Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2009-012 March 30, 2009

Representing Small Group Evolution

Whitman Richards and Nicholas Wormald

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

Representing Small Group Evolution

Whitman Richards [a,*] , Nicholas Wormald [b]

[a] CSAIL, Massachusetts Institute of Technology, 32-364, Cambridge, MA 02139

[b] Dept. of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, Canada N2L 3G1

Abstract
Understanding the dynamics of network evolution rests in part on the representation chosen to characterize the evolutionary process. We offer a simple, three-parameter representation based on subgraphs that capture three important properties of social networks: leadership, team alignment or bonding among members, and diversity of expertise. When plotted on this representation, the evolution of a typical small group such as start-ups or street gangs has a spiral trajectory, moving toward a tentative fixed point as membership increases to two dozen or so. We show that a simple probabilistic model for recruitment and bonding can not explain these observations, and suggest that strategic moves among group members may come into play.

1.0 Introduction

Small groups are defined here as a collection of less than 100 individuals or agents. They are typically formed by one or two individuals, who then enlist other colleagues for support and expertise. Examples are start-ups, non-profit initiatives, small businesses, street gangs and terrorist cells. The most obvious goal of the leadership is to foster a shared vision by expanding the group's capabilities through recruitment, at the same time increasing alignments between members to improve group effectiveness. At some juncture, the capabilities of the group may also be broadened through diversification (Page, 2007). We propose three parameters that capture these aspects of group evolution, and provide a useful representation for studying differences between small groups. Using this representation, the observed evolution of a typical small group is shown to have a spiral trajectory. This result is not consistent with recruitment and alignments (bonding) among members occurring with fixed probabilities. Rather, we suggest that small group formation likely involves more complex processes, such as members engaged in strategies to improve their influence in the group.

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 graphs for reciprocal relationships; for 12 individuals, there are over 100 billion; for 16 members the number explodes to $O[10^{23}]$. Hence pictorial representations are implausible and must be replaced by focusing on a few key parameters that capture regularities underlying classes of graphical forms. Over the past decade or so, popular characterizations have included degree distributions, edge probabilities, characteristic path lengths, clique number, diameter, chromatic numbers, spectral coefficients – there are dozens choices. (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) and peer-to-peer graphs (Bourassa & Holt, 2003.) 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; 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 formation and development.

The main proposal is that the evolution of a group $G_n^+ \rightarrow G^+_{n+1}$ entails the interplay of leadership, team building, and heterogeneity in expertise. For example, the members of a football team include those proficient in running and catching, others are effective blockers, there is the quarterback, the kicker, etc. All of these aptitudes must be highly coordinated, with not many, but only one leader calling the plays. Similarly, a start-up company needs not only the initial visionary leadership, but usually venture capital financing, special expertise for product development, etc. Again, the successful start-up functions smoothly as a highly motivated team. There is leadership, close alignments among team members, and a range of different talents. 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 of a group:

2.1 Leadership ***L***: This parameter is based on vertex degree, d_i, with the vertice(s) having the highest degree 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 (1977), the leadership index 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 $G_n = S_n$.

