Using Gossips to Spread Information: Theory and Evidence from Two RCTs
Banerjee, A., Chandrasekhar, A. G., Duflo, E. and Jackson, M. O.

by Balazs Kertesz

Central European University

October 30, 2018
Motivation

Policy and business relevance in *how to start spreading information efficiently* in networks

*Introduce:* Gmail, Microcredit and new agricultural products

*Seeds:* Leading bloggers, leading farmers, based on demographic or occupation characteristics or single officer (might not be even local)

*Limitation:* Can fail

Theory predicts

Central persons should be seeded according to an appropriate measure

*Limitation:* Expensive to discover network structure

Question

How to identify highly central persons cheaply without network structure?

*Idea:* Ask few individuals in community which other members would be best to spread info

*Despite:* Little knowledge about neighbors’ neighbors
Motivation

Experimental Evidence
  Phone or Cash RCT
  Immunization RCT

Theory

Evidence Pro Theory

Conclusion
Experimental Evidence
Phone or Cash RCT

Goal
Spread info about a non-rival promotion of a cell phone or cash lottery among on average 200 HHs/village

Task
Initiate a missed call to win lottery, \{\{50 : 25 : 275, Phone\}, P_{\sum \text{two dice rolls}}\}

Treatment Conditions
In each 71 village/TC up to 15 HHs interviewed in order to gather recommendations about possible seeds

<table>
<thead>
<tr>
<th>random</th>
<th>3/5 totally random seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>elder</td>
<td>3/5 seeds among well-respected village elders randomly</td>
</tr>
<tr>
<td></td>
<td>&quot;Who is a well-respected elder in the village?&quot;</td>
</tr>
<tr>
<td>gossip</td>
<td>3/5 seeds among gossip nominees randomly</td>
</tr>
<tr>
<td></td>
<td>&quot;Whom should we speak to about spreading information in the village?&quot;</td>
</tr>
</tbody>
</table>
Experimental Evidence

Results

\# Calls = \theta_0 + \theta_1 D_{Gossip(Hit)} + \theta_2 D_{Elder(Hit)} + \theta_3 \# Seeds + \theta_4 \# Gossip + \theta_5 \# Elder + \varepsilon

<table>
<thead>
<tr>
<th></th>
<th>(OLS1)</th>
<th>(OLS2)</th>
<th>(IV1)</th>
<th>(OLS3)</th>
<th>(OLS4)</th>
<th>(IV2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># Calls</td>
<td># Calls</td>
<td># Calls</td>
<td># Calls/# Seeds</td>
<td># Calls/# Seeds</td>
<td># Calls/# Seeds</td>
</tr>
<tr>
<td>(D_{Gossip})</td>
<td>3.651</td>
<td>1.053</td>
<td>1.053</td>
<td>0.952</td>
<td>1.979</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.786)</td>
<td>(0.698)</td>
<td>(0.518)</td>
<td>(0.501)</td>
<td>(1.071)</td>
<td></td>
</tr>
<tr>
<td>(D_{Elder})</td>
<td>-1.219</td>
<td>-0.116</td>
<td>-0.116</td>
<td>0.309</td>
<td>-0.677</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.053)</td>
<td>(0.518)</td>
<td>(0.511)</td>
<td>(0.588)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(D_{GossipHit})</td>
<td>3.786</td>
<td>7.436</td>
<td>0.952</td>
<td>1.979</td>
<td>1.979</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.858)</td>
<td>(4.266)</td>
<td>(0.501)</td>
<td>(1.071)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(D_{ElderHit})</td>
<td>0.792</td>
<td>-3.475</td>
<td>0.309</td>
<td>-0.677</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.056)</td>
<td>(2.259)</td>
<td>(0.511)</td>
<td>(0.588)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># of Obs.</td>
<td>212</td>
<td>212</td>
<td>212</td>
<td>212</td>
<td>212</td>
<td>212</td>
</tr>
<tr>
<td>E(Control)</td>
<td>8.077</td>
<td>5.846</td>
<td>5.805</td>
<td>1.967</td>
<td>1.451</td>
<td>1.317</td>
</tr>
</tbody>
</table>

⇒ 40+% gossip treatment effect
Goal
Improve immunization rates via spreading info about monthly immunization camps

