The Hare and the Tortoise: Network structure in collaborative problem solving

David Lazer
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge MA 02138
David_Lazer@harvard.edu

Allan Friedman
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge MA 02138
Allan_Friedman@ksgphd.harvard.edu

ABSTRACT
Policy makers often face similar problems, and tackling them is often a social effort, where individuals can share information through network ties about how to solve the problems they face. In particular, it has been amply documented that actors tend to emulate successful others that they observe. We present a simulation demonstrating that more efficient communication can actually lower total performance in the long run. This effect is manifest in complex (e.g. rugged) problem spaces, where extensive searches are expensive for the individual. When actors can communicate easily, average performance improves initially, but harder-to-find optimal solutions are less likely to be discovered. Small-world networks and cliques in communication have important implications for the social outcomes in a collaborative network.

Categories and Subject Descriptors

General Terms
Management, Economics, Human Factors,

Keywords
Networks, diffusion, collaboration, small world

1. INTRODUCTION
A critical component of IT-enabled government is communication and collaboration among policy actors as they tackle similar problems. It is often assumed that more information sharing is better. Clearly the internet has facilitated an enormous increase in information sharing, and great effort has been expended to facilitate the spread of information about "best practices" (or, in Europe, "good practices"). This paper critically examines the assumption that more information sharing is better. In particular, we explore the implications of the network architecture through which information can be shared between collaborating policy agents.

We demonstrate, using simulation models, how network architecture affects the balance between exploration and exploitation [7] within a system. It finds that small world networks (networks where the maximum number of degrees of separation between any two actors is small-- e.g., "six degrees"-- see [8]) are better at exploitation—they quickly converge on the best solution that exists in the network at the beginning of the simulation. However, like the speedy hare that takes a nap, small worlds do not improve after that initial burst, because they get stuck on the best local optimum that was reachable early in the history of the simulation. Larger world networks (the tortoise) get closer to the global optimum in the long run because they explore more of the solution space. This effect is greater the more complex is the solution space. Further, controlling for connectedness, we find that the smaller the world the better the system in the short run, and the worse it is in the long run. Finally, we observe how social cliques and network density affect the social outcome.

We argue that these findings are robust across a wide variety of settings, but are particularly important for IT-enabled policy networks. Information technology changes social networks in two ways: (1) it increases the velocity of information through the system, and (2) it increases the formation of distant links, which have a disproportionate impact on decreasing the degrees of separation in a system [10]. The process of learning from the experiences of others is particularly relevant to policy actors, who face common, complex problems with a huge range of adjustable parameters. All are trying to find the best solution but none have the capacity to try every possible combination. Strategic actors compensate by trying to learn from others facing similar problems. In a collaborative network, many actors exchange ideas about this common problem along specific patterns of communication and emulate those who have found success. Ironically, our results suggest that the more efficient the network at spreading information, the lower the long run performance of the system.

2. RESEARCH DESIGN
We propose an information diffusion model based on networked nodes on a complex space, where actors do not know the overall topography of the space, but can observe the success associated with the solutions that they and the nodes with which they communicate have chosen. Each node can be thought of as a policy actor in a population facing similar problems with a common solution. We assume that actors will tend to imitate any superior strategies that they observe, and absent a superior alternative, explore solutions close to its current strategy for an improvement. Over time all the nodes in such a system inevitably converge on some local optimum (which also may happen to be a global optimum).

Information diffusion does not imply information aggregation—the accumulation of unique signals allowing the system to, over time, hone in on "reality." The information cascade literature demonstrates how, if only adoption decision diffuses in the system, fads will result [1],[2]. The reason for this is that the public adoption decisions may, under certain circumstances, outweigh private signals people have that (if aggregated) would point the opposite direction. Strang and Macy, in a simulation model, have extended this essential finding to the circumstances where actors can observe success of other actors but not attribute the origins of that success[9]. In a collaborative network, however, actors can see the underlying mechanics of a successful solution. The question we focus on here is how does the pattern of communication affect the quality of a cooperative solution, and the speed with which they find it?

2.1 NK Space

The first component of our model is the problem space. A complex problem space can be envisioned as a hilly terrain on which myopic actors wander. They are myopic in that they can only see the world immediately around them, to discern in which direction, if any, they must go to ascend. An actor can tell whether they have reached a peak (any movement will make things worse), but have no way of knowing how good that peak is relative to other peaks. A peak can be thought of as one solution among many; the real world is filled with examples of problems where solutions can be identified but actors lack the tools to discern whether a solution is truly optimal.

