	*Cpl.*	***L***	***B***	***D***	***r***
1	Seed+	Initial Start-up	.47	2	2.3	.8	.43	.33	.91
2	SmallGroup	Start-up (avg)	.14	4	2.4	.45	.36	.21	.59
3	LosAlamos	Scientists	.18	4	2.05	.69	.37	.29	.80
4	Karate	Club Members	.14	5	2.4	.40	.26	.24	.42
5	Enron	Employee email	.075	7	3.4	.24	.24	.15	.30
6	LesMiserable	Characters*	.135	4	2.71	.26	.63	.30	.38
7	HIV	Core Group	.072	10	4.8	.11	.26	.18	.13
8	Bright	Words	.12	5	2.6	.23	.38	.26	.28
9	Dolphins	Dolphins	.08	8	3.36	.12	.31	.26	.13
10	Enron	Employee email	,048	7	3.3	.17	.23	.11	.24
11	PolVotes	Senate 2009	.072	8	4.1	.13	.24	.17	.15
12	PolBooks	Books	.08	7	3.1	.16	.35	.19	.22
13	Adj-Noun	Adjectives&Nouns	.07	5	2.5	.38	.16	.13	.42
14	SantaFe	Scientists	.026	15	6.7	.17	.22	.07	.30
15	Enron	Enron Employees	.06	8	2.9	.24	.36	.14	.38
16	JJATT	Terrorists	.03	13	5.9	.14	.49	.07	.37
17	C.Elegans	Neurons	.05	5	2.5	.40	.18	.11	.51
18	Linux2001	Kernel mailing list	.017	7	3.2	.25	.15	.03	.52
19	Linux2008	ditto	.021	6	2.8	.34	.19	.04	.74
20	PolBlogs	Poltical Blogs	.015	8	2.7	.22	.23	.03	.61
		“Non Cognitive”							
21	BinaryTree	Binary tree	.016	12	8.3	.008	0	.04	0
22	InfxDisease	HIV spread	.009	24	8.4	.076	.03	.02	.09
23	Football	College football	.094	4	2.5	.01	.04	.24	.004
24	SmallWorld	Ring seed	.02	5	3.3	.006	.54	.06	.02
25	Barabasi_2	Multi-scale	.008	7	3.9	.078	.02	.02	.08
26	Barabasi_5	Multi-scale	.02	5	2.8	.115	.05	.05	.12
27	Erdös-Rényi	Random graph	0.1	3	2.2	.06	.11	.25	.04
28	Erdös-Rényi	Random graph	0.4	2	1.6	.14	.40	.41	.14
29	BuddedTree	Triangle Buds	0.02	11	9.9	.065	.029	.05	.05
30	RandomTree		0.02	11	10.5	.05	0	.05	0

Appendix 2: Random Graph ***LBD*** regularity

Consider only large *n*, which allows the binomial degree distribution to be approximated by a Gaussian. The approximations for ***L, B, D*** are as follows, using $n = 100$ as the example.

L: Let $d_{\max}$ be the maximum degree for the Erdös-Rényi graph with edge probability *p* and order (size) *n*. The variance in the degree distribution is $np(1-p)$. For $n = 100$, take the upper limit for the tail of the zero-mean Gaussian at $t = 0.495$. This occurs for 2.7 standard deviations. Hence $d_{\max} \simeq 2.7 \bullet (100\,p(1-p))^{1/2} + 100\,p$. To sum $(d_{\max} - d_i)$ over $i = [1, n,]$ note the Gaussian is symmetrical about 100*p*. Hence

$$L = \frac{1}{n^2}\sum (d_{\max} - d_i) = \frac{2.7}{100} \bullet (100\,p(1-p))^{1/2} \qquad 12$$

$$= 0.27(p(1-p))^{1/2} \qquad 13$$

Note that the only "free" parameter is the number of standard deviation units (2.7). This parameter for determining $d_{\max}$ can be eliminated if necessary (see Palmer pg xx.)

B: The value for large random graphs will be *p* (Newman, 2003) This can easily be derived. The number of triangles is $p^3\,{}_nC_3$, which is divided by the number of two paths, namely $p^2\,{}_nC_3$.

