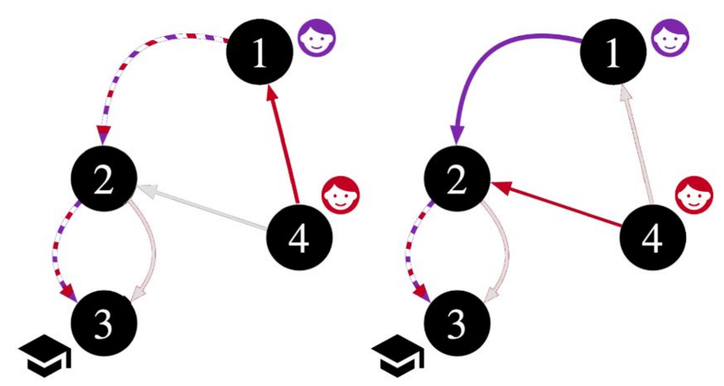
GAEA: Graph Augmentation for Equitable Access via Reinforcement Learning

Govardana Sachithanandam Ramachandran, Ivan Brugere, Lav R. Varshney, Caiming Xiong

Overview

- **Motivation**: Disparate access to resources by different subpopulations is a prevalent issue in societal and sociotechnical networks
- **Method**: Budgeted discrete edge augmentation to improve equity
- **Challenge**: We prove GAEA is the class of NP-hard, and cannot be approximated within a factor of $(1 - \frac{1}{3e})$

Toy Schematic of disparate groups access to resource



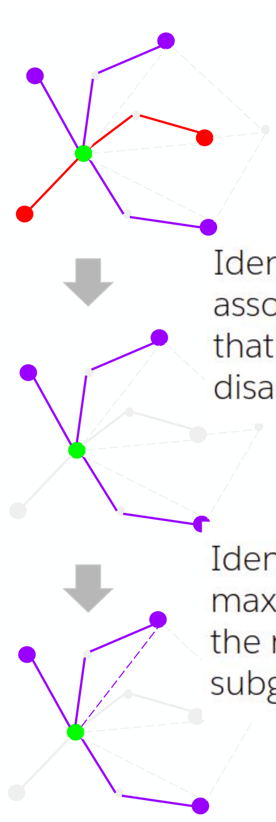
Before

After equitable graph augmentation

- Purple and Red are two disparate population distributed spatially different
- Individuals from each group traverse along the edges of their respective colors
- Red group is disadvantageous in accessing resources, compared to Purple group as they need to travel to further to access resource node#3
- Gray arrows represent, edges that are editable

Method#1 Greedy Equitable Centrality Improvement

- Two Groups are considerable, when their corresponding sub-graph have similar resource node centrality
- Hence discrete graph augmentation for equitable access can be reduced to:
  1. Unconstrained: Maximizing centrality of the resource nodes
  2. Constraints: Matching centrality across group

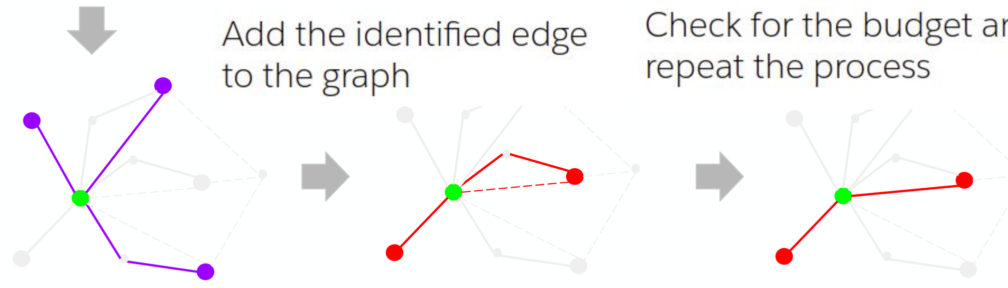


Identify the sub-graph associated with group that's most disadvantageous

Identify an edge that'd maximize the certality of the resource in the subgraph

Add the identified edge to the graph

Check for the budget and repeat the process



Method 2: Greedy equitable centrality improvement

Input: A directed graph $\mathcal{G} = (V, E)$, neighborhood function N_g, and budget B