To model this space, we use Kauffman's NK space ([3], [4]). Originally developed to discuss interlocking genetic attributes, an NK space is represented as a series of components, with each component influencing the score of how "high" the point is. Lateral movement in the space is made by altering the components. A string of N components, each of which offers 2 options creates a space of 2^N possible solutions. Each component may also affect the score contribution of K other components. The NK model's interlocked contributions allow for both gradual and rapid changes the overall score, emulating a multi-peaked space.

Beyond being a standard representation of a complex space, the NK space has been used to model organizational problem solving (e.g. [5]). It is particularly appropriate for modeling policy actors facing complex problems: each component can be thought of as a single policy "lever" in a series of policy options. Adjusting a single lever can result in a simple improvement—or degradation—of the current situation, but it could also have interactive effects with other components and create a drastic change the outcome.

2.2 The Networked Agents

We simulate collaborative actors trying to find the optimal score in an NK network. The simulation is composed of a population of agents, each of whom have an NK string. Agents are also connected as nodes in networks, but each agent is only aware of its immediate neighbors. A simulation lasts for 100 time-steps. For each time step, every agent examines its neighbors looking for a better score. If it finds one, it copies the NK string from the best neighbor. If the agent is unable to find a better solution, it will "explore" by changing one random digit in its NK string. If the resultant string is better, the agent keeps the new string; otherwise, it will revert to the older string. Thus, an agent will tend to mimic other successful agents, and when there is no one to mimic, they will attempt to adapt. New successful adaptations will subsequently be copied by neighbors, and so on.

2.3 Implementation Parameters

Macy and Willer ([6]), among others, have noted that agent-based models with a large parameter space can tempt a researcher to "mine" for attractive results. To create as experimental an environment as possible, we try to minimize the number of variable model parameters. The model is implemented in Repast[1], an object-oriented software library for agent-based modeling. We use 100 agents who interact for 100 time steps: the population is large enough for diffusion processes to occur, but small enough to credibly represent real world phenomena such as organizations. Preliminary testing showed that most models reached equilibrium after 100 turns. For the NK model, we kept N=19, and used K=5, except where noted below. The interlinking factor of 5 produced a space that was sufficiently complex, but not so rugged that gradual adaptation was made futile. All simulations below were run 30 times: 6 randomly seeded starting points on the same five NK landscapes across the different network configurations.

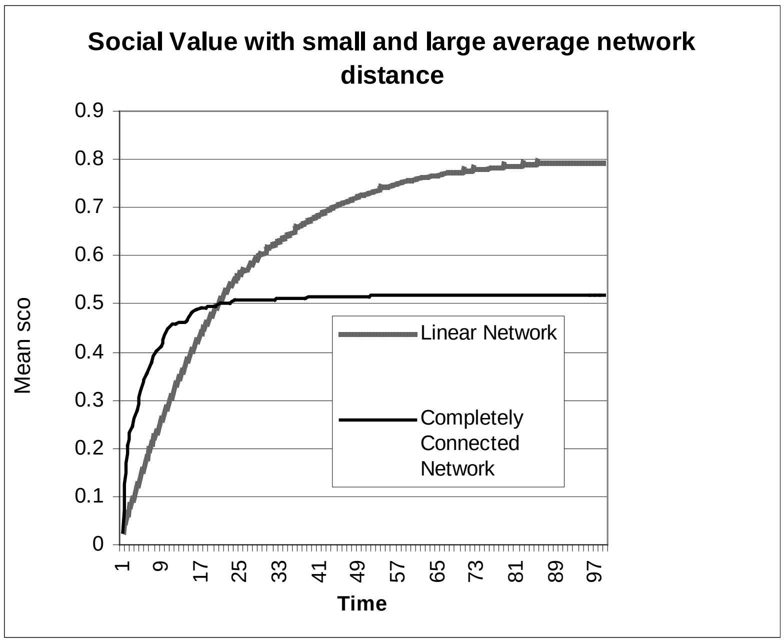

Figure 1 - The mean score in a fully connected network—the ultimate small world network—grows faster, but converges quickly, preventing further exploration in the problem space. The linear network cannot converge as fast, so actors exploring from sub-optimal positions have time to find global maxima.

---

[1] Repast is available here: http://repast.sourceforge.net/ Simulation code was written in Java.

