Welfare-guided Seeding with Time-constrained Data Collection

Carlos Hurtado* (Pitt), M. Amin Rahimian* (Pitt), Md Sanzeed Anwar (UMich),& Dean Eckles(MIT).
*{cah259,rahimian@pitt.edu}

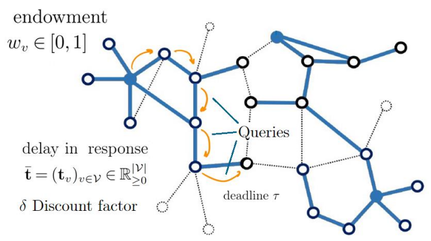
Motivation

Existing works on fair influence maximization are often reliant on one or more simplifying assumptions:

- Network structure is fully known.
- Utility received from influencing a node does not depend on time.
- Individuals can be easily categorized by groups to define a measure of inequality.

Project Description

We propose a seeding algorithm that operates with a network query phase bounded by some deadline τ and allows us to set the level of prioritization of the influence over nodes with the low prior endowment. Our theoretical performance guarantees demonstrate how the quality of the outcome depends on measurement resources such as sample size and time spent in data collection

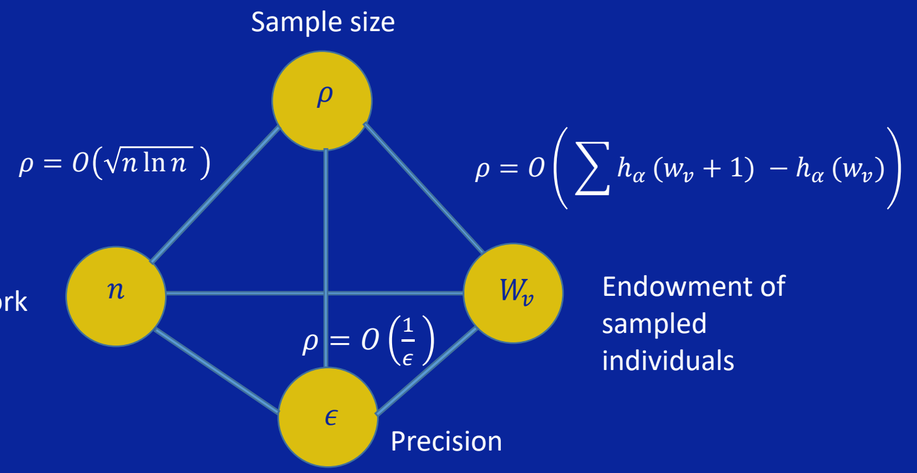


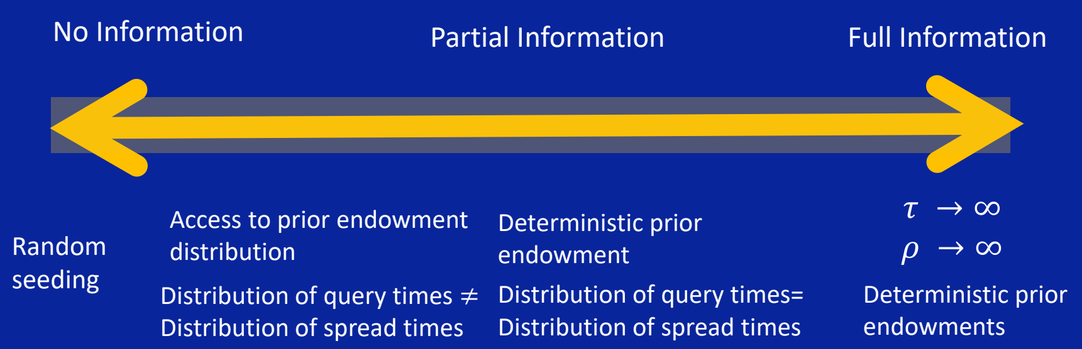
Context

- Previous work addresses problem of seeding with partial network information.
- Several fairness schemes has been proposed to address fair allocation in the influence maximization problem:
  1. Max-min fairness
  2. Disparity in utility across groups
  3. Combining fairness and efficiency through a welfare-theoretic perspective
- There exist works where time constrained optimization is considered only in the spread phase.

University of Pittsburgh

We propose an influence model that captures the time-sensitive nature of real-world influence maximization problems subject to fairness schemes that can be adapted to different degrees of **aversion to inequality.**





Main Contributions

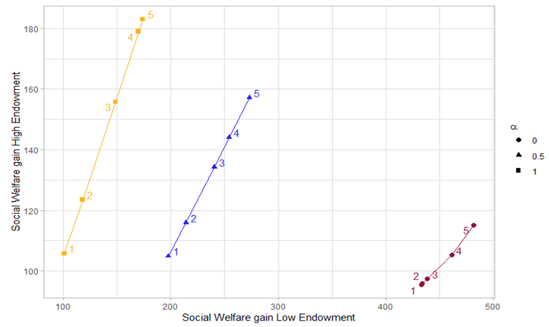
- We provided theoretical guarantees on sample size and seeding performance in scenarios where the data collection can either account with certainty for the time delay distribution or when the time delay distribution is partially known with or without the knowledge of prior endowment of the individuals.
- We provided insights into asymptotic behavior controlling variables such as time constraint in data collection, sample size, and aversion to inequality in the social welfare gain.

Numerical Experiments

Objective function:

$$f(S) = \mathbb{E}_{\mathcal{G},\mathcal{T}}\left[\sum_{v\in\mathcal{V}} h_\alpha\left(w_v + \delta^{\mathbf{t}_{v|S}}\right)\right]$$

$$h_\alpha(\cdot) = \begin{cases} (\cdot)^\alpha/\alpha, \text{ if } 0 < \alpha \leq 1, & \text{Pure utilitarian criterion } \alpha = 1 \\ \log(\cdot), \text{ if } \alpha = 0 & \text{Nash bargaining solution} \end{cases}$$



We test our algorithm on the Discovery of Higgs boson re-tweeting track data set. We test over different time deadline settings {1,2,3,4,5} and for different aversion to inequality {0,0.5,1} where 0 tends to benefit the lowest endowed individuals. The threshold for low and high endowed individuals are 0.25 and 0.75, respectively. Additionally, correlation between node degree and prior endowment was set at 0.

Potential Impact

- Public policies aimed to reduce structural inequalities
- Advertising job training/openings
- Time-critical viral marketing
- Financial assistance programs

References/ Acknowledgement

- Eckles, D., H. Esfandiari, E. Mossel, and M. A. Rahimian. 2022. "Seeding with Costly Network Information." *Operations Research*, forthcoming. https://doi.org/10.1287/opre.2022.2290

Rahimian acknowledges support from Pitt Momentum funds.