

The Mathematics of Misinformation

Philip Cate, Math 498 Math in Social Context, Spring 2022

The Problem

The spread of false or misleading information online is destructive to our society and democracy.

From a mathematical perspective, how do social media networks and algorithms enable the spread of misinformation on the internet, and how can it be fixed?

Background

- Misinformation spread had a significant impact on 2016 & 2020 elections
- Factors that influence spread: Information overload, Time limitation (limited attention span), and Network structure ("superconnected" and "echo chamber" clusters as opposed to random connections)
- It's the system, not the viral nature of information, that causes spread
- How can math & statistical modeling be used to curb the spread of misinformation?

References

[Science Daily](#)

[PNAS](#)

[American Univ \(via Scientific American\)](#)

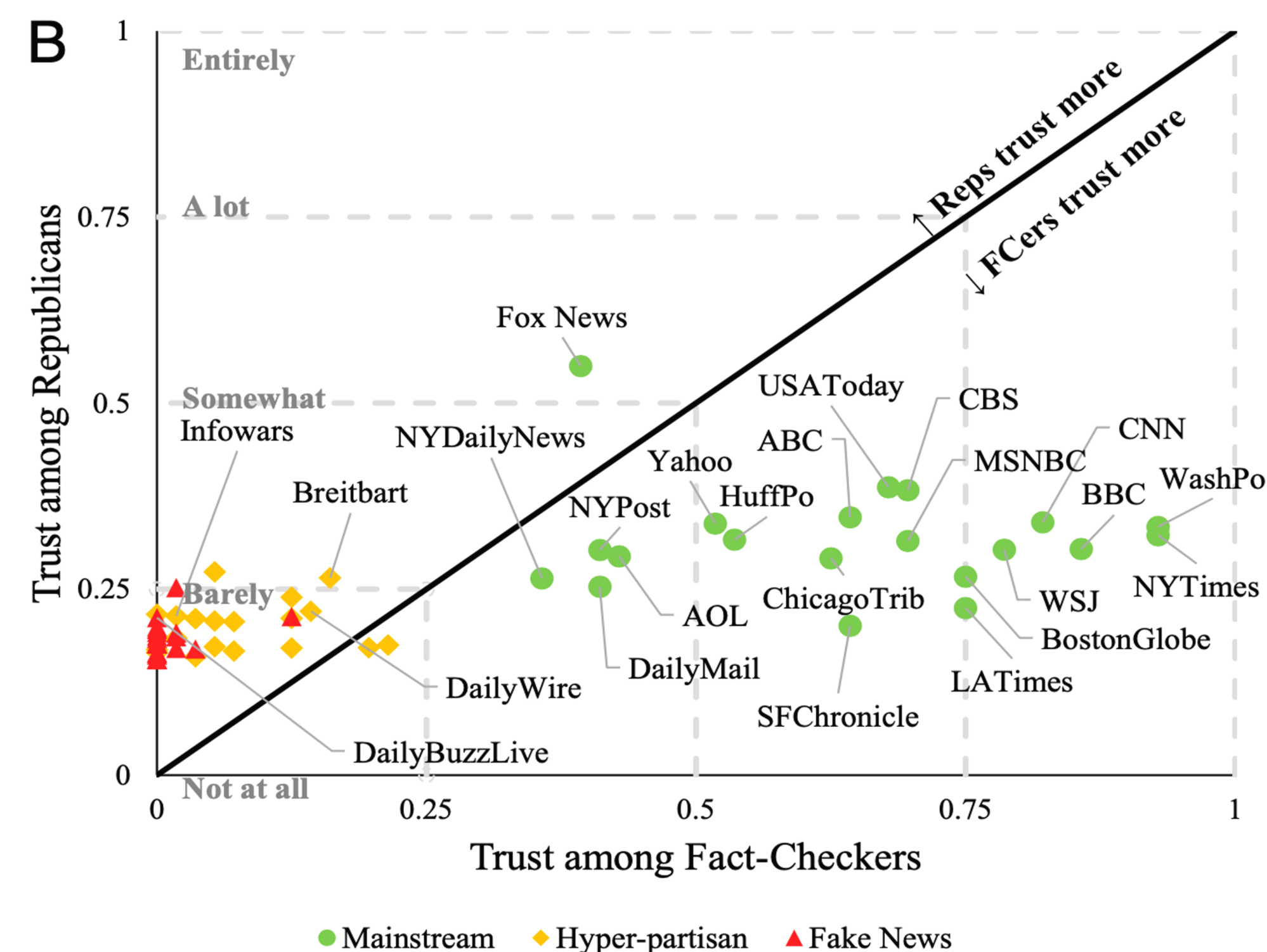


Figure 1: Comparing trust of outlets between Republicans & fact checkers. In the aggregate (including Dems) there is a 0.90 correlation in trust, suggesting crowd-sourcing can be effective in prioritizing trustworthy news outlets (PNAS)

Modeling

Adopting disease models to track and understand misinformation spread:

- Agent-based model (dots represent individuals, connected via lines)
- Using proxies for information load (probability of posting) & attention span (items viewed before sharing), simulate information spread under various conditions

Criticism: human behavior is complicated to model, and information is not a virus!