Treatment Conditions
500+ villages tracked over 12 months
random 6 totally random seeds
  gossip 17 HHs asked, 6 gossip seeds
  most nominations received
trusted 17 HHs asked, 6 trusted seeds
  most nominations received
trusted gossip combined Q, same method

\[
\# \text{Ch. rec. vacc.} = \theta_0 + \theta_1 D_{\text{Gossip}} + \theta_2 D_{\text{Trusted}} + \theta_3 D_{\text{Tr. Goss.}} + \theta_4 \text{Incentives} + FE + \varepsilon
\]

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># Ch. rec. measles</td>
<td># Ch. att. session</td>
</tr>
<tr>
<td>(D_{\text{Gossip}})</td>
<td>1.078</td>
<td>4.903</td>
</tr>
<tr>
<td></td>
<td>(0.500)</td>
<td>(2.503)</td>
</tr>
<tr>
<td>(D_{\text{Trusted}})</td>
<td>0.439</td>
<td>1.849</td>
</tr>
<tr>
<td></td>
<td>(0.408)</td>
<td>(2.047)</td>
</tr>
<tr>
<td>(D_{\text{Tr. Goss.}})</td>
<td>0.444</td>
<td>2.379</td>
</tr>
<tr>
<td></td>
<td>(0.376)</td>
<td>(1.917)</td>
</tr>
<tr>
<td># of Obs.</td>
<td>6697</td>
<td>6712</td>
</tr>
<tr>
<td>E(Control)</td>
<td>4.31</td>
<td>18.11</td>
</tr>
</tbody>
</table>

\(\Rightarrow\) 25+% gossip treatment effect
**Question**

How is it possible ppl nominate right seeds in terms of info diffusion even if they are uncertain about the network structure?

**Model**

Strongly connected network $w \in [0, 1]^{n \times n}$

Largest eigenvalue $\lambda$, eigenvector $v \in \mathbb{R}^n$

Time allowed to spread information $T \in \mathbb{N}$

**Concepts**

Diffusion centrality (DC) & network gossip (NG)
Message = (info, source tag)

Diffusion Centrality (Sender’s Perspective, ability to send info)

$$DC(w, T) = H(w, T) \cdot 1 = \left( \sum_{t=1}^{T} w^t \right) \cdot 1 \quad (n \times 1)$$

Intuition: $E$(# of times info originated from row is heard by anyone)

Network Gossip (Receiver’s Perspective, tracks reception of info)

$$NG(w, T)_j = H(w, T)_j \quad (n \times 1)$$

Intuition: $E$(# of times $j$ hears info originated from row)
Theory

Results

Result

Members of society can estimate who is diffusion central w/o knowing network structure

Proposition

If $T = 1 \Rightarrow \text{DC} \propto \text{degree centrality}$
If $T \to \infty \ & \ & \lambda < 1 \Rightarrow \text{DC} \equiv \text{Katz-Bonacich centrality}$
If $T \to \infty \ & \ & \lambda > 1 \Rightarrow \text{DC} \to \text{eigenvector centrality}$

Theorem 1

$$\sum_j \text{cov}(\text{DC}(w, T), \text{NG}(w, T)_j) = \text{var}(\text{DC}(w, T)) > 0$$

Theorem 2

$\forall j$ ranking of others under $\text{NG}(w, T)_j \overset{\propto}{\to} \text{DC}(w, T)$ as $T \to \infty$, provided $\lambda > 1$ and $w$ aperiodic.

Intuition

$j$ hears exponentially more often about diffusion (eigenvector) central nodes
Evidence Supporting the Theory

Discrete choice model:

\[ u_i(j) = \alpha + \beta x_j + \gamma z_j + \mu_v + \epsilon_{ijv}, \]

where

- \( x \) – centrality measures
- \( z \) – demographic characteristics
- \( \mu_v \) – village FE

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion</td>
<td>0.61</td>
<td>0.64</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.13)</td>
<td>(0.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>0.46</td>
<td>-0.04</td>
<td>-0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigenvector</td>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.09)</td>
<td>(0.19)</td>
</tr>
</tbody>
</table>
Conclusion

Take-aways
Small scale exploration of a network can be efficient in diffusing information
Simply by asking members who they deem the researcher should seed first
Despite having imperfect knowledge about own network
Without unnecessary data collection
Sizable effect

Limitations
Pure info transmission, no trust-based seeding strategies
Findings may not scale up