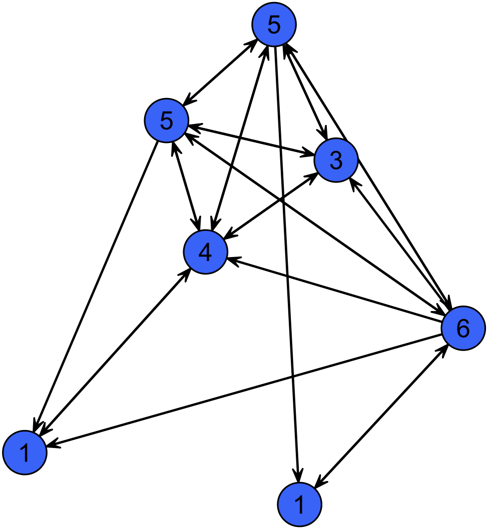
Not ‘Won’t you be my neighbor?’ But ‘Should you be my neighbor?’

Optimal Information Gathering in Social Networks

Eleanor Brush

September 26, 2012



5
5
3
4
6
1
1

What's the optimal way to gather information from a social network?

What's the optimal way to gather information from a social network?

1. a criterion of performance

What's the optimal way to gather information from a social network?

1. a criterion of performance
2. a strategy to maximize that criterion

Overview

1. a criterion for the whole network
2. a criterion for individuals
3. consequences of and strategy for optimizing individual performance
4. a measure of informational burden
5. future directions

$$\dot{x}_i(t) = \sum_j w_{ij}\big(x_j(t) - x_i(t)\big) + \xi_i(t)$$

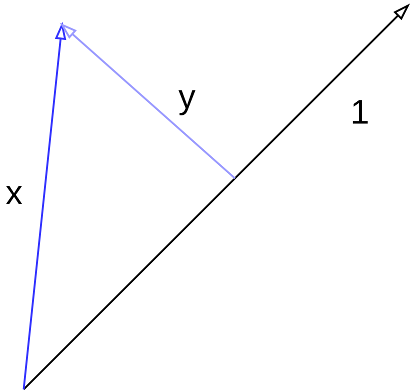
where $\sum_j w_{ij} = 1$ and $\xi_i(t) \sim N(0,1)$ i.i.d.

$\implies$ consensus with every $x_i = c$

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

$$\text{where } L_{ij} = \begin{cases} 1 & \text{if } i = j \\ -w_{ij} & \text{if } i \neq j \end{cases}$$

$\Longrightarrow$ consensus with $x \propto \vec{1}$


x
y
1

Deviation from Consensus

Define the orthogonal transformation $Q \in \mathbb{R}^{n-1 \times n}$ with

1. $Q\vec{1} = 0$
2. $QQ^T = I_{n-1}$ and
3. $Q^T Q = I_n - \frac{1}{n}\vec{1} \times \vec{1}^T$

Define $y = Qx$ and $\overline{L} = QLQ^T$.

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

Starling flock networks manage uncertainty in consensus at low cost,
George F. Young, Luca Scardovi, Andrea Cavagna, Irene Giardina, and Naomi E. Leonard. 2012.

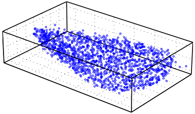
Network Optimization

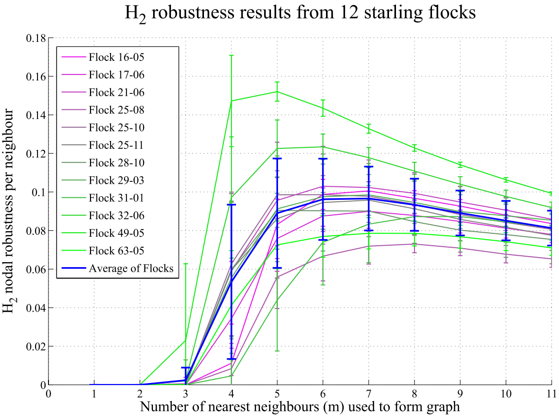
One measure of performance:

$$H(L) = \lim_{t \to \infty} E[||y(t)||]$$

Objective: choose the number of neighbors m that maximizes

$$\frac{1}{\left(\frac{H}{\sqrt{N}}\right) m}$$





Starling flock networks manage uncertainty in consensus at low cost,
George F. Young, Luca Scardovi, Andrea Cavagna, Irene Giardina, and Naomi E. Leonard. 2012.

