Information Aggregation in Social Networks

A. Jadbabaie*  P. Molavi**  A. Tahbaz-Salehi†

*Electrical and Systems Engineering, Computer and Information Science Operations and Information Management, UPenn

**Electrical and Systems Engineering, UPenn

†Decision, Risk, and Operations Columbia Graduate School of Business

Penn University of Pennsylvania
Motivating Example: Climate Change

General features:

- There is a *common* underlying state of the world.
Motivating Example: Climate Change

General features:

- There is a *common* underlying state of the world.
- Agents’ optimal choices depend on the realization of the state.
Motivating Example: Climate Change

General features:

- There is a *common* underlying state of the world.
- Agents’ optimal choices depend on the realization of the state.
- Agents have heterogeneous *information endowments.*
Motivating Example: Climate Change

General features:

- There is a *common* underlying state of the world.
- Agents’ optimal choices depend on the realization of the state.
- Agents have heterogeneous *information endowments*.
- Agents communicate their *opinions* with agents in their social neighborhood.

Other examples:
- Adoption of microfinance (Banerjee et al. (2012))
- Agricultural methods (Hagerstrand (1969) and Rogers (1983))
- Medical procedures
Motivating Example: Climate Change

General features:
- There is a *common* underlying state of the world.
- Agents’ optimal choices depend on the realization of the state.
- Agents have heterogeneous *information endowments*.
- Agents communicate their *opinions* with agents in their social neighborhood.

Other examples:
- adoption of microfinance (Banerjee *et al.* (2012))
- agricultural methods (Hagerstrand (1969) and Rogers (1983))
- medical procedures
The Problem

Research question

How do the network structure and agents’ information endowments determine the extent of information aggregation?

Our model: an extension of DeGroot’s learning model with
- continuous flow of new information
- heterogenous observations
- asymptotic agreement with the Bayesian benchmark
DeGroot Learning Models

DeGroot based models are tractable:

- DeGroot (1974)
- Tsitsiklis (1984)
- Jadbabaie, Lin, Morse (2003)
- Acemoglu, Ozdaglar, Parandeh-Gheibi (2010)
- Golub and Jackson (2010)
- Jadbabaie, Molavi, Sandroni, Tahbaz-Salehi (2012)
DeGroot Learning Models

DeGroot based models are tractable:

- DeGroot (1974)
- Tsitsiklis (1984)
- Jadbabaie, Lin, Morse (2003)
- Acemoglu, Ozdaglar, Parandeh-Gheibi (2010)
- Golub and Jackson (2010)
- Jadbabaie, Molavi, Sandroni, Tahbaz-Salehi (2012)

There is empirical evidence in favor of DeGroot models:

- Chandrasekhar, Larreguy, Xandri (2012)
Problem Setup

- Finite set of agents $\mathcal{N} = \{1, \ldots, n\}$. 
- Agents are organized in a social network. 
- $\mathcal{N}_i \subseteq \mathcal{N}$ is the set of neighbors of agent $i$. 
- Agent $i$ can observe the beliefs of her neighbors, i.e., $j \in \mathcal{N}_i$. 
- Finite set of conceivable states $\Theta$. 
- Agents have beliefs over $\Theta$ and repeatedly update them.
Timing of the Events

At time 0:

- A state $\theta^* \in \Theta$ is realized.
- Agent $i$ starts with the prior $\mu_{i,0}$ over $\Theta$. 
Timing of the Events

At time 0:
- A state $\theta^* \in \Theta$ is realized.
- Agent $i$ starts with the prior $\mu_{i,0}$ over $\Theta$.

At time $t$:
- Agent $i$ observes $\{\mu_{j,t-1}\}_{j \in N_i}$.
- Agent $i$ also privately observes signal $\omega_{i,t}$.
- Agents use the new information to update their beliefs:
  \[ \mu_{i,t} = f_i(\mu_{i,t-1}; \{\mu_{j,t-1}\}_{j \in N_i}; \omega_{i,t}) \]
Agents’ Private Observations

- Signal profile $\omega_t = (\omega_{1,t}, \ldots, \omega_{n,t})$ is generated at time $t$.
- $\omega_1, \omega_2, \omega_3, \ldots$ are i.i.d.
- $\omega_t \sim \ell(\cdot | \theta)$, given the realized state is $\theta \in \Theta$.
- $\omega_{i,t}$ belongs to the finite set $S_i$.
- The marginal of $\ell(\cdot | \theta)$ over $S_i$ is $\ell_i(\cdot | \theta)$. 
Information Content of Agents’ Observations