$E^u := \emptyset$

for $b = 1, 2, \ldots, B$ **do**
  $E^e := E \cup E^u$
  $g_{\min} := \text{argmin}\{U_g(E^e) | g \in G)\}$
  for $u \in V | u \in N_{g_{min}}(E^e)$ **do**
    for $v \in V | v \in E^e$ **do**
      Compute $U_{g_{\min}}(E^e \cup \{(u,v)\})$
  $u_{\max}, v_{\max} := \text{argmax}\{U_{g_{\min}}(E^e \cup \{(u,v)\})\}$
  $E^u := E^u \cup \{(u_{\max}, v_{\max})\}$
return E^u

Method#2 Equitable Mechanism Design in MRP

- Unlike classical RL, where to objective is to optimize for policy. Here the human's are the agents whose policy may not be changeable.
- Hence we design as Mechanism design problem where the Dynamics is augmented to reduce inequity among groups and improve overall utility
- Given:
  - Graph $\mathcal{G} = (V, E)$
  - Set of disparate Groups $G = \{g_1, \ldots, g_k\}$
  - Set of resource nodes R
- We define: A particle p_g as an instance of $g \in G$ spawned at node s_0 as per node distribution of $\mu(g)$
- Value function of the particle p_g, is:

$$v_g(s_0) = \sum_{t=0}^{T-1} \gamma^t R P^t s_0$$

- Value function for the group, g, is:

$$V^g = \mathbb{E}_{s_0 \sim \mu(g)} [v^g(s_0)]$$

- We parameterize the transition probability as

$$P = D^{-1} E^e$$

$$E^e = E + A \odot E^u$$

$$D(i,i) = \sum_j E_{(i,j)}$$

- $A \in \{0,1\}^{|S| \times |S|}$ mask adjacency matrix that restrict the candidate edges for edit
- Continuous relaxation of edge edits using Gumbel sigmoid

$$E^u(i,j) = \frac{1}{(1 + \exp(-(\phi(\vec{0}) + g_i)/\tau)}, \forall i,j \in S$$

- The problem objective framed as MRP optimization

$$E^u = \underset{\substack{s.t. \sum_{g \in G} |V_g - \overline{V}_G| = 0 \\ \sum_{g \in G} \|E^u\|_0 < B}}{\text{argmax}} \sum_{g \in G} V_g$$

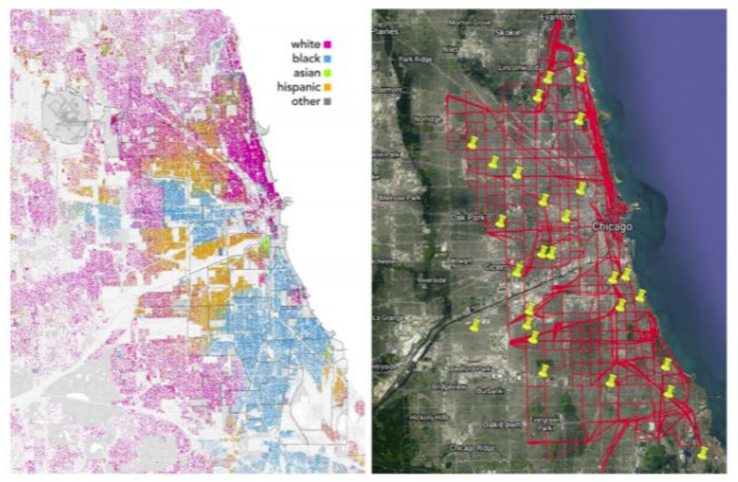
- Un-constrained objective using augmented lagrangian.

$$J = -\min_{E^u} \sum_{g \in G} V^g - \mu_1 \Big(\sum_{g \in G} V_g - \overline{V_G}\Big)^2 - \mu_2 \Big(\min(0, \sum_{g \in G} \|E^u\|_0 - B)\Big)^2 - \lambda_1 \Big(\sum_{g \in G} |V_g - \overline{V_G}|\Big) - \lambda_2 \Big(\min(0, \sum_{g \in G} \|E^u\|_0 - B)\Big).$$

$$\lambda_1^{new} \leftarrow \lambda_1^{old} + \mu_1 \Big(\sum_{g \in G} |V_g - \overline{V}_G|\Big),$$

$$\lambda_2^{new} \leftarrow \lambda_2^{old} + \mu_2 \Big(\max(0, \sum_{g \in G} \|E^u\|_0 - B)\Big)$$