Learning about the world

One node gets an external signal so that

$$
\begin{aligned}
x_i(t) &= \text{signal} \;\; \text{for all } t \text{ and} \\
\dot{x}_j(t) &= \sum_k w_{jk}\big(x_k(t) - x_j(t)\big) + \xi_j(t) \text{ for } k \neq i \\
\delta_j(t) &:= \;\text{signal}\; - x_j(t) \\
\Rightarrow \dot{\delta}(t) &= -L^i\delta(t) + L^i\xi(t)
\end{aligned}
$$

where $L^i = L$ without the i^{th} row and column

We can approximate $\delta_j(t)$ with

$$\delta_j(t) \sim e^{-\lambda^i t} v_j^i$$

where λ^i is the smallest eigenvalue of L^i and v^i is its eigenvector.

Deviations

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

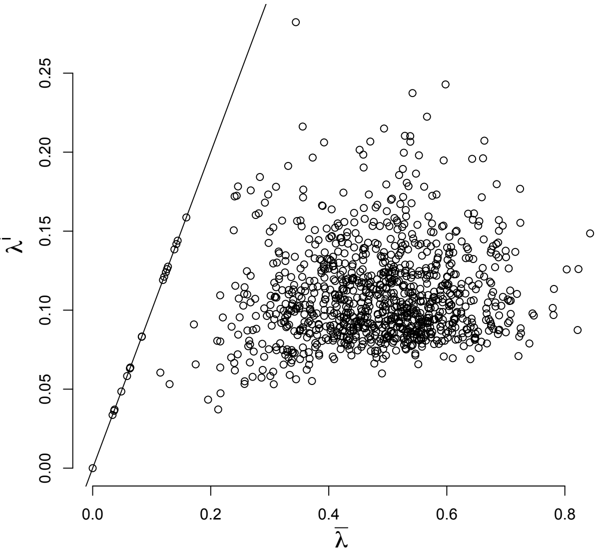
from external signal: $\dot{\delta}(t) = -L^i\delta(t) - L^i\xi(t)$

Minimizing

$$\delta_j(t) \sim e^{-\lambda^i t} v_j^i$$

usually optimizes consensus in the whole group because

1. the rate of convergence to consensus is given by $\overline{\lambda}$, the smallest eigenvalue of $\overline{L}$, and $\lambda^i \leq \overline{\lambda}$



0.00
0.05
0.10
0.15
0.20
0.25
λ^i
0.0
0.2
0.4
0.6
0.8
$\bar{\lambda}$

Deviations

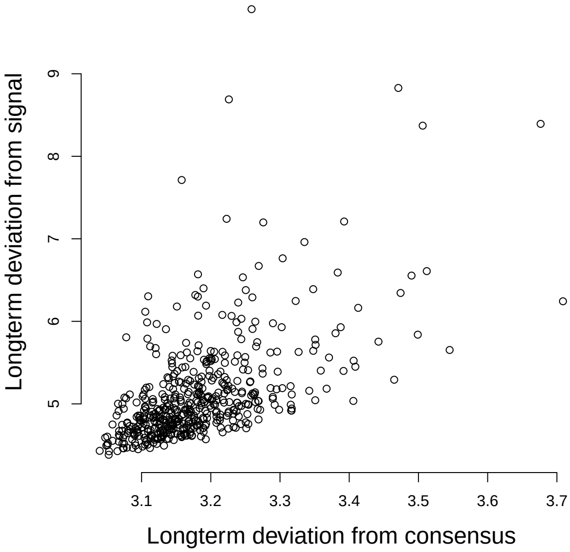
from consensus: $\dot{y}(t) = -\overline{L}y(t) + Q\xi(t)$
from external signal: $\dot{\delta}(t) = -L^i\delta(t) - L^i\xi(t)$

Minimizing

$$\delta_j(t) \sim e^{-\lambda^i t} v_j^i$$

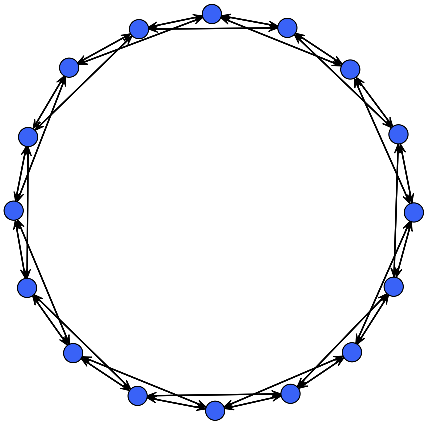
usually optimizes consensus in the whole group because