D. For large random graphs, the diameter is two if $0 < p < 1$. (Bollabas, 2001). The only subgraph with disjoint dipoles with diameter 2 is the bow tie (Fig 1.) More directly, the probability of two pairs of edges is p^2, and the probability that the 4 vertices are otherwise not connected is $(1-p)^4$. We have ${}_nC_4$ such pairs. Hence from Eqn 3,

$$D = ((p^2(1-p)^4\,{}_nC_4)/(n^4/64))^{1/2} \qquad 14a.$$

$$= ((64n^4p^2(1-p)^4)/24n^4)^{1/2} = 2.7^{1/2}\,p(1-p)^2 \qquad 14b.$$

where again, a reminder that the 2.7 parameter arises from the setting to determine is set to determine $d_{\max}$ for 100 nodes.

Note for large n, $\boldsymbol{B}$ and $\boldsymbol{D}$ are independent of n, whereas $\boldsymbol{L}$ is not (because of the $d_{\max}$ condition. Hence

$$r_{100} = L(B/D)^{1/2} = (2.7/10)(p(1-p)^{1/2} \bullet p^{1/2} / (2.7^{1/2} p^{1/2}(1-p)).$$
$$= 2.7^{1/2}(p/(1-p))^{1/2} / 10 \qquad 15.$$

Table 3 below shows the monotonic relation between r and p for $n = 100$.

EdgProb	L	B	D	maxDeg	r	model
10	0.06	0.11	0.25	17	0.04	0.05
20	0.11	0.20	0.37	37	0.08	0.08
40	0.14	0.40	0.40	51	0.14	0.13
60	0.13	0.60	0.26	73	0.20	0.20
80	0.11	0.81	0.082	91	0.35	0.32
90	0.06	0.91	0.02	95	0.41	0.49

Appendix 3: Examples of Scale-based analysis of "cognitive" social Networks

For each simplex, the full graph ***lbd*** location is indicated by an asterisk. The colors simply indicate the node type closest to each point. The histograms show distributions of raw ***LBD*** values. See Macindoe (2010) for further details. Part of the raw network data are available at http://people.csail.mit.edu/owenm/netdata.html. The balance are in Newman's collection at http://www.personal.umich.edu/mejn.netdata/

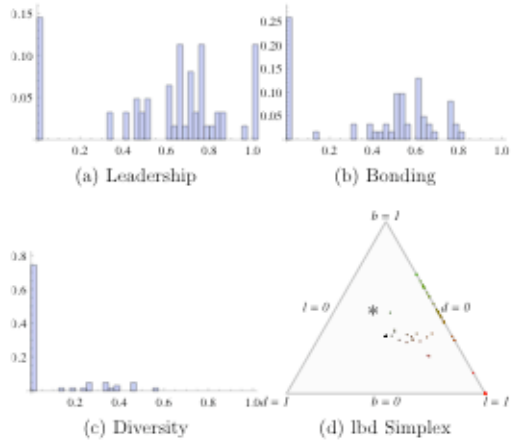


Figure A-16: LBD distributions and *lbd* simplex for the Dolphins graph at radius 1.

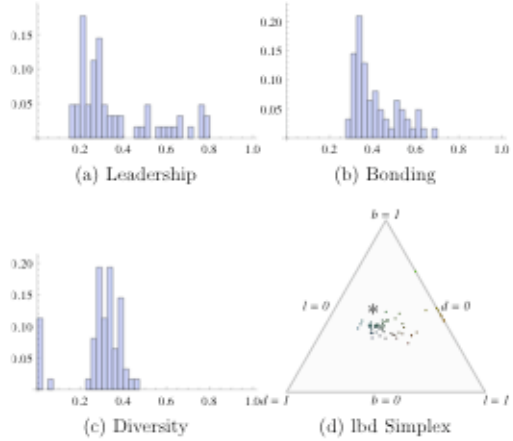


Figure A-17: LBD distributions and *lbd* simplex for the Dolphins graph at radius 2.