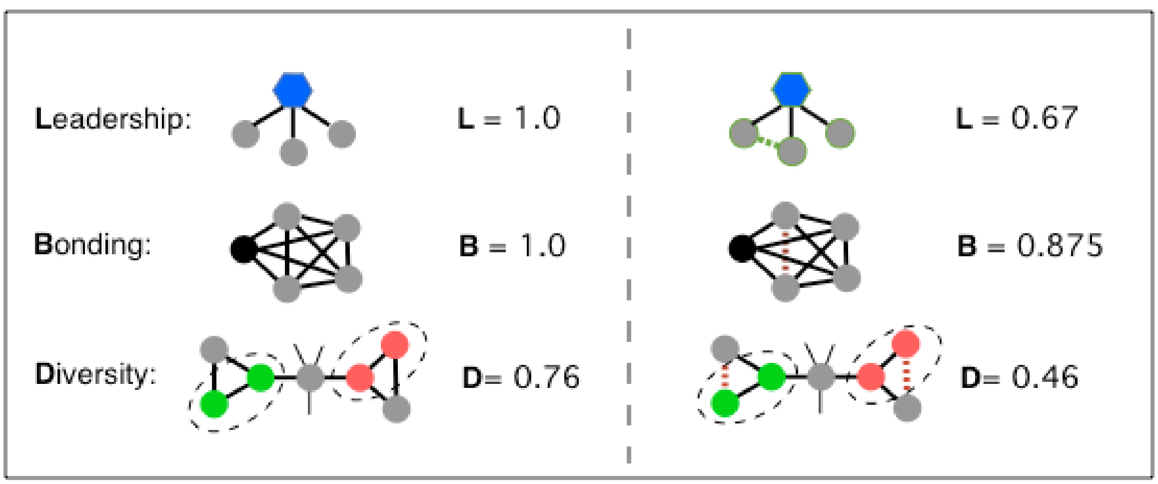

Fig 1: Subgraphs that capture the key properties of groups and their evolution. Left: Graphs with maximal or near-maximal values of the ***L, B, D*** parameters. Right: Graphs revised with an edge addition (green) or deletion (red), with new ***L, B, D*** values.

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 should be revised to count the total in-degree for each vertex.

2.2 Bonding $\boldsymbol{B}$: 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 of your friend is also your friend, then you belong to a tightly bonded clique.) A popular measure for this clustering of vertices about a vertex is the number of triangles about that 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 $\boldsymbol{B}$ 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 $\boldsymbol{L}$ is one, $\boldsymbol{B}$ will be zero, and vice versa. This interdependence among these two parameters, and also the third parameter $\boldsymbol{D}$ 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) $\boldsymbol{D}$: 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 independent dipoles. Diversity increases with an increasing number of such pairs of independent dipoles. Note that we have excluded counts based on pairs of individual vertices separated by two or more edge steps, as occurs in the star S_n graph above. These vertices of degree one will be considered as "new recruits" – a category discussed later.

To obtain the diversity measure, the number of pairs of independent dipoles in G_n with $n \geq 4$ are normalized using the count of the number of induced squares in the complete bipartite graph $K_{F[n/2],\, C[n/2]}$ as follows:

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

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, overall 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. Like our diversity measure, the intent is to unveil hidden, higher-order properties of complex networks. But the motivations underlying the two measures are quite different.

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:

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

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 $\boldsymbol{l} >> \boldsymbol{b}, \boldsymbol{d}$, then the region abuts the $\boldsymbol{l} = 1$ point and includes variations on star-like subgraphs $\mathbf{S_n}$. Similarly, near $\boldsymbol{b} = 1$, we find the complete graphs $\mathbf{K}_n$, and near $\boldsymbol{d} = 1$ 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.

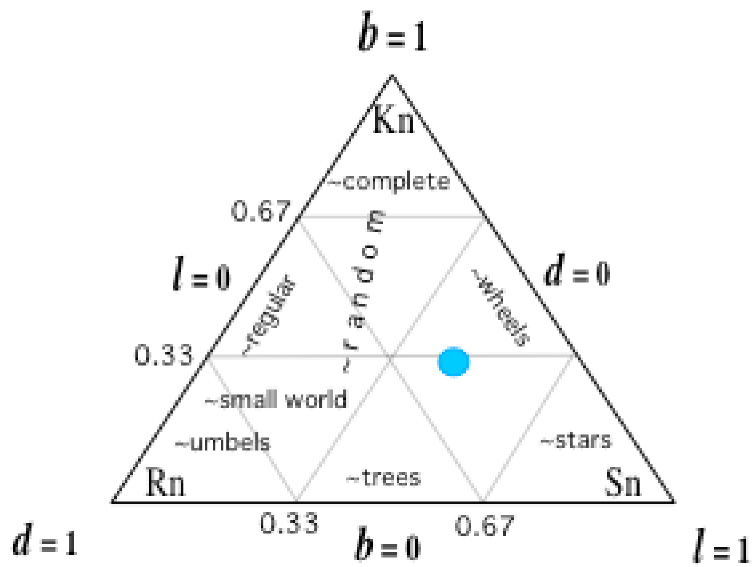

Fig 2: Regions of some familiar graphs are indicated on the projection of ***L, B, D*** onto the 1, 1, 1 plane. The blue circle indicates the terminal equilibrium location of evolving small groups.