Exp: Infrastructure Network - Equitable School Access in Chicago



(a) Chicago demographics by race/ethnicity (b) Chicago transit network and school locations

- Chicago demographics and infrastructure
- (a) shows demographics, demonstrating highly segregated areas of the city by race and ethnicity
- (b) shows a transit network (**red**) we collected for this work, induced from Chicago Transit Authority bus routes.
- (b)also show (yellow) the location of schools within our dataset from the Chicago Public Schools.

- Nodes # 2011
- Edges # 7984
- Resource Nodes: schools over the 90th percentile
- Budget : 100 edges
- Data:
  - https://www.transitchicago.com/data/
  - https://cps.edu/SchoolData/

Chicago Schools	Initial	GECI	EMD-MRP
Avg. distance	6.85	6.80	3.67
Var. b/w groups	0.131	0.102	0.033

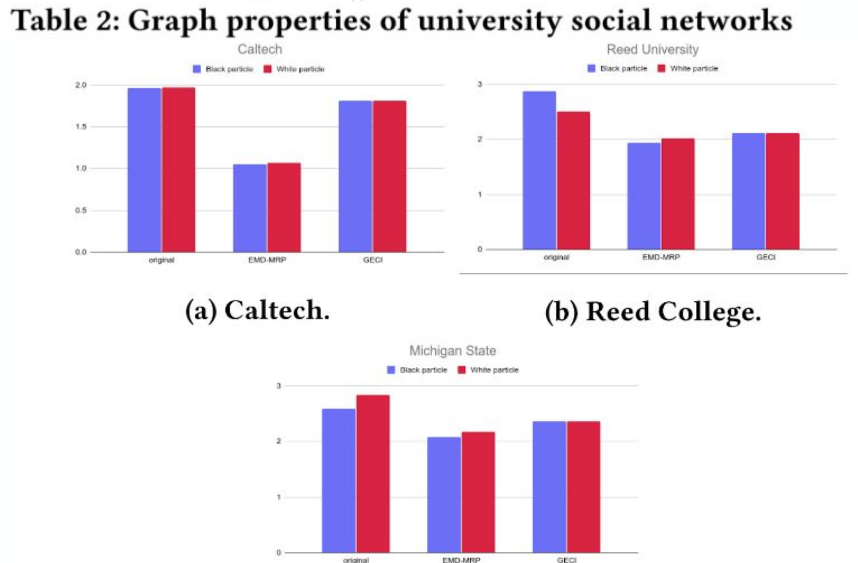
Table 1: South Chicago Public School with budget 100

Exp: Social Network

- Social networks within universities and organizations may enable certain groups to more easily access people with valuable information or influence
- On Facebook100 dataset, We define popular seniors as the reward nodes and the objective is for freshmen of both genders to have equitable access to these influential nodes

	Caltech	Reed	Mich. State
num. nodes, $\|V\|$	770	963	3749
num. edges, $\|E\|$	33312	37624	163806
num. editable edges $\|A\|$	336597	474439	3958660

