Information Gathering Strategies in Social Networks

Eleanor Brush

December 27, 2012

Contents

1 The Model

We consider a network where each agent has an interval variable, keeping track of, for example, direction or velocity of movement or a desired location, and uses the variables of the neighbors to whom it is connected to update its variable. More concretely, there are n nodes, which each have a variable x_i and receive information from its neighbors weighted by a factor w_{ij} such that $\sum_j w_{ij} = 1$ for all $i = 1, \ldots, n$. For instance, node i may have k connections, each of which it values equally so that $w_{ij} = 1/k$ for those nodes from whom i receives information and $w_{ij} = 0$ otherwise. In continuous time, the dynamics of the model are given by

$$\frac{dx_i}{dt}(t) = \sum_j w_{ij}(x_j(t) - x_i(t)) + \xi_i$$

where ξ_i is random noise. Let L be the Laplacian of the weighted matrix given by the w_{ij} so that

$$L_{ij} = \begin{cases} 1 & , \quad \text{if } i = j \\ -w_{ij} & , \quad \quad \text{else} \end{cases}$$

so that the model can be written as

$$\dot{x}(t) = -Lx(t) + \xi$$

Suppose there is external or environmental information that the agents can gather and that the agents perform better if they gather more accurate information. In particular, suppose that a node receives an external signal s, which the rest of the agents want to learn as quickly and as accurately as possible. We will suppose that the agent that receives the external signal is chosen at random. For now, we assume that the first node is chosen. Having received the external signal, that agent no longer

listens to the other agents in the system, so that its dynamics are changed to $x_1(t) = s$ for all t. The dynamics for the other nodes become

$$\frac{dx_i}{dt} = w_{i1}(s - x_i(t)) + \sum_{j \neq 1} w_{ij}(x_j(t) - x_i(t)) + \xi_i(t).$$

A consensus state is again the equilibrium, but now the consensus value is fixed at s, the value of the external signal.

2 Measures of deviation

Different nodes in the network might be subject to different sources and magnitudes of noise. We will always assume that $\mathbb{E}[\xi_i] = 0$. Let $d_i = Var[\xi_i]$ and $D = diag(d_1, \ldots, d_n)$. In particular, we will set d_i equal to the number of neighbors influencing node i, so that nodes with more neighbors make noisier estimates.

Let y represent the distance between the vector of opinions and consensus, i.e. $y = QxQ^T$, where $Q \in \mathbb{R}^{n-1 \times n}$ is such that $Q\vec{1} = 0$, $QQ^T = I_{n-1}$, and $Q^T Q = I_n - \frac{1}{n}\vec{1} \times \vec{1}^T$. Then

$$\dot{y}(t) = -\overline{L}y(t) + Q\xi(t)$$

where $\overline{L} = QLQ^T$. As shown in [3], if $\Sigma_y(t) = \mathbb{E}[y(t)y(t)^T]$, Σ_y at equilibrium solves

$$0 = -\overline{L}\Sigma_y - \Sigma_y \overline{L}^T + D. \tag{1}$$

If we instead consider deviation from the external signal, $\delta_i(t) = x_i(t) - s$,

$$\dot{\delta}_i(t) = \dot{x}_i(t)$$
$$= \sum_j w_{ij}(x_j(t) - x_i(t)) + \xi_i(t)$$
$$= \sum_j w_{ij}(\delta_j(t) - \delta_i(t)) + \xi_i(t)$$
$$= \sum_{j \neq 1} w_{ij}\delta_j(t) - \delta_i(t)\sum_j w_{ij} + \xi_i(t) \text{ since } \delta_1(t) \equiv 0$$

so that $\dot{\delta}(t) = -L^1 \delta(t) + \xi(t)$

where L^1 is L with the first row and column removed and $\delta(t) = (\delta_2(t)\ldots\delta_n(t))$. If $\Sigma(t) = \mathbb{E}[\delta(t)\delta(t)^T]$, Σ at equilibrium solves the analogous Lyapunov equation,

$$0 = -L^1\Sigma - \Sigma(L^1)^T + D. \tag{2}$$

3 Relationship between consensus and external signal models

Claim 1 TO BE PROVED: Let L be the Laplacian of the weighted matrix and L^1 be L without its first row and column. Let the second smallest eigenvalue of L be λ and the smallest eigenvalue of L^1 be λ_1. Then $1 \geq \lambda \geq \lambda_1 \geq 0$. In particular,