Definition

Agent $i$’s information for $\theta$ in favor of $\hat{\theta}$ when $\theta$ is realized:

\[ I_{i}^{\theta}(\hat{\theta}) \overset{\text{def}}{=} D_{KL}(\ell_{i}(\cdot|\theta)\|\ell_{i}(\cdot|\hat{\theta})) \geq 0. \]
### Definition

Agent $i$’s **information** for $\theta$ in favor of $\hat{\theta}$ when $\theta$ is realized:

$$I^\theta_i(\hat{\theta}) \overset{\text{def}}{=} D_{KL}(\ell_i(\cdot | \theta) \| \ell_i(\cdot | \hat{\theta})) \geq 0.$$ 

### Definition

The realized state $\theta^*$ is **identifiable** if $I^*_i(\theta) > 0$ for some $i$ and all $\theta \neq \theta^*$. 
The Model

Belief update

\[ \mu_{i,t} = a_{ii} \text{BU}(\mu_{i,t-1}; \omega_{i,t}) + \sum_{j \in \mathcal{N}_i} a_{ij} \mu_{j,t-1} \]
The Model

Belief update

\[ \mu_{i,t} = a_{ii} \text{BU}(\mu_{i,t-1}; \omega_{i,t}) + \sum_{j \in N_i} a_{ij} \mu_{j,t-1} \]

- Agents aggregate the beliefs of their neighbors linearly.
- Agents process their private signals in a Bayesian way.
- \( a_{ii} \) is the self-reliance of agent \( i \).
- \( a_{ij} \) is trust of agent \( i \) in agent \( j \).
Sufficient Conditions for Learning

**Proposition**

Assume

(a) the network is strongly connected;
(b) agents’ priors have full support;
(c) all the states are identifiable;
(d) agents have strictly positive self-reliance; i.e., \( a_{ii} > 0 \) for all \( i \).

Then all agents learn the realized state asymptotically almost surely; i.e.,

\[
\mu_{i,t}(\theta^*) \longrightarrow 1 \quad \forall i \in \mathcal{N}.
\]
Definition

Eigenvector centrality of agent $i$ is the normalized fixed point of

$$v_i = \sum_{j=1}^{n} a_{ij} v_j.$$
**Definition**

Eigenvector centrality of agent $i$ is the normalized fixed point of

$$v_i = \sum_{j=1}^{n} a_{ij} v_j.$$ 

**Proposition**

$e_t$: the total expected error at time $t$.

$$\lim_{t \to +\infty} \frac{1}{t} \log e_t = -r,$$

where

$$r \approx \min_{\theta} \min_{\hat{\theta} \neq \theta} \sum_{i \in \mathcal{N}} v_i a_{ii} I^\theta_i (\hat{\theta}).$$
\[
    r \approx \min_{\theta} \min_{\hat{\theta} \neq \theta} \sum_{i \in N} v_i a_{ii} \mathcal{I}_i^\theta(\hat{\theta})
\]

- \( \min_{\theta} \min_{\hat{\theta} \neq \theta} \mathcal{I}_i^\theta(\hat{\theta}) \) is the rate of learning for a Bayesian agent.
- \( 1 - a_{ii} \) is the reduction in the rate as a result of agent’s bias.
- \( v_i \) captures the effective attention that agent \( i \) receives.
Definition

- \( I_i \succeq I_j \) if
  \[
  I_i^\theta(\hat{\theta}) \geq I_j^\theta(\hat{\theta}) \quad \text{for all} \quad \theta, \hat{\theta} \in \Theta.
  \]

- \( I_i \succ I_j \) if \( I_j \succeq I_k \) and
  \[
  I_i^\theta(\hat{\theta}) > I_j^\theta(\hat{\theta}) \quad \text{for some} \quad \theta, \hat{\theta} \in \Theta.
  \]

\( \succeq \) captures a notion of dominance.

- \( I_i \) provides more information than \( I_j \prec I_i \).

- This is regardless of which state is realized.
**Definition**

An allocation rule is a bijective mapping

$$\sigma : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}.$$ 

**Question**

Given signals with information $$\{I_j\}_{j \in \mathcal{N}}$$, which allocation rule maximizes the rate of learning?
Proposition

Assume that \( \{I_j\}_{j \in \mathcal{N}} \) are fully ordered with respect to \( \succeq \). Then for any optimal allocation \( \sigma \),

\[
\begin{align*}
    a_{ii} v_i & \geq a_{jj} v_j & \Leftrightarrow & \quad I_{\sigma(i)} \succeq I_{\sigma(j)}, \\
    a_{ii} v_i & > a_{jj} v_j & \Leftrightarrow & \quad I_{\sigma(i)} \succ I_{\sigma(j)}.
\end{align*}
\]