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 $\boldsymbol{b} = 1$, roughly on the partition $\boldsymbol{l} > \boldsymbol{b},\boldsymbol{d}$. 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 $\boldsymbol{l} = 0$.)

3.0 Probabilistic Evolution

An obvious question is whether the evolution of small groups can be modeled probabilistically. The ***l, b, d*** simplex provides a convenient representation for testing this possibility. We consider one very simple probabilistic model defined as follows:

Let the connected graph G^{+}_{n} with n vertices represent the group structure, with vertex v_1 of maximal degree taken as the leader. Also let *pR* be a fixed probability for recruitment and $pB > 0$ the fixed probability for choosing new edges in G^{+}_{n}. For each iteration in the evolutionary growth add only one new vertex v_{n+1}, with v_{n+1} linked to only one old vertex, chosen uniformly with probability *pR* from the *n* nodes. (Note that for small *pR* values and small *n*, several iterations may be required to add a node.) During this same iteration, also sample the available free edges in G^{+}_{n+1} and with probability $0 < pB < 1$ add new edges to G^{+}_{n+1}. Finally, if any vertex v_k has a degree larger than that of v_1, then add a new edge $\{v_1, v_j\}$ where v_j is the newest vertex of minimal degree (typically the vertex v_{n+1}.)

Although other models for group evolution might be proposed, our choice above is meant to be a fair representation of actual practice. With the exception of member drop-outs (see discussion to follow), variations were found to make little difference in the general form of the evolutionary trajectories.

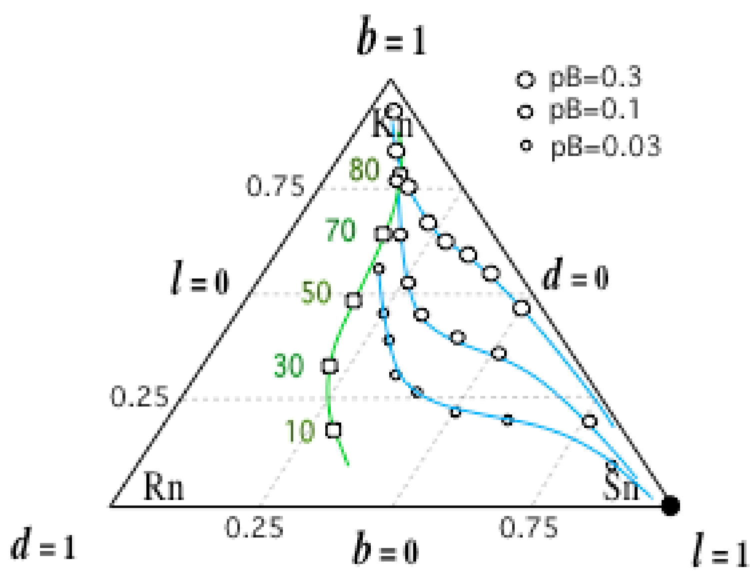

Fig. 3: Three evolutionary trajectories (blue) show different probabilities of bonding (30% largest circles; 10% middle circles; 3% smallest circles) and constant member recruitment of 30%. The green locus shows the ***l, b, d*** values for 20 vertex Erdös-Rényi random graphs, with edge probabilities as indicated.