1. $\lambda_1 = 0$ if and only if $w_{i1} = 0$ for all i and

2. if $w_{i1} = 1$ for all i then $\lambda_1 = 1$.

Claim 2 $||\delta(t)|| \geq ||y(t)||$ for all t.

Proof. For the following, let $\langle x \rangle = \frac{\sum_i x_i}{n}$.

$$y = x - \langle x \rangle \text{ and } \delta = x - s$$

$$\Rightarrow ||y||^2 = \sum_i (x_i^2 - 2x_i\langle x \rangle + \langle x \rangle^2) \text{ and } ||\delta||^2 = \sum_i (x_i^2 - 2x_i s + s^2)$$

$$\Rightarrow ||y||^2 = \sum_i x_i^2 - 2n\langle x \rangle^2 + n\langle x \rangle^2 \text{ and } ||\delta||^2 = \sum_i x_i^2 - 2ns\langle x \rangle + ns^2$$

$$\Rightarrow ||\delta||^2 - ||y||^2 = ns^2 - 2ns\langle x \rangle + n\langle x \rangle^2$$

$$= n(\langle x \rangle - s)^2$$

$$\Rightarrow ||\delta||^2 - ||y||^2 \geq 0 \text{ with equality iff } \langle x \rangle = s$$

$$\Rightarrow ||\delta||^2 \geq ||y||^2 \text{ with equality iff } \langle x \rangle = s$$

$$\Rightarrow ||\delta|| \geq ||y|| \text{ with equality iff } \langle x \rangle = s$$

∎

Therefore, $\lim_{t \to \infty} E[||\delta(t)||] \geq \lim_{t \to \infty} E[||y(t)||]$.

4   Measures of Performance

There are two general measures of performance: the speed with which the external signal can be learned and the steady state variance around the signal. There is a tradeoff between these two measures. As above let L be the graph Laplacian of the network, $\overline{L}$ be the transformation of L, and L^i be L with the ith row and column removed, which describes how a signal from individual i propagates across the network. Let λ be the smallest eigenvalue of $\overline{L}$ and let λ^i be the smallest eigenvalue of L^i and v^i its associated eigenvector. Let Σ_y and Σ be the solutions to (1) and (2) respectively. The eigenvalues λ and λ^i reflect how quickly consensus is reached or the external signal is learned and Σ_y and Σ are informative about the variance around consensus or the external signal. We introduce a parameter α to represent the tradeoff between speed and variance.

Our measure of group performance will be

$$p_G = \left(\frac{1}{e^{-\lambda}}\right)^{(1-\alpha)} \left(\frac{\sqrt{N}}{\sqrt{\text{Tr}(\Sigma_y)}}\right)^{\alpha} \tag{3}$$

where we take reciprocals so that a higher value of p_G is associated with better performance (or fitness). Our measure of individual performance for node j will be

$$p_j = \left(\frac{1}{e^{-\lambda^i}v_j}\right)^{(1-\alpha)} \left(\frac{1}{\sqrt{\Sigma_{jj}}}\right)^{\alpha} \tag{4}$$

5   Measure of a Node's Position Predicts Individual Performance

6   Equilibrium Strategies

6.1   Group Optimization If we assume the whole group has the same strategy, we can ask what strategy gives the best performance for the group (and for every individual, since they are assumed to be identical and thus receive the same payoffs), for various levels of the tradeoff α between speed and variance (Figure 1). Unsurprisingly, when speed is more important (low α), it is best for the group for every node to have many neighbors. When variance is more important (high α), it is best for every node to have few neighbors. What is somewhat surprising is that at intermediate levels of tradeoff, there is not a optimum intermediate between low and high strategies. Instead, the landscape is rather flat and every strategy performs about equally well.