Table 2: Graph properties of university social networks

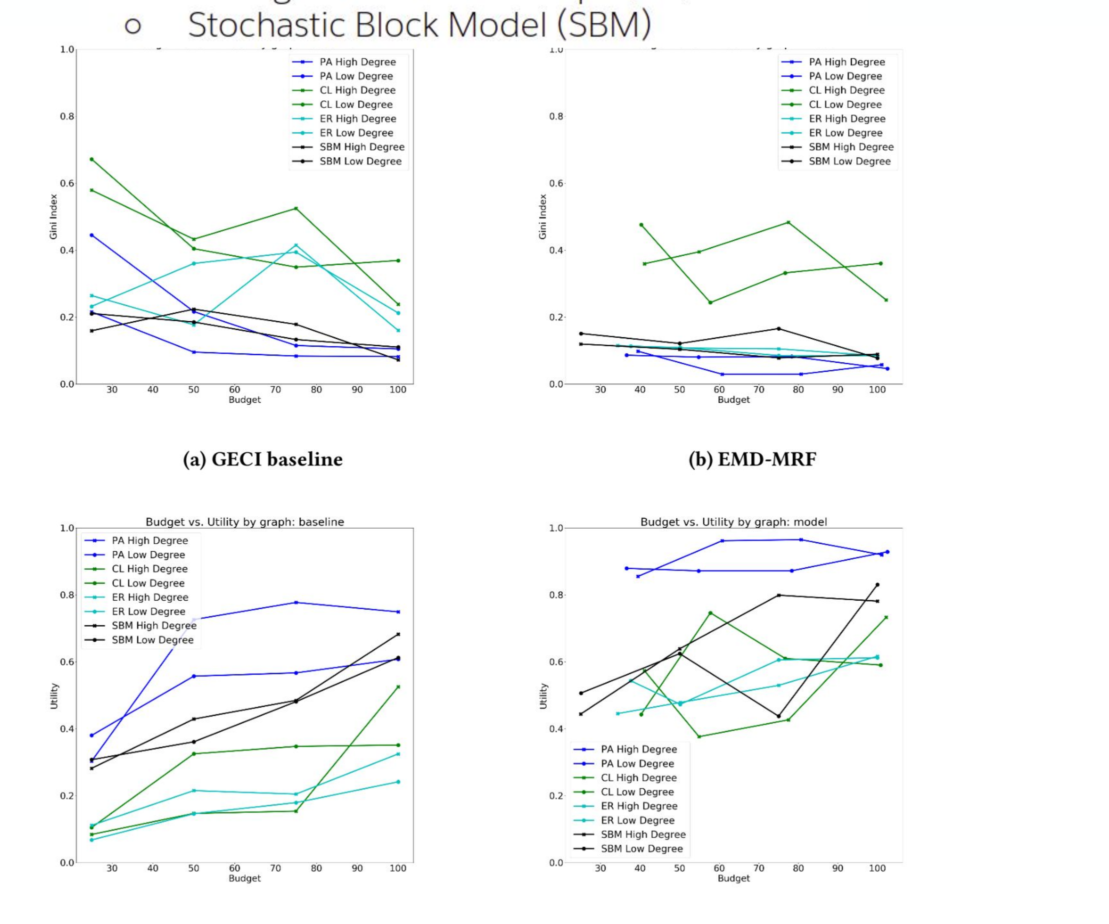


(a) Caltech. (b) Reed College. (c) Michigan State.

Figure 6: Mean shortest path of gender groups from the influence nodes

Exp: Synthetic Networks

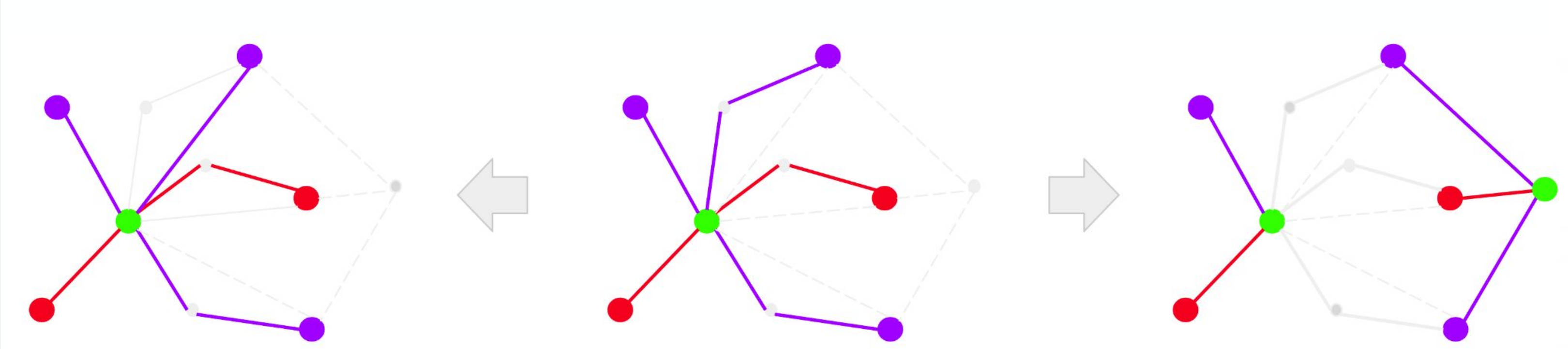
We evaluated on the following Synthetic graphs:
- Erdös-Rényi Random Graph (ER)
- Preferential Attachment Cluster Graph (PA)
- Chung-Lu Power Law Graph (CL)
- Stochastic Block Model (SBM)