3.1 Simple Example

Fig. 3 illustrates the behavior. For each iteration G^{+}_{n} -> G^{+}_{n+1} let pB = 10% probability of two unlinked members (vertices) bonding, and pR = 30% be the probability of adding one new member (vertex) for each iteration. The seed for our example is the star graph $\mathbf{S_3}$ with v_1 having degree 2 – the most common seed observed for group formation. A sequence of iterations sampled for n= 4, 5, 6, 8,10, 20, 40,100 is shown by the middle blue curve. Note that this middle curve, as well as the other two curves, one for pB=0.03, the other for pB=0.3, all eventually head to $\boldsymbol{b}$ = 1. Other simulations (such as seeds along the $\boldsymbol{b}$ = 0 border) show that for fixed probabilities pR and $pB > 0$, all evolutionary trajectories eventually will move toward $\boldsymbol{b}$ = 1. We formalize this intuition below.

3.2 Asymptotic Behavior

To support the inference that the probabilistic trajectories illustrated in Fig. 3 will move toward $\boldsymbol{b}$ = 1, we prove that for bonding probability $pB > 0$, and recruitment probability $pR > 0$, the fraction of free edges in the evolving graph G_n^+ will tend to zero for sufficiently large n. This implies that the location of G_n^+ can be approximated by $\boldsymbol{b}$ = 1 (i.e. a complete graph K_n) for very large n. To see why this implication holds, assume the number of free edges (i.e. the number of edges in the complement of G_n^+) is at most εn^2. There are at most n^2 squares in the complement of G_n^+ containing any given edge (n choices for each of the two other vertices in the square), so the number of squares in the complement of G_n^+ is at most εn^4. Hence $\boldsymbol{D}$, which has a normalizing factor of order n^4, must be a small multiple of ε. Similarly, the numerator in the definition of $\boldsymbol{L}$ is bounded as follows

$$\sum_{i=1}^{n}(d_{\max}-d_i) \le \sum_{i=1}^{n}(n-1-d_i)$$

which is the sum of its vertex degrees in the complement of G_n^+, i.e. twice its number of edges, which is at most $2\varepsilon n^2$. Thus $\boldsymbol{L}$ is also bounded above by a small constant times ε. Thus, as $\varepsilon \to 0$, $\boldsymbol{D}$ and $\boldsymbol{L}$ tend to 0.

Theorem 1: Let the connected graph G_n^+ grow iteratively, adding with probability pR = 1 a new vertex v_{n+1} to one of the vertices already in G_n^+, choosing that vertex from a uniform distribution. In the same iteration also add with probability $pB > 0$ new edges between unlinked vertices in G_{n+1}^+. Then as n increases, the fraction of free edges in G_{n+1}^+ is likely to approach 0 as n -> inf.

Proof: Begin with a $\mathbf{S_3}$ seed with vertex v_o joined to v_1, v_2, and a recruitment probability pR =1. Consider then any edge that can possibly be added to a vertex v_i in G_n^+ that is not added by recruitment. After *i* iterations, there will be *i* –2 such edges to lower numbered vertices. For each such edge, the probability in each successive iteration that an edge is not added is (1 – *pB*). Hence the probability that edge is not present after the *n*th iteration is $(1-pB)^{(n-i)}$. By linearity of expectation, the expected number of these edges absent after iteration *n* is thus $(i-2)\,(1-pB)^{(n-i)}$. Summing over all iterations from *i* = *3* to *n* gives the number $\overline{\beta}_n$ of non-recruited edges absent after iteration *n*. Using the well-known inequality $(1-p) < e^{-p}$, the sum $\overline{\beta}_n$ is bounded as follows:

$$Est(\overline{\beta}_n) = \sum_{i=3}^{n} (i-2) * (1-pB)^{n-i} < \sum_{i=3}^{n} (i-2) * e^{pB(n-i)} < \sum_{i=3}^{n} n * e^{-pB(n-i)} < n \sum_{j=0}^{n-3} e^{-pB*j}$$

where *j* = *n* –*i*. (1)

Note that the sum of the exponential in the last term is some constant, *c*. Recall that excluding the *n* edges recruited, there will be *n*(*n*-3)/2 possible edges that could be absent. Hence the fraction of absent edges will be at most $c*n/\mathrm{O}[n^2]$, which tends to 0 as n goes to infinity. □