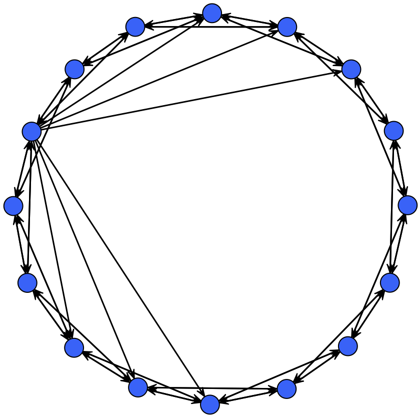
1. the rate of convergence to consensus is given by $\overline{\lambda}$, the smallest eigenvalue of $\overline{L}$, and $\lambda^i \leq \overline{\lambda}$
2. $\lim_{t\to\infty} E[\|\delta(t)\|] \geq \lim_{t\to\infty} E[\|y(t)\|]$


Longterm deviation from signal
5
6
7
8
9
3.1
3.2
3.3
3.4
3.5
3.6
3.7
Longterm deviation from consensus

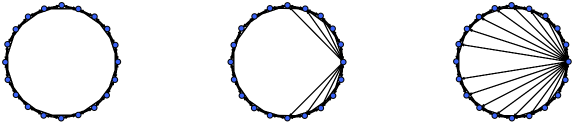
Individual Optimization

Objective: each node j wants to minimize $s_j = \langle e^{-\lambda^i} v_j^i \rangle_i$

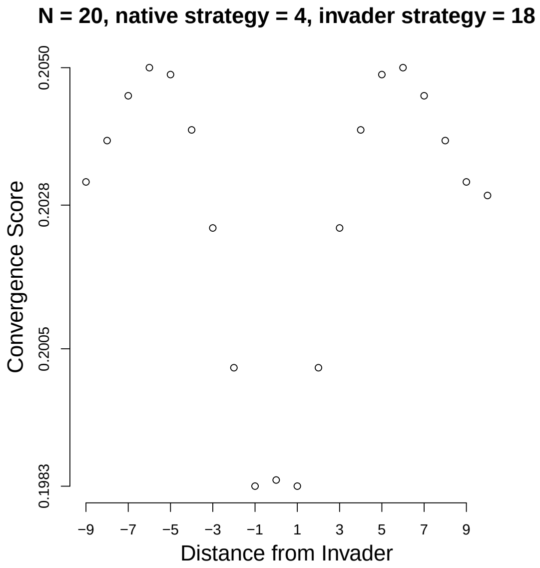


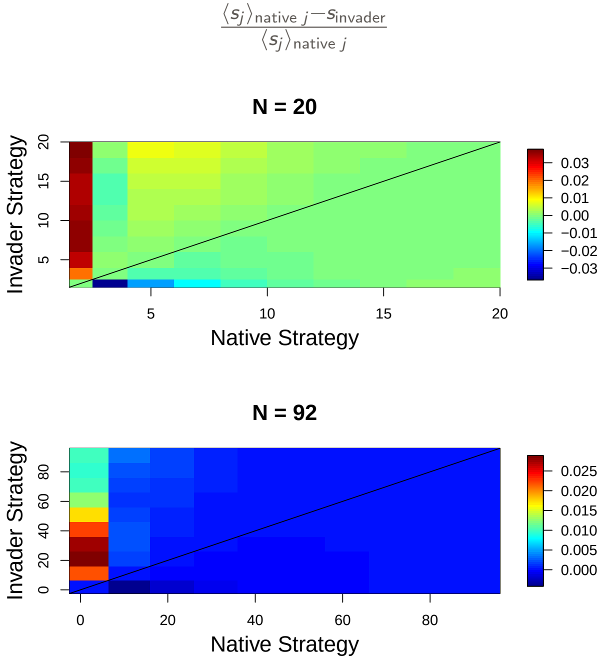


Optimizing s_j affects everyone.