Group Performance, N=20
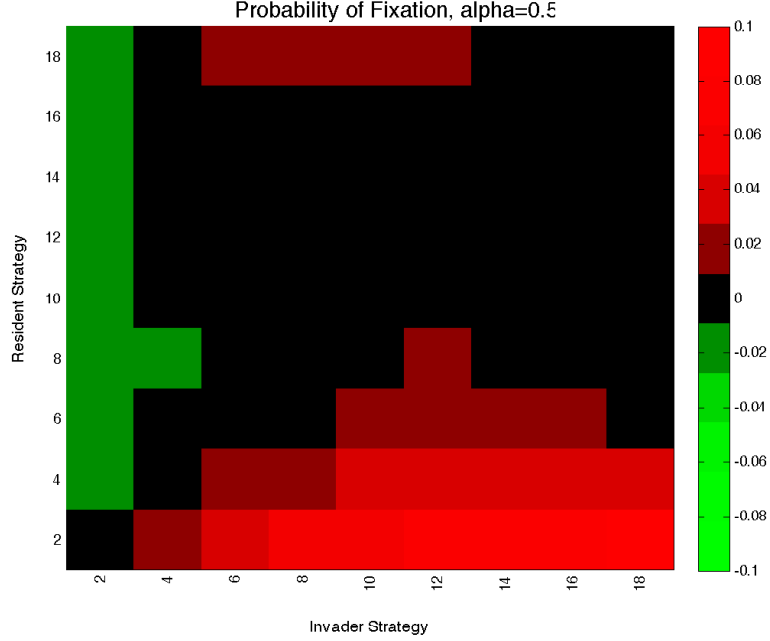
(a) $N = 20$
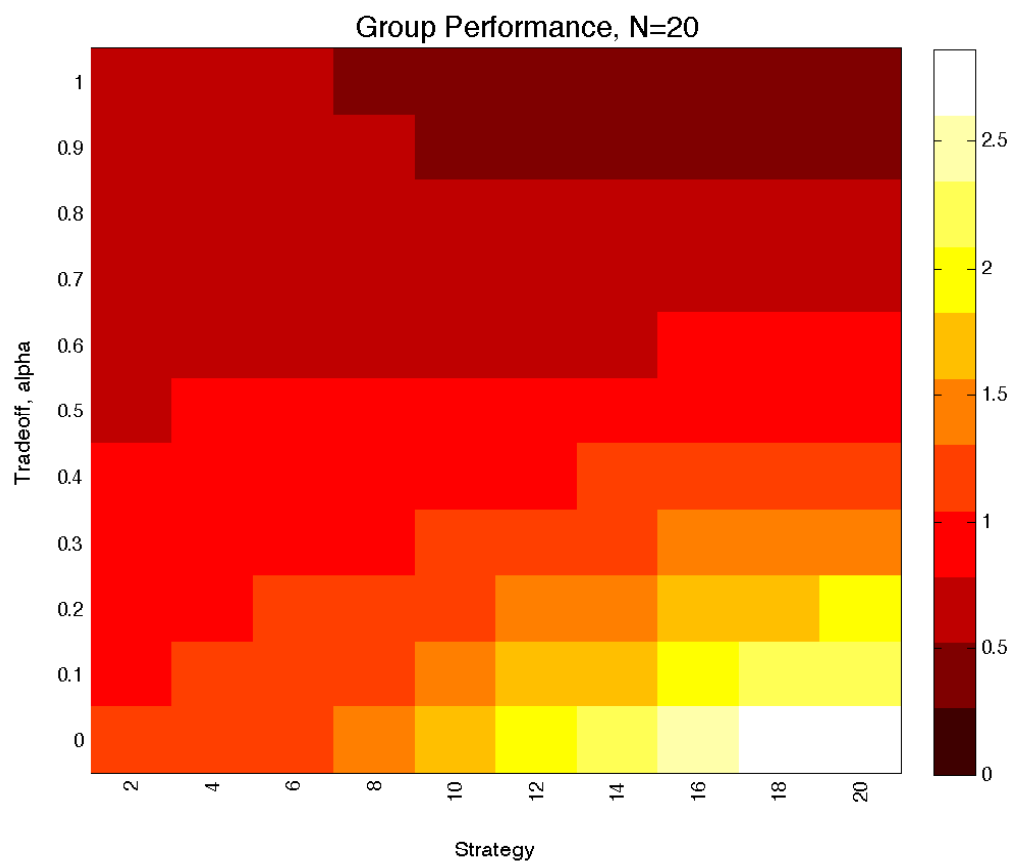
Group Performance, N=40
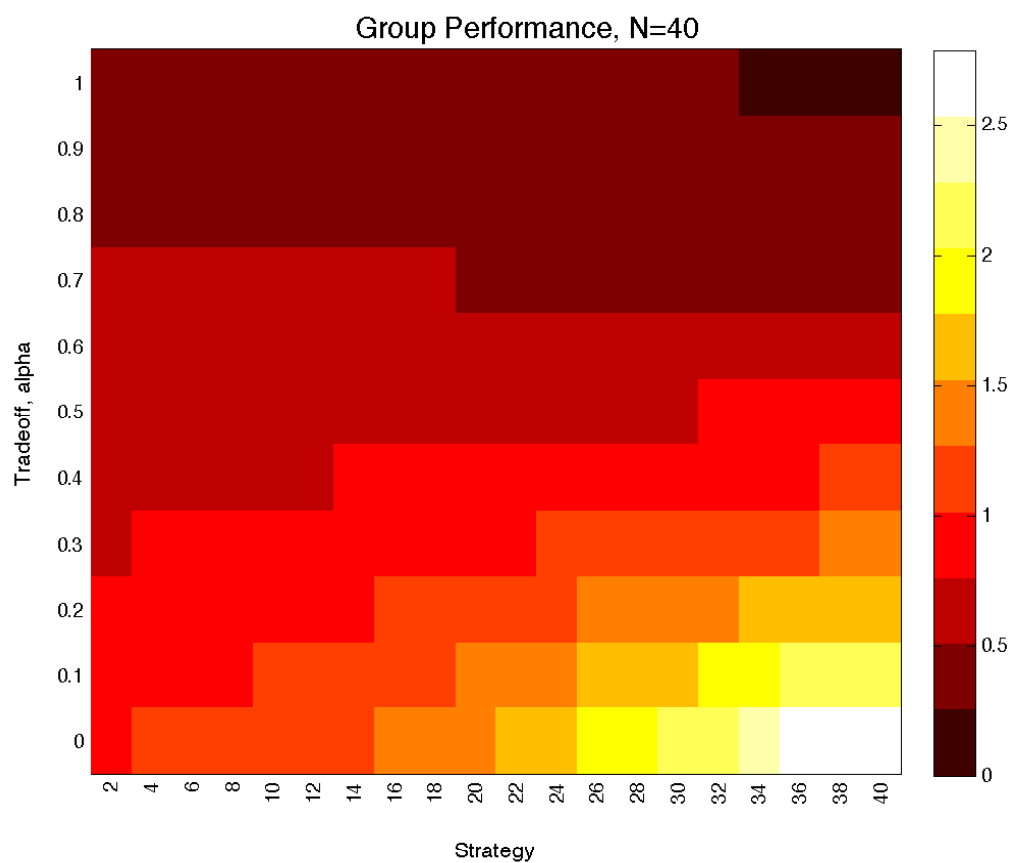
Figure 2: This figure shows, for each pair of resident and invader strategies, the difference between the probability of the invader strategy reaching fixation and the fixation probability under neutral dynamics, $p_f - \frac{1}{N}$.