Remark: For any fixed pair {*pB* > 0 and *pR* > 0}, the trajectory of small graph evolution can be closely approximated by the following mapping to a new pair {*pB*' > 0 and *pR*' = 1}:

$$pB' = 1-(1-pB)pR/(1-(1-pR)(1-pB)) \qquad (2)$$

For this more general case, we need to revise the $(1 - pB)^{n-i}$ probability used in the first sum of equation (1) to include *pR* < 1. Consider the pair of vertices in G{ v_i, v_j }. The continued absence of edge E_{ij} between two recruitments requires that no bonding occurred over the sequence of iterations between those two recruitments. Equation (3) gives the geometric series, which is simplified to (4). This expresses the probability $p(\overline{E}_{ij} / v_i)$ of not adding E_{ij} between two consecutive recruitments, expressed as the sum of the probabilities of increasing numbers of iterations passing between recruitments with no bonding occurring between v_i and v_j.

$$p(\overline{E}_{ij} / v_i) = (1-pB) \bullet pR + (1-pB)^2 (1-pR) pR + (1-pB)^3 (1-pR)^2 pR + \ldots (3)$$

$$\cong (1-pB)pR/(1-(1-pR)(1-pB)). \qquad (4)$$

Hence just before the $(n-i)^{\text{th}}$ vertex is added, the probability no edge having been formed between v_i and v_j is given by

$$p(\bar{E}_n) \cong (pR(1-pB)/(1-(1-pR)(1-pB)))^{n-i} = (1-pB')^{n-i} \qquad (5)$$

Solving for B' gives the desired approximation (2). Note that if this mapping is applied, then for any $pR > 0$ and $0 < pB$, the asymptote in the simplex for G_n^+ will approximate ***b*** = 1 in the ***l b d*** simplex.

With respect to the above theorem and remarks, the addition of a vertex that created a new dipole with probability *pD* will not change the asymptotic result. For example, the new vertex could be joined to G_n^+ with two added edges that created an induced hexagon C_6. In the limit, two such edge additions will have no effect on the evolving trajectory, because as n increases, bonding will eventually link dipole members to other vertices, reducing the diversity count, and moving the trajectory on ***l,b,d*** toward one determined by an {*pB*, *pR* = 1} pair for some *pB*.

3.3 Edge Deletions

If edges in G_n^+ can be deleted, Theorem 1 and the remarks do not apply. Such deletions could arise by imposing conditions on edge densities or vertex degrees (Watts, 2002.) An extreme case would be to include a fixed probability *pQ* for removing edges, but again keeping G_n^+ connected. The two examples in Fig. 4.0 were both created in this manner using fixed probabilities for *pR, pB, pD* , *pQ* of {0.30, 0.30, 0, 0.30} and {0.60, 0.30, 0, 0.40} respectively.

Note reversals in direction are now common, leading to chaotic-like behaviors. This haphazard behavior can continue for several hundred iterations for appropriate choices of parameters. Further studies are needed to establish fixed points and the chaotic-like regime. However, for some parameters there is a trend toward a stable edge probability whenever the size of G_n^+ begins to increase monotonically. In this case, the trajectory moves toward a fixed point in the ***l,b,d*** simplex, as suggested especially by the left panel in Fig. 4. The behavior of