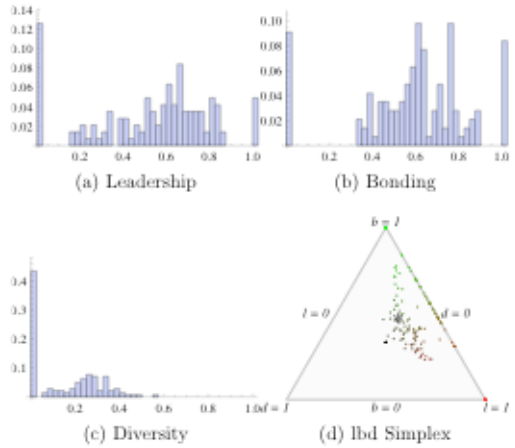


Figure A-19: *LBD* distributions and *lbd* simplex for the Enron graph at radius 1.

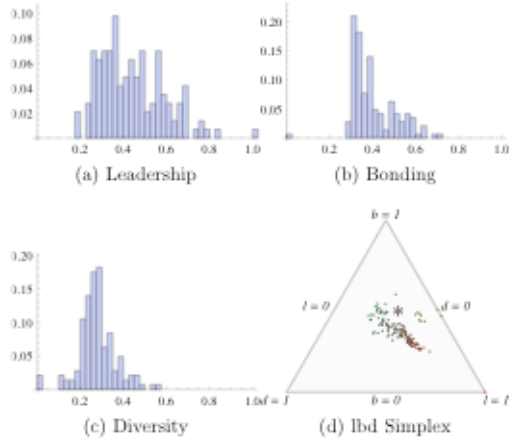


Figure A-20: *LBD* distributions and *lbd* simplex for the Enron graph at radius 2.

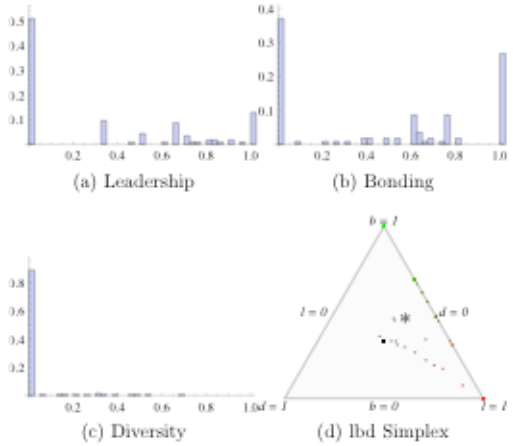


Figure A-22: *LBD* distributions and *lbd* simplex for the Santa Fe graph at radius 1.

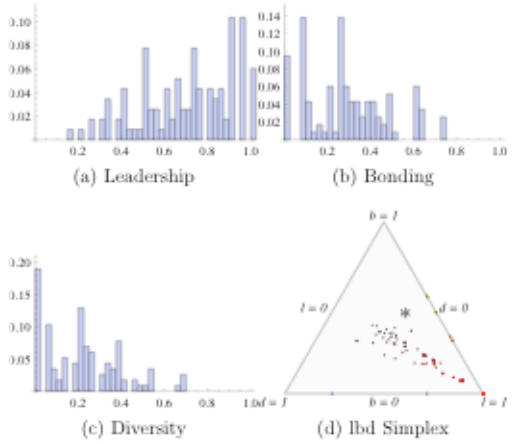


Figure A-23: *LBD* distributions and *lbd* simplex for the Santa Fe graph at radius 2.

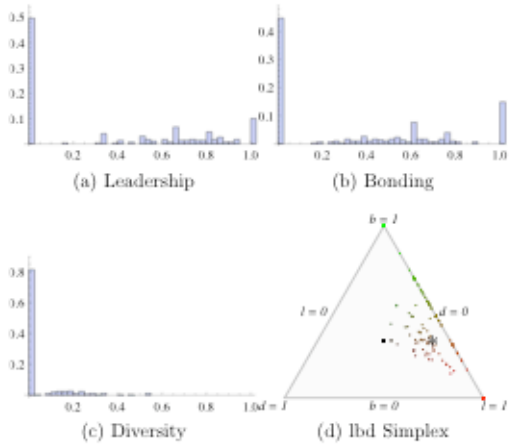


Figure A-28: *LBD* distributions and *lbd* simplex for the Linux 2001 graph at radius 1.