2.4 Measuring Success

To study social exploration and exploitation, we examine how the mean score from each agent's NK string changes over time: we are interested in maximizing social welfare. To normalize the scores across different NK spaces and produce a distribution with a long positive tail, we divide the each agent's score by the global maximum for its space, and then exponentiate this score to the 8^{th} power. This transformation does not alter the ranking of solutions, but distributes them numerically to show that some are better than others more clearly.

3. RESULTS

In these simulations, we use archetypical networks with certain specific properties to analyze the effect of those properties on the efficacy and speed of social problem solving. The most obvious comparison is between extremes of connectivity. Figure 1 compares the social outcomes of the smallest world possible, where every node can talk to every other node, and the largest possible (connected) network, where each actor can only talk to two other agents to form a line. The maximum shortest distance between any two nodes on the former is 1; the maximum distance for the latter is equal to the population size. There is a dramatic difference between the two networks: the completely connected network converges very quickly on a decent peak, but once all the actors have converged to that solution, it becomes difficult to find a better one using random exploration. Convergence in the linear network, on the other hand takes an average 50 time steps without any exploration. During that time, the heterogeneous population can explore from their starting points, and some actors are likely to find higher peaks, to which the population gradually converges. The final convergence point in a fully connected graph will never be far from one of the original starting points.

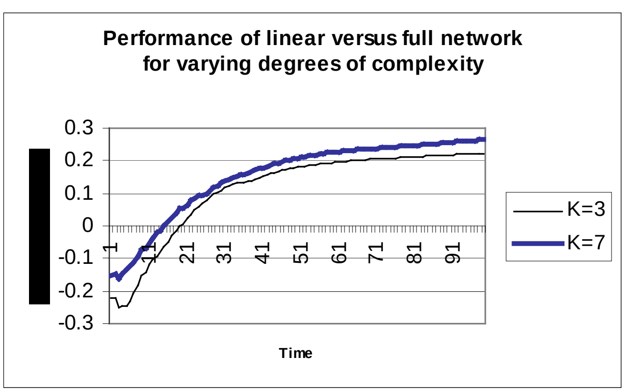

Figure 2: A linear network's performance improves against a complete network both in short term and long term searches as the problem space grows more complex.

At first, this is counter-intuitive: more communication should result in an inferior solution. The superior outcome of a sparser network is tied to the complex nature of the NK problem space. In the trivially simple case of a single peak, the small world network would be able to climb faster, but both models would reach the maximum. As K approaches N and the space gets incredibly rugged, the value of "walking" uphill diminishes because there are so many peaks. In a moderate range, however, the success of a network is related to the complexity of the space. We test the effects of using a fairly rugged space and a simpler space, and find that more complex spaces increase the power of the sparse network. Figure 2 compares the linear and fully connected models for k=3 and k=7. At the higher level of complexity, the early advantage of the connected model diminishes, and the linear model produces superior aggregate outcomes earlier.

Between the archetypes of a fully connected network and a network that maximizes social distance, there are a myriad of network configurations. One such configuration is a star network, with one hub and 99 spokes (maximum degrees of separation two), like that seen in Figure 3. A star network is a stylized model of centralized information flow, with all actors talking to a central authority. This network provides a level of performance between the fully connected and line networks. If two "leaf" nodes in a stable system both find new peaks, the center node will copy the better of the two in the first turn. That means the un-copied leaf can explore for one more turn, trying to find a still higher peak from its new vantage point.

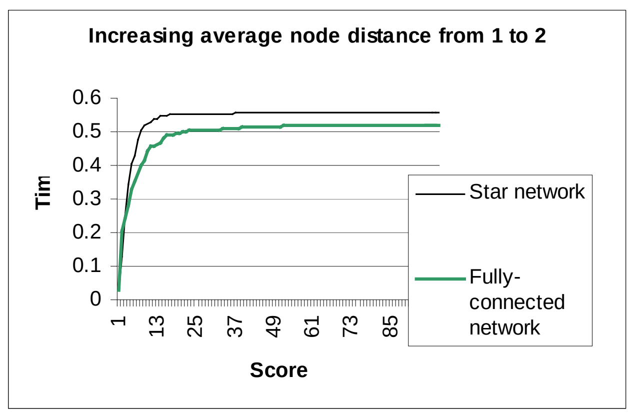

Figure 3: A star network is still a very small world, but has a diameter of 2, rather than 1. This dramatically improves the solution in which the population converges, although it is still suboptimal to (and faster than) the linear network.