(a) GECI baseline (b) EMD-MRF (c) GECI baseline (d) EMD-MRF

Figure 3: (a-b) Budget vs Gini Index. This shows the Gini Index for varying budgets for (a) the GECI baseline and (b) our proposed method. Our proposed model performs better, particularly at a smaller budget. (c-d) Budget vs Utility. This shows the utility for varying budgets for (c) the GECI baseline and (d) our proposed method. Our proposed method outperforms the deterministic baseline on all graphs at all budgets.

Method/Exp: Facility Placement



Equitable discrete graph augmentation

$$E^u = \underset{\substack{s.t. \sum_{g \in G} |V_g - \overline{V}_G| = 0 \\ \sum_{g \in G} \|E^u\|_0 < B}}{\text{argmax}} \sum_{g \in G} V_g$$

$$E^u(i,j) = \frac{1}{(1 + \exp(-(\phi(\vec{0}) + g_i)/\tau)}, \forall i,j \in S$$

Equitable facility placement

$$R^u = \underset{\substack{s.t. \sum_{g \in G} |V_g - \overline{V}_G| = 0 \\ \|R\|_0 < B}}{\text{argmax}} \sum_{g \in G} V_g$$

$$R^u(s) = \frac{1}{(1 + \exp(-(\phi(\vec{0}) + g_s)/\tau)}, \forall s \in S$$

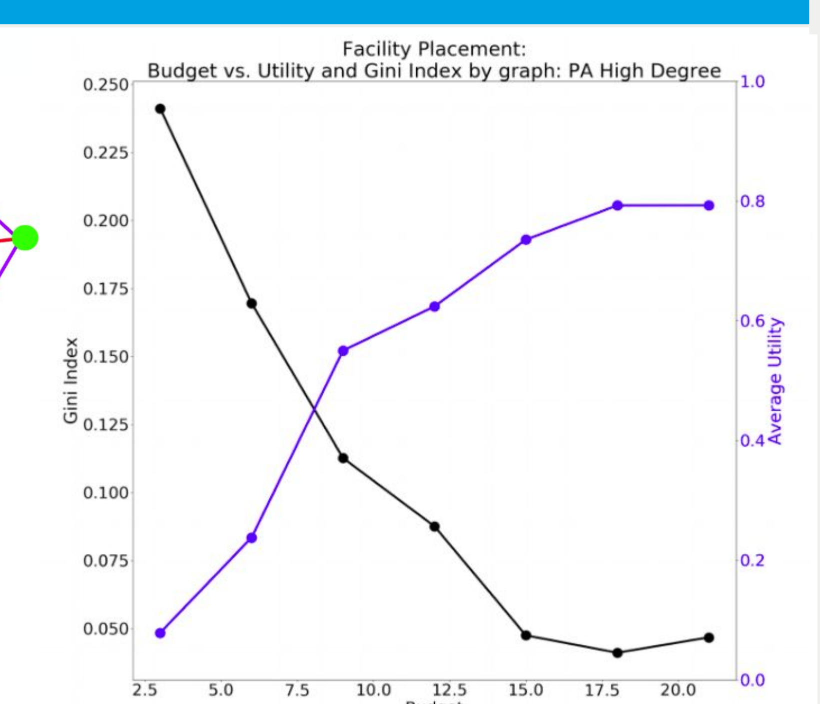


Figure 5: Facility placement results showing varying budget (facilities) vs. total Gini index and utility.

Contact
- gramachandran@salesforce.com

salesforce