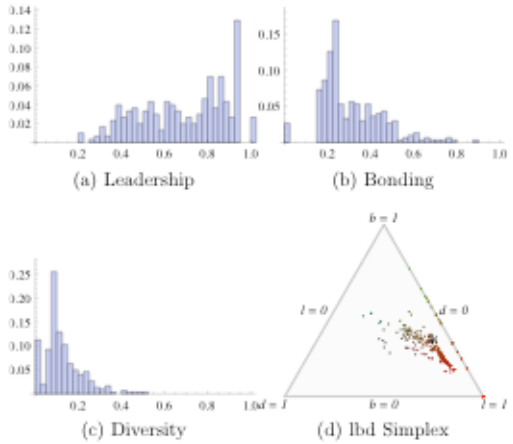


Figure A-29: *LBD* distributions and *lbd* simplex for the Linux 2001 graph at radius 2.

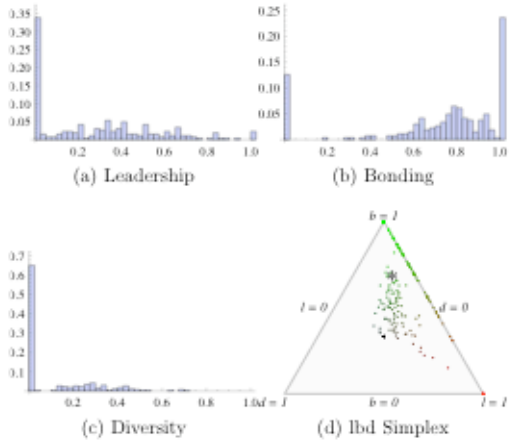


Figure A-25: *LBD* distributions and *lbd* simplex for the JJATT graph at radius 1.

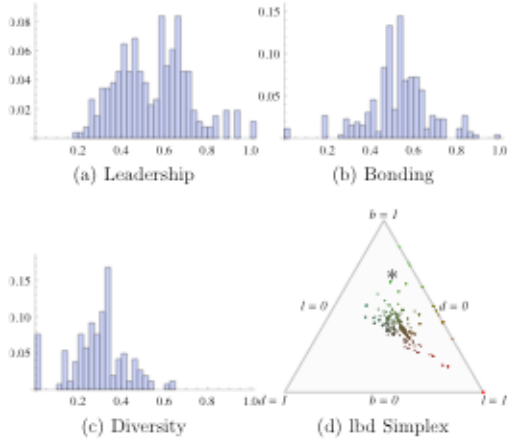


Figure A-26: *LBD* distributions and *lbd* simplex for the JJATT graph at radius 2.

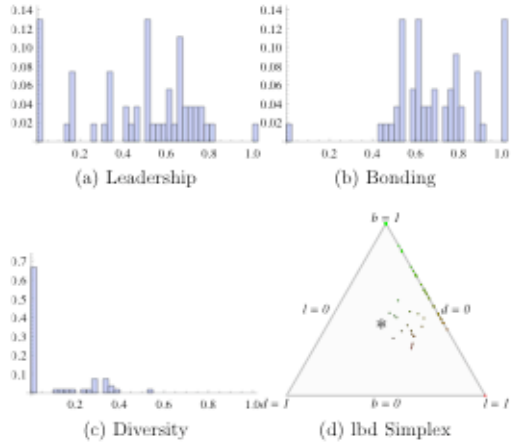


Figure A-34: *LBD* distributions and *lbd* simplex for the Bright graph at radius 1.

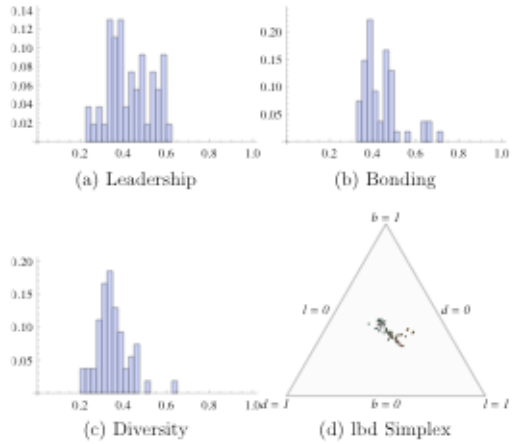


Figure A-35: *LBD* distributions and *lbd* simplex for the Bright graph at radius 2.