As discussed above, a grid network is often used in social simulation models as a stylized way of representing social interactions. Each actor has four contacts, who have three other contacts. As indicated in Figure 4, the grid (actually a torus in this simulation) is worse in the short run than the star and fully connected graphs, but better in the long run; and, conversely better than the linear graph in the short run, but worse in the long run.

In this grid, the maximum diameter is one less than the square root of the population, or nine hops for our population of 100. To increase this diameter, we divide the population into four cliques with more common internal ties and fewer external connections. Each of these populations is connected as a grid (not a torus), and then the clusters are connected by a single link. Now the shortest path between the two most distant nodes is 36 links. Moreover, the population has been divided into subpopulations with limited interactions. Inside a cluster, it is easy to converge on the local maximum, but the nodes that border the neighboring cluster can introduce a higher point. Figure 4 shows that this inter-population sharing is an improvement over the more fluid diffusion of the grid.

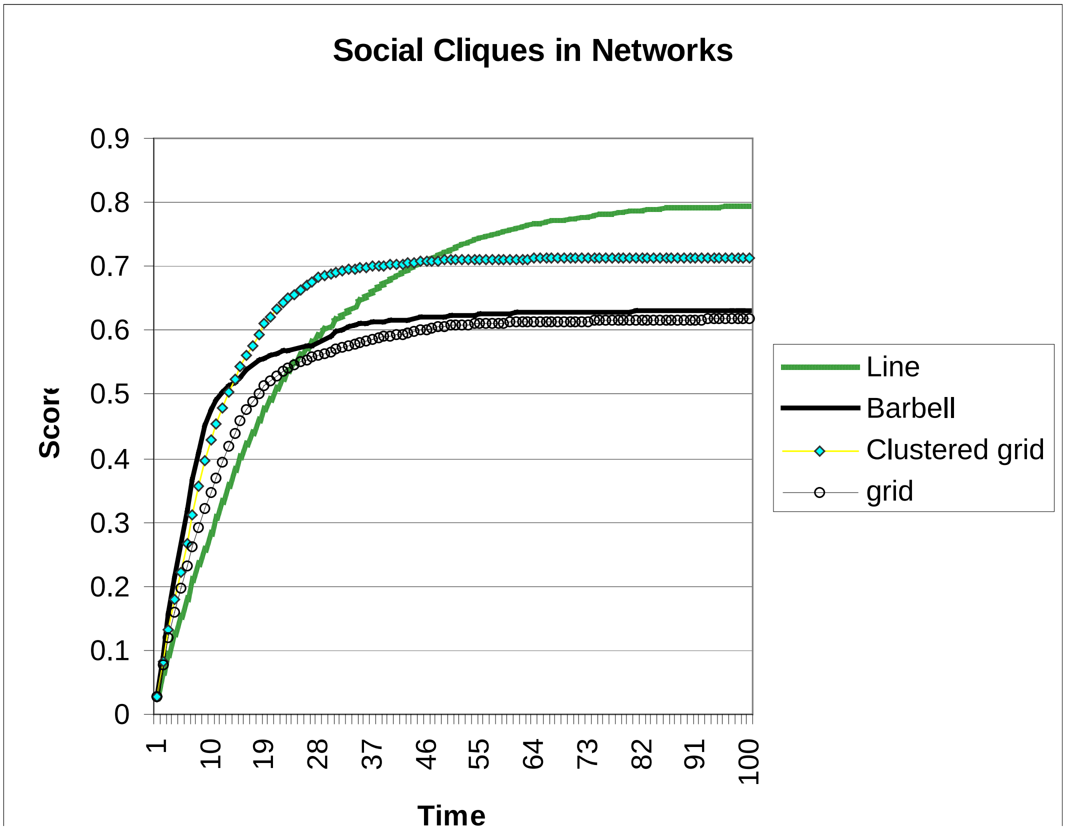

Figure 4: Breaking up a grid to decrease connectedness improves long-term performance at the expense of speed of search; adding cliques to a line increases short term gains but reduces the quality of the final solution.

