Responsible Data Collection & Intervention Design

Amin Rahimian
Department of Industrial Engineering
University of Pittsburgh

Motivation

- Social networks are very efficient for advertising job openings and other economic opportunities, as well as public health and educational outreach but collecting data about social networks poses a critical bottleneck.
- It is important to note the vulnerability of individuals to being left out of the collected data or being unduly exposed in the collected data.
- Carefully optimized designs (even though they have a broad reach) can amplify the existing inequalities that shape the structure of a social network.

Project Description

- We address the challenges of reliable and responsible social network algorithm design through the compound lens of algorithmic fairness, differential privacy, and adversarial robustness.
- We develop a welfare-guided, privacy-aware and adversarially robust codesign framework for data collection and algorithmic intervention in large-scale, limited-information and resource-constrained environments of complex sociotechnical systems.

Context

- Public health professionals collect social network data to engage with vulnerable populations for improving effectiveness of HIV and suicide prevention programs.
- These community relations can be harmed if public health efforts pose excessive privacy risks to the individuals or fail to effectively address historical and structural inequalities in the population.

University of Pittsburgh

Developing **socially responsible** approaches for **data collection** is essential for designing **sustainable** public health and development **interventions**.

Project Deliverables

- Manuscripts on algorithmic fairness and privacy guarantees of network interventions
- NSF proposal for developing new data collection and intervention methodologies based on the preliminary results

Potential Impact

- This proposal adopts a welfare-guided perspective to social network data collection and intervention design, giving policy makers explicit choices to balance the net benefits of a design against its unequal impacts and privacy risks.
- It can lead to new methods for collecting social network data that are privacy-preserving and provide adequately high resolution to ensure intervention effects reach all communities.

Reference

- Eckles, D., H. Esfandiari, E. Mossel, and M. A. Rahimian. 2022. “Seeding with Costly Network Information.” *Operations Research*, forthcoming. https://doi.org/10.1145/3328526.3329651