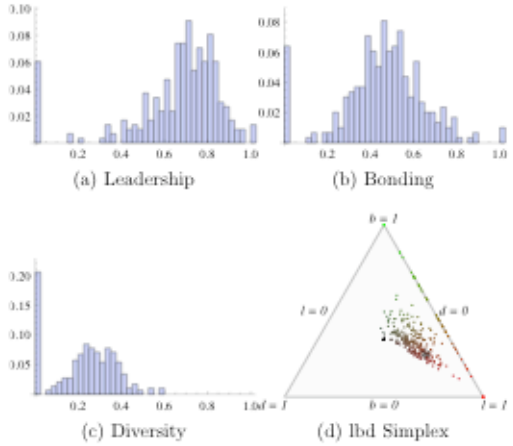


Figure A-49: *LBD* distributions and *lbd* simplex for the C-Elegans graph at radius 1.

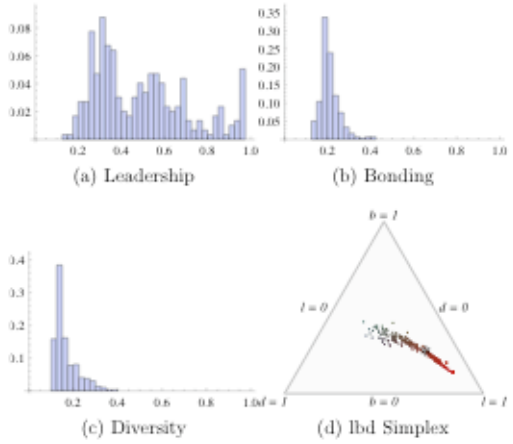


Figure A-50: *LBD* distributions and *lbd* simplex for the C-Elegans graph at radius 2.

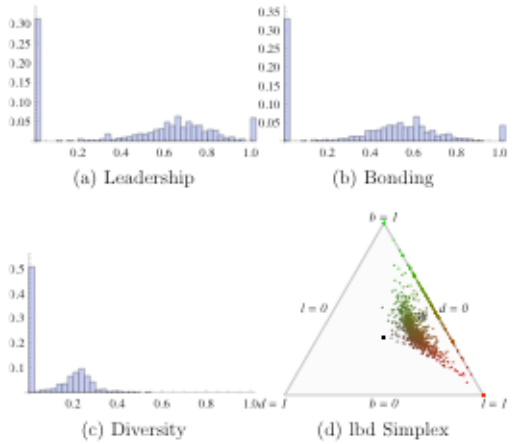


Figure A-52: *LBD* distributions and *lbd* simplex for the PolBlogs graph at radius 1.

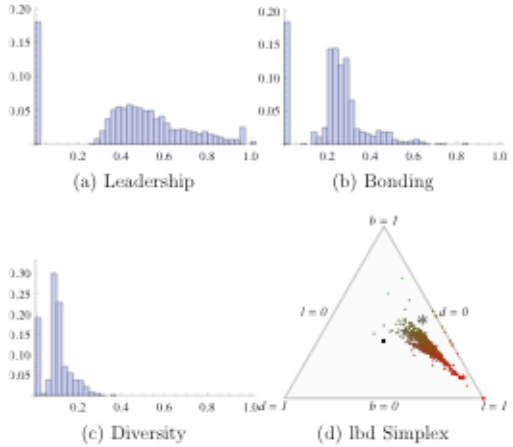


Figure A-53: *LBD* distributions and *lbd* simplex for the PolBlogs graph at radius 2.

Appendix 4: Examples of scale-based analyses of “non-cognitive” social networks.

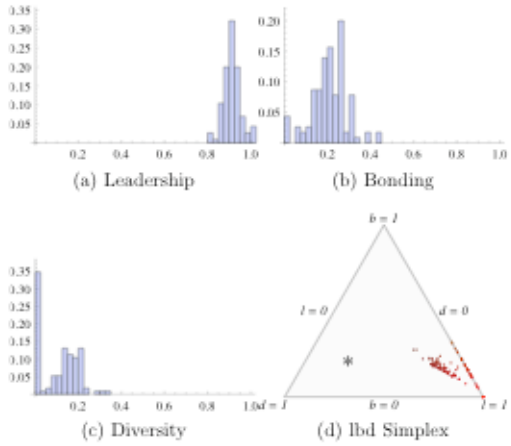


Figure A-7: *LBD* distributions and *lbd* simplex for the Erdős-Rényi graph at radius 1.