6.2 Individual Optimization

6.2.1 Adaptive Dynamics

As shown in [2], if f_i is the fitness of an invader and g_i is the fitness of a resident when there are i invaders, the probability of the number of invaders increasing by one is given by

$$p_{i \to i+1} = \frac{i(N-i)f_i}{N(if_i + (N-i)g_i)} \tag{5}$$

and the probability of the number of invaders decreasing by one is given by

$$p_{i \to i-1} = \frac{i(N-i)g_i}{N(if_i + (N-i)g_i)} \tag{6}$$

Together these give the probability that the invader will come to fixation in the population:

$$p_f = \frac{1}{1 + \sum_{k=1}^{N-1} \Pi_{j=1}^{k} \frac{p_{i \to i+1}}{p_{i \to i-1}}} \tag{7}$$

Here, individual fitnesses are given by (4). Since individual fitness depends on where the node is in the network in addition to its strategy, we say that $f_i = \langle p_j \rangle_{j: \, j \text{ is an invader}}$ and $g_i = \langle p_j \rangle_{j: \, j \text{ is an resident}}$ when there are i invaders. For an intermediate level of tradeoff between speed and variance, invaders with intermediate strategies have the highest probability of reaching fixation (Figure 2).

A reasonable definition of an equilibrium is a situation in which, for a given pair of resident and invader strategies, the probability of the number of invaders increasing by one is equal to the number of invaders decreasing by one so that the population is at a probabilistic steady state. The probabilities may never be exactly equal. However, there is a number of invaders such that the probability of increasing by one is greater than the probability of decreasing by one but if the number increases by
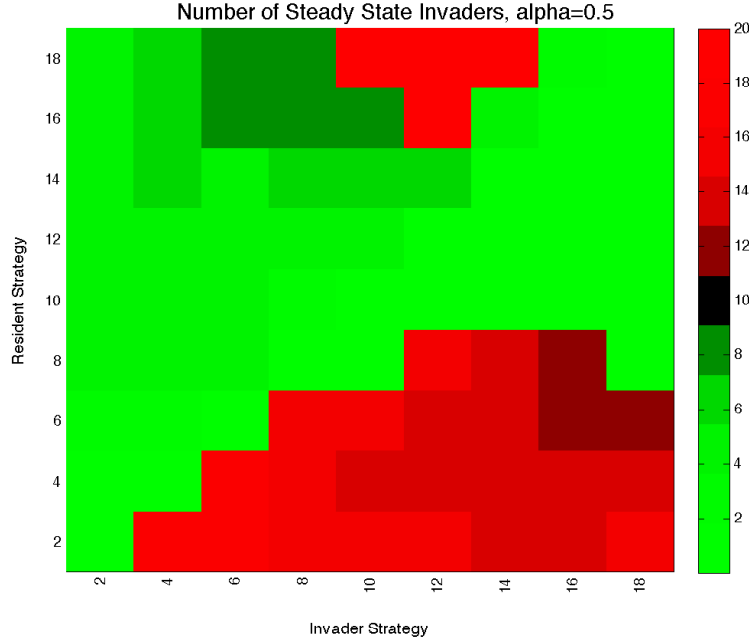
Figure 3: This figure shows, for each pair of resident and invader strategies, the steady state number of invaders.

one the probability of decreasing becomes greater than the probability of increasing so that the number of invaders will go back and forth between these two sizes. This is our provisional definition of a steady state number of invaders. For a given level of tradeoff between speed and variance ($\alpha = .5$) and network size ($N = 20$) we find this steady state number (Figure 3).

