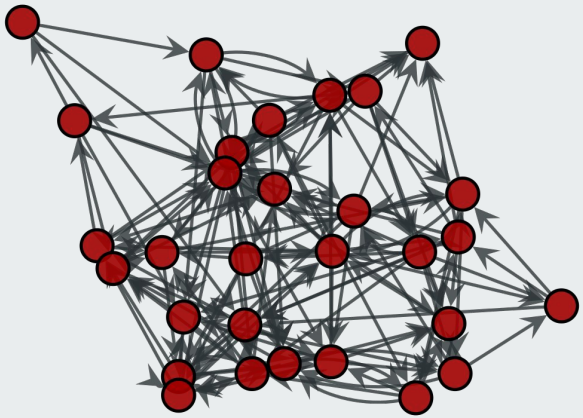
Truth Learning in Social and Adversarial Settings

By Amanda Wang, Julia Krizanova, Filip Uradnik, and Rhett Olson
Mentor: Professor Jie Gao

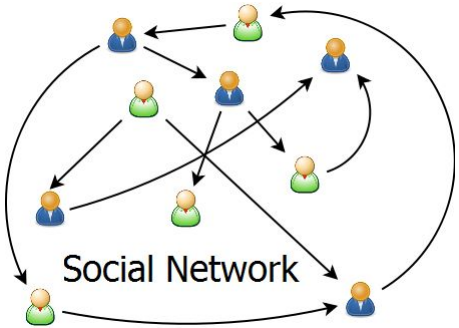


Motivation

- You have imperfect information about the world
- You are influenced by other people, also with imperfect information
- Examples:
  - Social media
  - Panel Discussions

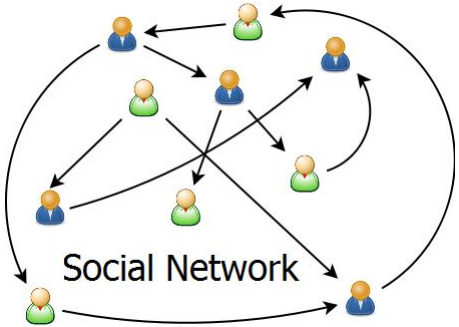
Setup

- Social network structured as a graph.



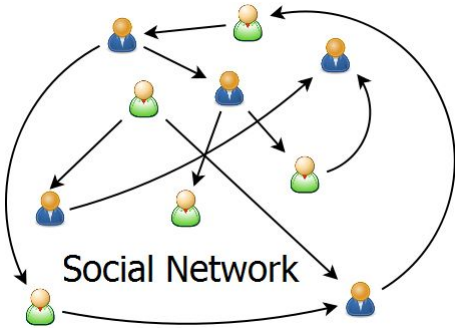
Setup

- Social network structured as a graph.
  - Vertices = agents.
  - Edges = which agents know each other.

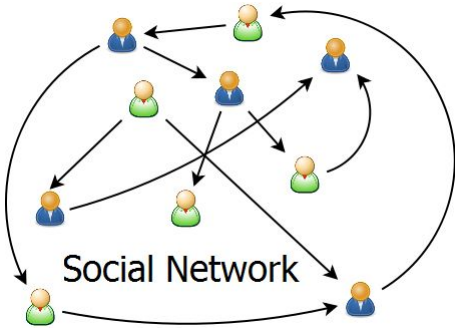


Setup

- Social network structured as a graph.
  - Vertices = agents.
  - Edges = which agents know each other.
- Binary ground truth.

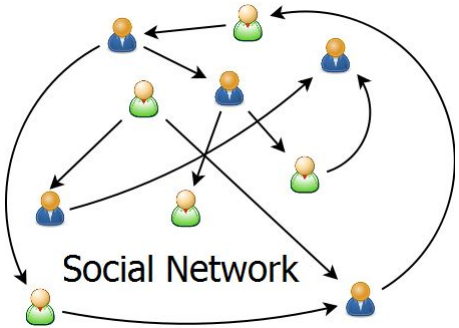


Setup



- Social network structured as a graph.
  - Vertices = agents.
  - Edges = which agents know each other.
- Binary ground truth.
- Agents have independent, private, and *noisy* measurements of the ground truth.
- Agents make predictions of ground truth sequentially.
  - based on private measurement & neighbors' predictions.

Setup



- Social network structured as a graph.
  - Vertices = agents.
  - Edges = which agents know each other.
- Binary ground truth.
- Agents have independent, private, and *noisy* measurements of the ground truth.
- Agents make predictions of ground truth sequentially.
  - based on private measurement & neighbors' predictions.

Q: *When is it possible for agents to learn a ground truth given a network topology, prediction order, and distribution of private measurements?*

Information Cascades/Herding

Agents can become biased if one opinion appears to dominate among those it has seen.

➔ Can lead nearly the entire group to stop using their own observations and instead copy previous decisions.

Cascade example

An urn contains either 1 blue + 2 red balls, or 1 red + 2 blue balls w/ equal probability.

Taking turns, each person:

- randomly picks a ball to observe in private (with replacement)
- publicly states if they think the urn is majority red or majority blue

Cascade example

Suppose 2 blue + 1 red.

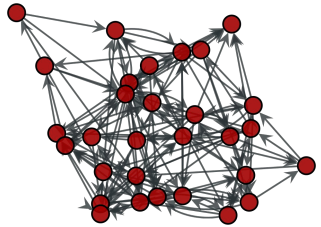
1. Person 1 observes red. Announces “majority red”.
2. Person 2 observes red. Announces “majority red”.

---

3. Person 3 observes blue. Announces “majority red”. **No new information!**
4. Person 4 observes xx. Announces “majority red”.

…

Directions



- Complexity: Decide for a given network whether truth learning can happen.
  - Is this NP-hard?
- Voting: Non-binary ground truth.
  - No “good” way of aggregating non-binary agent preferences.
  - Condorcet Paradox, Arrow’s Impossibility Theorem
- Adversaries: How can adversarial agents affect the outcome?
  - How does this depend on the network structure?
  - How can the remaining agents protect against this?

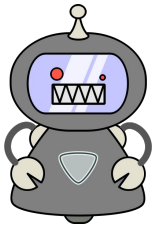
$$A \succ_1 B \succ_1 C$$

$$B \succ_2 C \succ_2 A$$

$$C \succ_3 A \succ_3 B$$

$$\implies A \succ B \succ C \succ A$$

Condorcet Paradox



Acknowledgements

This work made possible by the Rutgers DIMACS REU program. Thank you to faculty and staff who work to keep the program running. Thank you as well to the National Science Foundation for funding this project through the grant CNS-2150186 and the REU supplement to NSF 2208663 -Collaborative Research: AF: Small: Promoting Social Learning Amid Interference in the Age of Social Media.

This work was carried out while the authors Filip Uradnik and Julia Krizanova were participants in the 2024 DIMACS REU program at Rutgers University, CNS-2150186. The authors were supported by CoSP, a project funded by European Union's Horizon 2020 research and innovation programme, grant agreement No. 823748

Thank you as well to Professor Jie Gao for her help and leadership on this project.