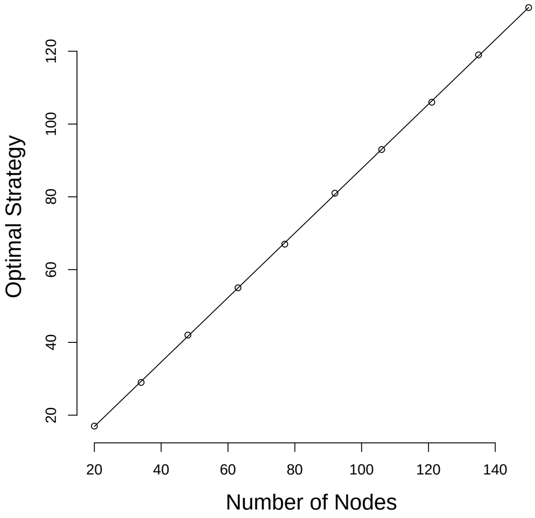


s_{invader}	0.2018	0.1993	0.1984
$\langle s_j \rangle_{\text{native } j}$	0.2018	0.2025	0.2029
$\lim_{t\to\infty} E[\lVert y(t) \rVert]$	4.10	3.99	3.89
$\lim_{t\to\infty} E[\lVert \delta(t) \rVert]_{i=10}$	5.98	5.48	5.09


N = 20, native strategy = 4, invader strategy = 18
Convergence Score
0.2050
0.2028
0.2005
0.1983
−9
−7
−5
−3
−1
1
3
5
7
9
Distance from Invader



⟨s_j⟩native j − s_invader / ⟨s_j⟩native j
N = 20
Invader Strategy
Native Strategy
5
10
15
20
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
N = 92
Invader Strategy
Native Strategy
0
20
40
60
80
0.025
0.020
0.015
0.010
0.005
0.000



Optimal Strategy
20
40
60
80
100
120
20
40
60
80
100
120
140
Number of Nodes

Future Directions

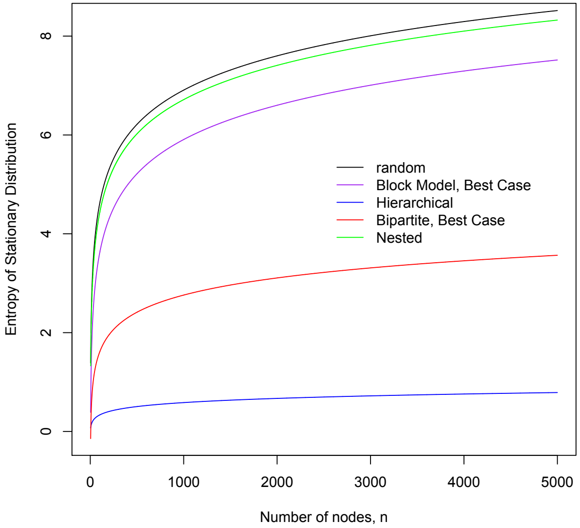
- allow for more variation among strategies across nodes and allow the strategies to evolve
- presence of multiple signals with different values
- incorporate a principled measure of cognitive costs / informational burden of paying attention to more neighbors
- extend to other systems
- real network data

Measure of Informational Burden

$w_{ij} =$ weight of edge from $i \rightarrow j$, $\sum_j w_{ij} = 1$

$\implies$ the vector p such that $pW = p$ is the stationary distribution

$\beta(W) :=$ the entropy of the distribution p , $-\sum_i p_i \log(p_i)$



Entropy of Stationary Distribution
0
2
4
6
8
0
1000
2000
3000
4000
5000
Number of nodes, n
random
Block Model, Best Case
Hierarchical
Bipartite, Best Case
Nested

Future Directions

- real network data
- how to minimize informational burden while maintaining network structure
- how informational constraints affects network inference

Acknowledgements

Simon Levin
Naomi Leonard
George Young

H_2 norm

Consider a dynamical system $\dot{v}(t) = Av + Bw$ and $z = Cv$. Let Σ be the solution to the Lyapunov equation

$$A^T\Sigma + \Sigma A + BB^T = 0.$$

Then the H_2 norm[1] of the system is

$$\sqrt{\mathrm{Tr}(C\Sigma C^T)}.$$

For y, this means that Σ solves

$$-\overline{L}^T\Sigma - \Sigma\overline{L} + I = 0 \text{ and } H_2 = \sqrt{\mathrm{Tr}(\Sigma)}.$$

It can be shown[1] that $H(L) = \lim_{t\to\infty} E[\|y(t)\|] = H_2(L)$.

[1] Robustness of Noisy Consensus Dynamics with Directed Communication, Proceedings of the American Control Conference, George Forrest Young, Luca Scardovi, and Naomi Ehrich Leonard, 2010.