If the population goes from a monomorphic one to one in which the invaders have reached their steady state size (as described above), does this improve or decrease the overall group performance? In Figure (4) we show this change in group performance.

6.2.2 Greedy Optimization

Rather than using adaptive dynamics to study how strategies evolve over time, we can allow the nodes themselves to optimize their strategies. Given the set of strategies being used, each node can evaluate its performance and what its performance would have been had its strategy been one higher or lower and change its strategy accordingly. We let the nodes perform this greedy optimization for several iterations until the strategies come to steady state. (The steady state may oscillate between two adjacent strategies). Figures 5-6 show the results of these optimizations, for different initial configurations of the network.

7   Future Work

1. prove Claim 1 about eigenvalues

2. find ESS and CIS

3. dependence of equilibria on the size of the network

4. Information centrality predicts the accuracy of each individual's opinion and eigenvector centrality predicts the speed of each individual's opinion. Is there a measure that predicts performance when both speed and accuracy are valued?
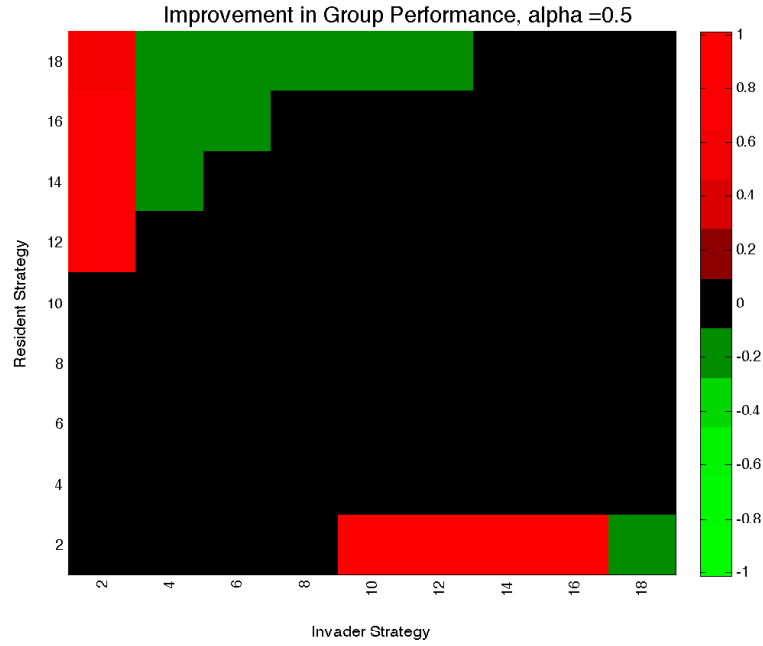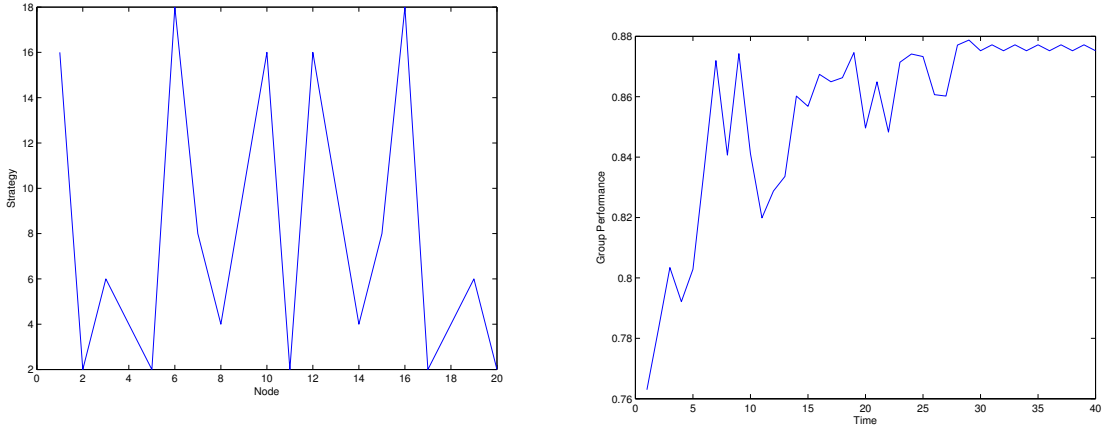
Figure 4: Here we show the effect on the group performance when a mutant invades and the invader population grows to the size that can be sustained by the group, as shown in Figure 3.
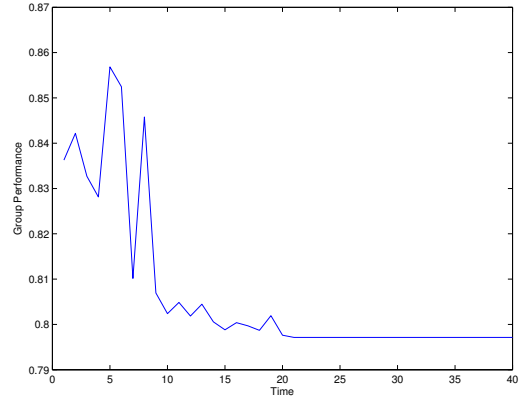