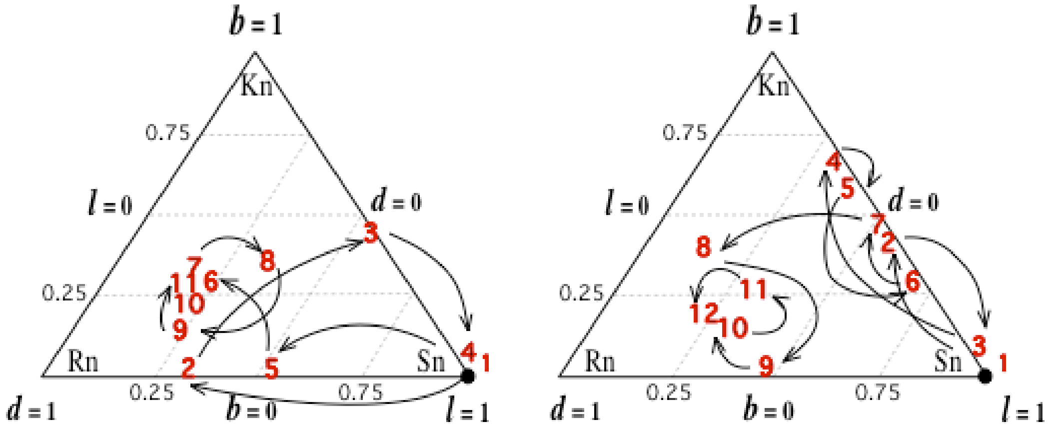

Fig 4: Loci of evolution from S_3 for two simulations, with values for *pR, pB, pQ* of 0.30, 0.30, 0.30 for the left panel, and 0.60, 0.30, 0.40 for the right panel. The behavior is very haphazard.

the trajectories in Fig. 4 suggests that a feedback condition could be imposed to control the evolution of a group toward any desired ***l,b,d*** fixed point. Such a global constraint on group evolution may be worth exploring.

4.0 Start-up Group Evolution

Figure 5 presents averaged data of the evolution of six small groups (e.g. start-ups). The left panel shows the ***L, B , D*** values for several stages of the group development; the trajectory on the ***l, b, d*** simplex is the red spiral on the right. Note the general trend is quite different from the probabilistic evolution given earlier in section 3. For these start-ups, bonding dominates in the early stages, causing a counter-clockwise movement in the trajectory; for the simple probabilistic evolution in Fig 3, however, all the curves move clockwise. The key point is that this trajectory is neither haphazard, nor does it follow any of the probabilistic curves in Figure 3.

Although simple in form, the dynamics of the start-up evolution depicted in Fig 5 is complex when analyzed in terms of the actual graph structures. To give a modicum of insight, we divide the evolution into seven stages (top of left panel.) The group begins with a leader and two recruits. The next step is

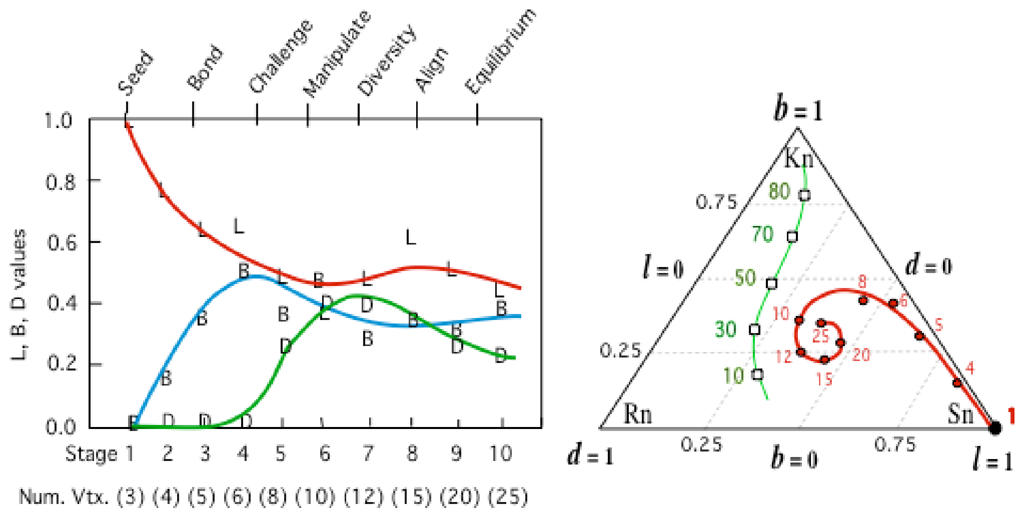

Fig 5. An example of how ***L, B, D*** and ***l, b, d*** vary in the evolution of a small group. In the left panel the size of the group is shown on the bottom line. The red spiral on the right panel shows typical small group evolution, averaged over six small groups. The green curve shows the positions of 20 vertex random graphs with edge probabilities as indicated.

to build a team by bonding and recruitment. This reduces leadership dominance, as shown by the decline in ***L*** in the left panel. As bonding increases, leadership is threatened by a competing maximal node. This requires additional recruitment by the leader, or manipulation of bonds through minor reorganizations. By stage four, additional recruitment of one or more small K_3 cliques has diversified the group. Further alignments (bonding) of these new members then ensues. In the final stages, a crude balance between leadership dominance, bonding and diversity is achieved.