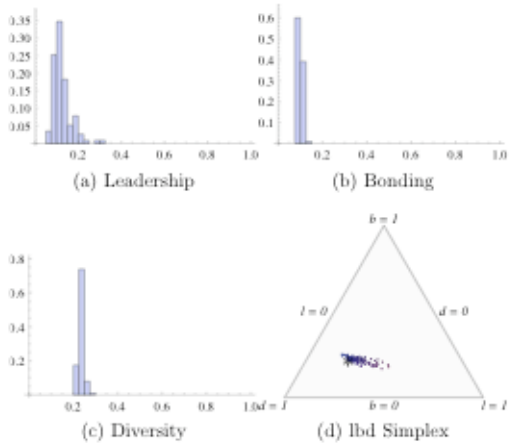
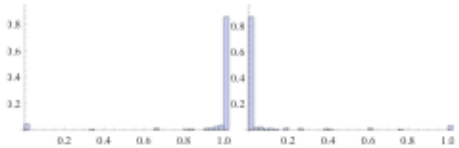
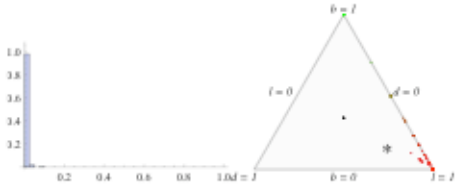


Figure A-8: *LBD* distributions and *lbd* simplex for the Erdős-Rényi graph at radius 2.

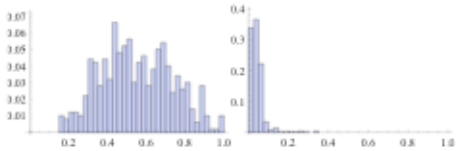


(a) Leadership (b) Bonding

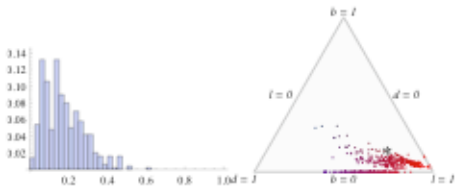


(c) Diversity (d) lbd Simplex

Figure 3: *LBD* distributions and *lbd* simplex for the Barabasi Scalefree [2] graph at radius 1.

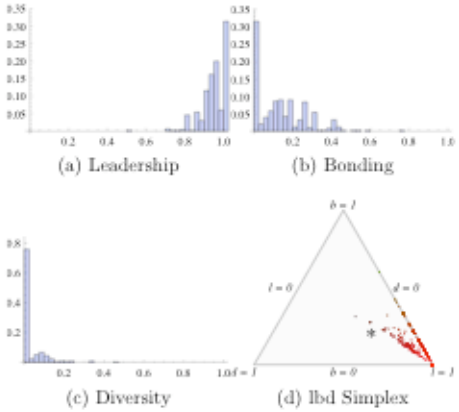


(a) Leadership (b) Bonding



(c) Diversity (d) lbd Simplex

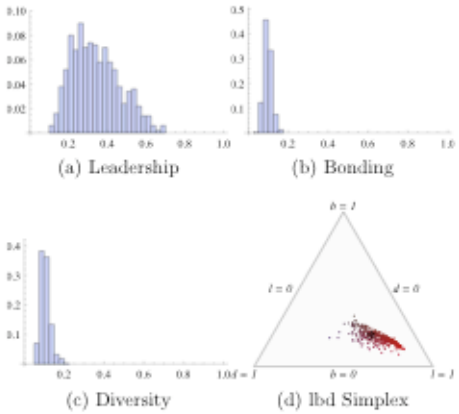
Figure 4: *LBD* distributions and *lbd* simplex for the Barabasi Scalefree [2] graph at radius 2.



(a) Leadership (b) Bonding

(c) Diversity (d) lbd Simplex

Figure 5: LBD distributions and lbd simplex for the Barabasi Scalefree [5] graph at radius 1.



(a) Leadership (b) Bonding

(c) Diversity (d) lbd Simplex

Figure 6: LBD distributions and lbd simplex for the Barabasi Scalefree [5] graph at radius 2.