(a) The end result of a greedy optimization, where the x-axis represents the nodes and the y-axis the final strategy. The node that started with a different strategy, 10 as opposed to 2 is first. The greedy optimization results in spatial periodicity in strategy.

(b) Group performance changes over time as the nodes perform the greedy optimization. The greedy optimization tends to increase overall group performance.

Figure 5: The results of a greedy optimization, where at each time step each node increases its strategy by 2, decreases its strategy by 2, or stays the same, depending on which performs best given that the rest of the network has the strategies they do. Initially, the population all has strategy 2, except for the first node which has strategy 10. ($\alpha = .5$, $N = 20$)
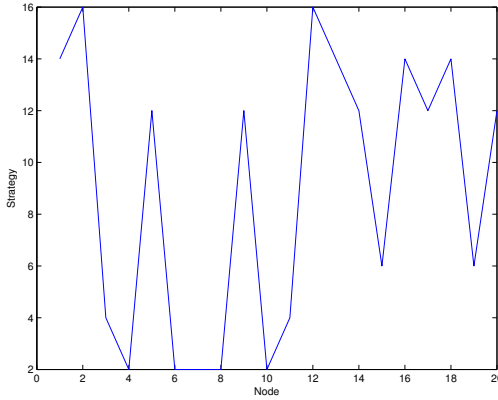
(a) The end result of a greedy optimization, where the x-axis represents the nodes and the y-axis the final strategy. The node that started with a different strategy, 10 in a population of 2, is first. The greedy optimization results in spatial periodicity in strategy.

(b) Group performance changes over time as the nodes perform the greedy optimization. In this case, greedy optimization tends to decrease group performance.

Figure 6: The results of a greedy optimization, where at each time step each node increases its strategy by 2, decreases its strategy by 2, or stays the same, depending on which performs best given that the rest of the network has the strategies they do. Initially, the population randomly drawn strategies. ($\alpha = .5$, $N = 20$)

5. prove that spatial periodicity in strategy is optimal for group performance

6. perform these analysis on more (empirical) networks

8 Conclusions

1. from the perspective of network design, measure X tells how to optimize individual performance

2. optimal topology for the group when we also consider speed to consensus

3. group performance does or doesn't or in certain conditions (i.e. size of group, tradeoff between speed and accuracy) emerge out of individual level considerations so group performance does or doesn't or in certain conditions have to be invoked

References

[1] Carl Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 2001.

[2] Martin Nowak. *Evolutionary Dynamics*. Belknap / Harvard, 2006.

[3] George Forrest Young, Luca Scardovi, and Naomi Ehrich Leonard. Robustness of noisy consensus dynamics with directed communication. *Proceedings of the American Control Conference*, 2010.