復旦大學

Fast Computation and Optimization for Opinion-Based Quantities of Friedkin-Johnsen Model

Haoxin Sun, Yubo Sun, Xiaotian Zhou and Zhongzhi Zhang
Fudan University, Shanghai, China

Contributions

FUDAN UNIVERSITY

- **Partial Rooted Forests:** Propose a fast sampling-based method to estimate opinion-based quantities efficiently.
- **Sublinear Algorithms:** Design scalable algorithms for two FJ optimization problems.
- **Experiments:** Achieve higher accuracy, faster speed, and better scalability than baselines.

Friedkin–Johnsen model

- Each node has a fixed internal opinion s_i and an expressed opinion z_i that evolves over time.
- Opinions are updated by averaging one's own internal opinion with neightbors' expressed opinion.

$$z_i(t+1) = \frac{s_i + \sum_{j \in N_i} a_{ij} z_j(t)}{1 + \sum_{j \in N_i} a_{ij}}.$$

- At equilibrium $\boldsymbol{z} = \boldsymbol{\Omega s}$ where $\boldsymbol{\Omega} = (\mathbf{I} + \mathbf{L})^{-1}$ captures influence propagation across the network.

Partial Rooted Forest

FUDAN UNIVERSITY

- **Forest Matrix:** The fundamental matrix $\mathbf{\Omega} = (\mathbf{I} + \mathbf{L})^{-1}$ of FJ model can be estimated using forest sampling.
- **Partial Rooted Forest:** It focuses on a subset of nodes and uses loop-erased absorbing random walks.
- **Key Idea:** It captures key structural information without visiting all nodes and the expected time complexity is $O(rpl), 1 \leq r \leq \bar{d}$.

FJ Opinion Quantities & Optimization

FUDAN UNIVERSITY

- **Opinion Quantities:** The FJ model defines key measures for opinion evaluation, for example: disagreement $D = \boldsymbol{z}^\top \mathbf{L} \boldsymbol{z}$ captures opinion gaps between connected individuals.
- **Optimization Problems:**
  - Opinion Minimization (OpMin):

$$H^* = \arg \min_{|H|=k} \frac{1}{n} \boldsymbol{c}^\top \boldsymbol{z}$$

  - Polarization–Disagreement Minimization (PDMin):

$$T^* = \arg \max_{|T|=k} [\mathcal{I}(\mathcal{G}) - \mathcal{I}(\mathcal{G} + T)]$$

Experiments

➢ Datasets

Network	Nodes	Edges	$\bar{d}$	r
Delicious	536,108	1,365,961	5.1	2.4
Youtube	1,134,890	2,987,624	5.3	2.4
Pokec	1,632,803	22,301,964	27.3	9.0
Orkut	3,072,441	117,184,899	76.3	13.6
Livejournal	7,489,073	112,305,407	30.0	3.3
Twitter	41,652,230	1,202,513,046	57.7	7.7

Experiments

➢ Higher accuracy and faster speed

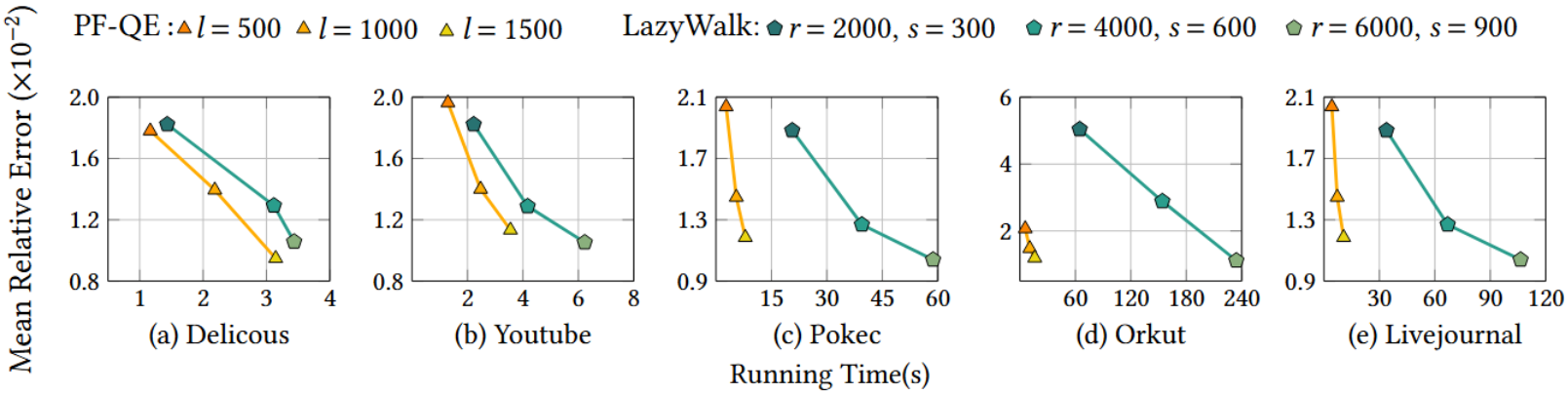

Figure 1: Comparison of mean relative error for z_i and running time under varying parameters for PF-QE and LazyWalk. Internal opinions are generated using the uniform distribution.

Experiments

Table 3: Mean relative errors and running time for three algorithms on six networks. Internal opinions are generated using the uniform distribution.

Network	Time(s)			Reletive Error in %									
				LazyWalk					PF-QE				
	LapSolver	LazyWalk	PF-QE	P	I	C	DC	D	P	I	C	DC	D
Delicious	7.2	3.1	2.2	0.90	0.48	0.83	0.42	3.58	1.38	0.42	0.55	0.32	4.05
Youtube	9.6	4.1	2.4	1.03	0.75	0.21	0.33	3.52	0.95	0.73	0.22	0.31	4.19
Pokec	49.8	39.5	5.3	1.83	0.93	0.25	0.50	11.01	2.01	0.96	0.21	0.51	10.63
Orkut	291.9	142.2	9.9	3.97	0.62	5.43	2.71	34.77	2.98	0.72	0.23	0.47	13.67
Livejournal	186.8	67.6	6.9	1.37	0.85	0.31	0.61	8.36	1.30	0.90	0.30	0.52	8.17
Twitter	-	160.6	32.3	-	-	-	-	-	-	-	-	-	-

Table 4: Running time and effectiveness of OPMIN (in terms of opinion decline θ) and PDMIN (in terms of P-D index decline δ).

Network	OpMin						PDMin			
	PF-OPMIN		FAST		EXACT		PF-PDMIN		FASTGREEDY	
	time	θ	time	θ	time	θ	time	δ	time	δ
Delicious	2.2	-6.21	12.4	-6.79	5.9	-7.80	160.3	-5.69	862.3	-5.63
Youtube	3.7	-6.56	33.2	-6.48	4.6	-7.79	154.3	-6.46	1814.3	-6.42
Pokec	6.3	-4.17	79.4	-4.13	60.2	-4.98	494.1	-3.18	18656	-3.17
Orkut	10.3	-1.88	163.1	-1.79	310.5	-2.43	774.2	-5.23	-	-
Livejournal	8.1	-6.56	296.8	-6.44	214.7	-7.12	436.5	-4.62	-	-
Twitter	33.7	-	625.0	-	-		2531	-	-	-

復旦大學

THANKS