Solutions

- 1) Using crowdsourced judgments of news source quality (PNAS): partisan differences in determining trust, but strong correlation with certified fact-checkers; difficult in practice since outlets & networks seek profit; issues with familiarity/picking users
- 2) Machine learning (American Univ.): using statistical modeling to classify misinformation through linguistic analysis; achieved 90% accuracy on 112 tweets

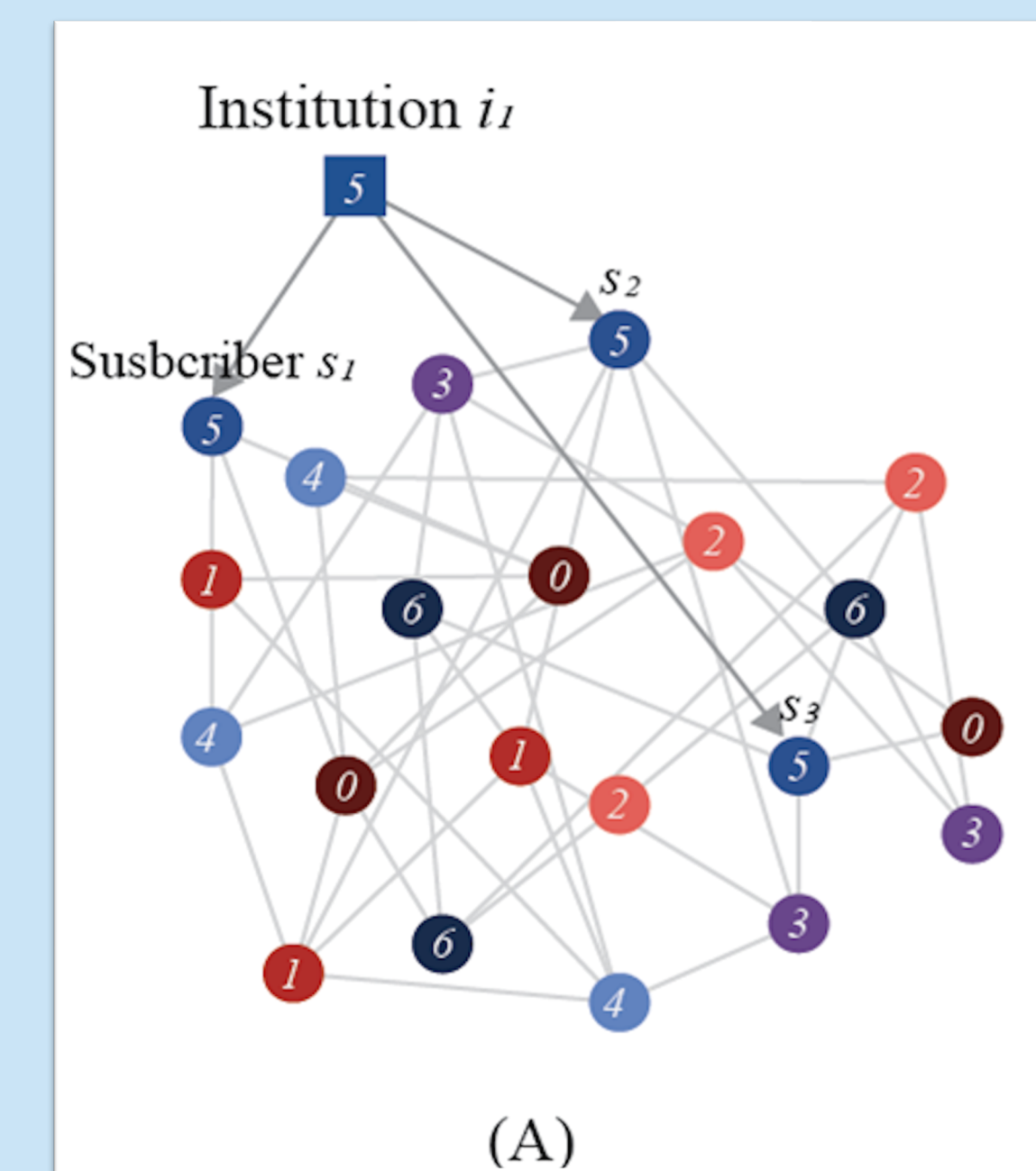


Figure 2: A diagram that exemplifies an agent-based model, demonstrating the various connections that can cause the spread of misinformation.

The Big Picture

Relates to our discussions on policing algorithms: how can info classification & elimination be equitable?