This corresponds to positive assortative matching.
Corollary

Assume that \( a_{ii} = a \) for all \( i \) and for some \( m \leq n \),

\[
I_j = \begin{cases} 
    I & \text{for } j = 1, \ldots, m, \\
    0 & \text{for } j = m + 1, \ldots, n.
\end{cases}
\]

Then in any optimal allocation the \( m \) most central agents make the informative observations.
Corollary

Assume that $a_{ij} = a$ for all $i$ and for some $m \leq n$,

$$I_j = \begin{cases} 
I & \text{for } j = 1, \ldots, m, \\
0 & \text{for } j = m + 1, \ldots, n.
\end{cases}$$

Then in any optimal allocation the $m$ most central agents make the informative observations.

Corollary

If $|\Theta| = 2$, in any optimal allocation the most central agents observe the most informative signals.
Corollary

Assume that \( a_{ij} = a \) for all \( i \) and for some \( m \leq n \),

\[
I_j = \begin{cases} 
I & \text{for } j = 1, \ldots, m, \\
0 & \text{for } j = m + 1, \ldots, n.
\end{cases}
\]

Then in any optimal allocation the \( m \) most central agents make the informative observations.

Corollary

If \( |\Theta| = 2 \), in any optimal allocation the most central agents observe the most informative signals.

These are consistent with the empirical observation of Banerjee, Chandrasekhar, Duflo, & Jackson (2012).
The information ordering which leads to positive assortative matching is very strong (dominance).

Other orderings could lead to dramatically different results.

In fact, the optimal allocation could correspond to the least central agent making the “best” observations.
An Example (1/4)

- \( a_{ii} = a \)
- \( \Theta = \{\theta_0, \theta_1, \ldots, \theta_n\} \)
- \( S_i = \{H, T\} \)

\[
\ell_j(\cdot|\cdot) : \begin{pmatrix}
\theta_0 & \begin{pmatrix} 1 - \alpha_j & \alpha_j \end{pmatrix} \\
\theta_1 & \begin{pmatrix} 1 - \alpha_j & \alpha_j \end{pmatrix} \\
\vdots & \vdots \\
\theta_j & \begin{pmatrix} \alpha_j & 1 - \alpha_j \end{pmatrix} \\
\vdots & \vdots \\
\theta_n & \begin{pmatrix} 1 - \alpha_j & \alpha_j \end{pmatrix}
\end{pmatrix}
\]
An Example (2/4)

\[ r \propto \min_{k \neq 0} \sum_{i=1}^{n} v_i I_i^0(\theta_k) \]

\[ I_j^0(\cdot) : \begin{cases} j = 1 & \begin{pmatrix} \theta_1 & \theta_2 & \ldots & \theta_n \end{pmatrix} \\
 j = 2 & \begin{pmatrix} l_1 & 0 & \ldots & 0 \\
 0 & l_2 & \ldots & 0 \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \ldots & l_n \end{pmatrix} \\
 j = n & \begin{pmatrix} \theta_1 & \theta_2 & \ldots & \theta_n \end{pmatrix} \end{cases} \]

- Agent with signal \( j \) is “expert” in state \( j \) and knows nothing about other states.
- Nobody is expert in state 0.
Lemma

Given the allocation rule $\sigma$, the rate of learning is proportional to

$$\min_i \nu_i l_{\sigma(i)}.$$

The rate of learning is determined by the agent with the smallest centrality $\times$ expertise.
Proposition

For any optimal allocation $\sigma$, 

$$v_i \geq v_j \iff I_{\sigma}(i) \leq I_{\sigma}(j),$$
$$v_i > v_j \iff I_{\sigma}(i) < I_{\sigma}(j).$$
Proposition

For any optimal allocation $\sigma$,

\[ v_i \geq v_j \Leftrightarrow I_{\sigma}(i) \leq I_{\sigma}(j), \]
\[ v_i > v_j \Leftrightarrow I_{\sigma}(i) < I_{\sigma}(j). \]

- In any optimal allocation, the least central agent has the highest expertise.
- There needs to be a mismatch between the attention and information bottlenecks.

This corresponds to negative assortative matching.
The interplay of information and network structures determine the extent of information aggregation.

When the information content of signals are fully ordered according to a notion of dominance, optimal allocation of signals entails central agents to receive better signals.

Using different notions of information ordering can result in significantly different conclusion.

In the optimal allocation, the least central agents could make the best observations.

Future research:
- a more complete characterization of the rate of learning