Clustering can both divide a population and make a small world network less so. To investigate whether clustering can help if it makes the network into a *smaller* world, we introduce two small loops to the simple linear network, one on either end to form a barbell-like shape. This creates two groups that can communicated faster, because the maximum path length inside the group is half that of a linear array. These two groups are linked by a linear network comprising half the population, which is still long enough to delay propagation. The two groups can explore independently, and then converge on the best solution. The effect is dramatic: the final solution is lower, but reached much faster. Figure 4 illustrates that even a small increase in the density of the network can increase . the nodes' ability to converge on a solution faster, while decreasing their ability to thoroughly search the solution space After examining some very basic constructed networks, we test our findings on an Erdos-Renyi random graph. E-R graphs are constructed by pairing any two nodes with probability P. P also serves as a density measure, since the number of links an average node will have is equal to P times the population. E-R graphs constructed in this method are not guaranteed to have every node connected to the others in a single component; nodes can be isolated, or in exclusive dyads, triads, and so on. As P goes up, the probability of all nodes being connected increases rapidly.

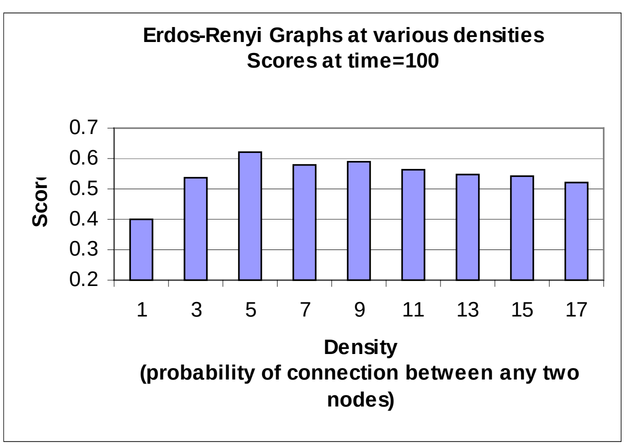

Figure 5: The relationship between network density and social outcome for random graphs is curvilinear. Low densities result in isolated nodes that cannot converge, while high clusters fall prey to the small-world over-connectedness discussed above.

Isolated nodes cannot converge with the population, nor can they share their own optima with others, so their presence can bring down the mean. On the other hand, as P grows, the network gets more dense, which in turn brings about faster, less optimal convergence. Figure 5 indicates a maximum density for effectiveness of search, where the system's ability to find the optimal solution after 100 time-steps peaks when two randomly selected nodes have a 5% chance of sharing a link. Before then, the network is less likely to be a single component; the same issues with small world networks discussed above appear when the random network grows too dense.

4. DISCUSSION

The core result above is that there is a trade-off between networks that are good at exploring solution spaces, and networks that are good at exploiting social spaces. Greater connectivity can result in finding a decent solution quickly, but it can prevent a collaborative network from finding the optimal solution.

In this age of connectedness, a great deal of energy is spent to further increase the connectedness of systems. In fact, much of the practical management application of social network analysis is spent attempting to break down silos and bridging the structural holes of organizations. However, this paper suggests that some of these interventions may be doing more harm than good—that connectedness creates a system-wide groupthink, a homogeneity that is beneficial in the short run, but dysfunctional in the long run. Removing a small number of connections in a dense group, or adding a few ties in a sparse group can dramatically improve the quality of the outcome, or the speed with which a decent outcome is found. Further work can focus on formalizing the relationship between specific network properties and the speed and quality of outcomes, as well as understanding the potential of individual nodes to affect the social outcome by acting as an "obstacle" or an "entrepreneur". In the process of understanding more about how people work together, we must revisit the advantages of collaboration and understand that all cooperative arrangements are not equal.

5. REFERENCES

[1] Banerjee, A V. 1992. "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107/3: 797-817.

[2] Bikhchandani, S, Hirshleifer D., and Welch, I. 1992. "A Theory of Fads, Fashions, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100/5: 992-1026.

[3] Kauffman, S. 1991. Antichaos and Adaptation. *Scientific American*, 265(2), 78-84

[4] Kauffman, S. 1995. *At Home in the Universe*.

[5] Levinthal, D.A. 1997. Adaptation on Rugged Landscapes. *Management Science* 43(7), 934-950

[6] Macy, M. W. and Willer, R. 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology* 28:143-66.

[7] March, J. 1991. "Exploration and exploitation in organizational learning." *Organization Science*, 2(1): 71-87

[8] Milgram, S. (1967) . The small-world problem. Psychology Today 1, 61-67.

[9] Strang, D., and Macy, M. 2001. "'In Search of Excellence': Fads, Success Stories, and Adaptive Emulation" *American Journal of Sociology*, 106.

[10] Watts, D. and Strogatts, S.. 1998, Collective dynamics of 'small-world' networks, Nature 393:440-42.Bowman, B., Debray, S. K., and Peterson, L. L. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst., 15,* 5 (Nov. 1993), 795-825.