At the right, the simplex shows the same trends, with a move upward along the $\boldsymbol{d} = 0$ locus as bonding increases, followed by a leftward turn toward $\boldsymbol{d} = 1$ as diversity emerges. This is the beginning of a counterclockwise motion toward a potential equilibrium point where a balance among leadership dominance, bonding and diversity emerges. Note that this evolutionary process does not enter the random graph region.

5.0 Discussion

The ***l,b,d*** simplex provides a simple, but revealing representation for small group evolution. Trajectories show whether trends are toward a dominant leadership (dictator), or to cohesion among members (with equal leadership), or to diversity (such as seen in many grass roots organizations.) Strategic moves among members are typically reflected by a change in the trajectory toward one of the vertices. In contrast, if group evolution is a (fixed) probabilistic process (for recruitment, bonding or diversity), all trajectories will move toward the random graph region in the simplex, and eventually toward $\boldsymbol{b} = 1$ if edges and vertices can only be added. With probabilistic edge deletions, haphazard trajectories result (Fig. 4.)

The relation between small group evolution ($n < 100$) and that for very large networks ($n > 1000$) has only begun to be explored. Some proposals for the structure of large networks use global constraints (i.e. probabilities linked to the degree distribution, preferential attachments, etc.), whereas others might use constraints that are more local such as degree conditions or edge probabilities (Newman, 2003, Barabasi & Albert 1999.) These latter influences are generally limited no more than two or three edge steps. Hence, for large networks, the simplex might be more properly applied to their subgraphs, with the subgraphs in turn being treated as independent sets of agents in a network with hierarchical structure. This raises the problem of defining more formally a "small group" and its threshold for the leap into one of the larger networks that others have characterized (Palmer, 1985; Newman, 2003; Dorogovtsev & Mendes, 2003; Pella et al, 2005). Relevant work are studies of changes in network structure initiated by a cascade of influence triggered by a particular small group structure (Watts, 2002; Lescovec et al, 2006).

Cascades are an iterative example of a local dynamics that propagates thru the network. One might envision an analogy in a small group where the connectivity (alignments) of one member enhances that member' status, or favors some particular goal. A simpler form of such dynamics would be for members to ponder possible alignments and recruitments, taking into account the goals of other members. Although the payoffs are unclear (financial, social status, vision for group, etc.), this scenario is one form of a game. The ***l,b,d*** simplex offers a potential tool for studying group evolution as a game, because the consequences of new alignments or manipulations can be calculated for any potential group structure. In principal, optimal trajectories can be discovered, thwarting others along the way. Obviously, however, given the explosion of possible undirected graph types as n increases (e.g. for n = 8, we already have 10K forms), even a two-step look-ahead game is very complex. Nevertheless, given the objectives of each group member (and the recruits), in principal one could calculate whether there is an equilibrium structure.

*Corresponding Author: wrichards@mit.edu
Acknowledgements: Supported by AF MURI, NSERC and Canada Research Chairs Program. Also appreciated were suggestions and data provided by Rajesh Kasturirangan, Kobi Gal, Scott Atran and Owen MacIndoe.

6.0 References

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

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

Bollabas, B. (2001) *Random Graphs*, 2nnd Ed. Cambridge Univ. Press.

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.

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.

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

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

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

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).

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and U.Alon (2002). Network Motifs: simple building blocks of complex networks. Science 298, 824-827.

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

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

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

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

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

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

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

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

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