Opinions and conflict in social networks: models, computational problems and algorithms

Lecture 3: Methods for mitigating polarization

Bertinoro International Spring School 2022

Aristides Gionis

argioni@kth.se

KTH Royal Institute of Technology

course overview

- lecture 1: introduction
  - polarization in social media; methods for detecting polarization
- lecture 2: mathematical background
  - submodular maximization; spectral graph theory
- lecture 3: methods for mitigating polarization
  - maximizing diversity, balancing information exposure
- lecture 4: signed networks; theory and applications
- lecture 5: opinion dynamics in social networks

overview of this lecture

- methods for mitigating polarization and echo chambers in social networks
- increasing awareness
- recommendation-based methods for increasing diversity and balancing information exposure
  - ideas inspired by influence-maximization setting
- other recommendation-based methods

improving awareness for alternate viewpoints

[Lahoti et al., 2018]

improve awareness

- develop tools for users to perceive their "news diet"
- visualize/navigate in the underlying ideology space, their position, the accounts they follow, the news they read
- offer functionalities such as

  "find a high-quality article on the same topic from the opposing viewpoint"

learning of ideological leanings

- infer ideological stances of users and content

  e.g., liberal–conservative space

- common latent space for users and content

- e.g., substitute ground-truth polarities in previous study with learned polarities

- joint non-negative matrix-factorization task

intuition

- map users and content in a joint latent ideology space

such that

- similar users are more likely to follow each other
- similar users are more likely to share similar content
- similar content is more likely to be shared by similar users

*similar means close in the latent ideology space

the problem setting

- social network $G = (V, E)$
  - adjacency matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$
- user–content matrix $\mathbf{C} \in \mathbb{R}^{m \times n}$
- latent matrix representing user ideology $\mathbf{U} \in \mathbb{R}^{n \times k}$
- latent matrix representing content ideology $\mathbf{V} \in \mathbb{R}^{m \times k}$
- decompose

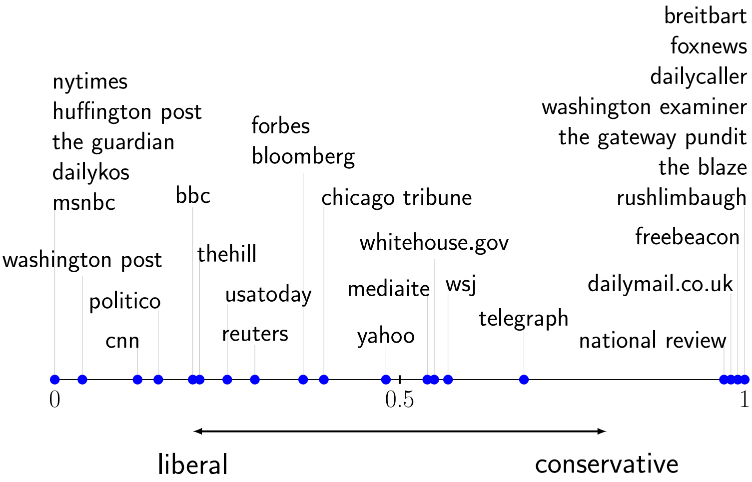
$$\mathbf{A} \approx \mathbf{U}\mathbf{H}_u\mathbf{U}^T \quad \text{and} \quad \mathbf{C} \approx \mathbf{U}\mathbf{H}_v\mathbf{V}^T$$

subject to orthonormal $\mathbf{U}$ and $\mathbf{V}$ and graph-regularization

evaluation

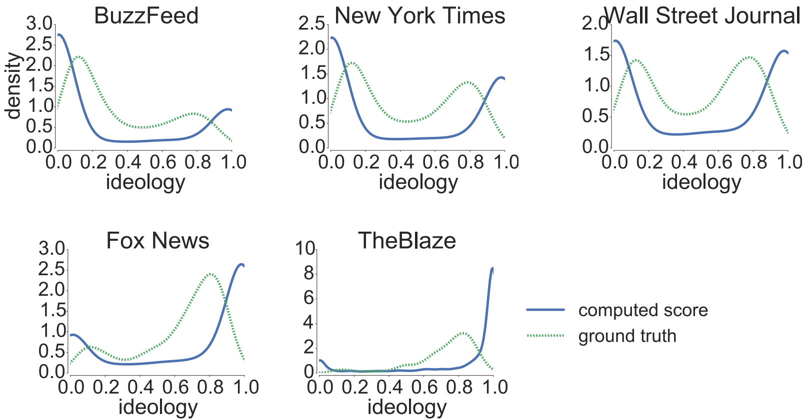
- twitter data from 2011 to 2016, focusing on controversial topics
  - gun control, abortion, obamacare, etc.
- 6 391 users and 19 million tweets
- gather ground-truth polarity scores
  - content polarity [Bakshy et al., 2015]
  - user polarity [Barberá, 2015]

content ideology scores



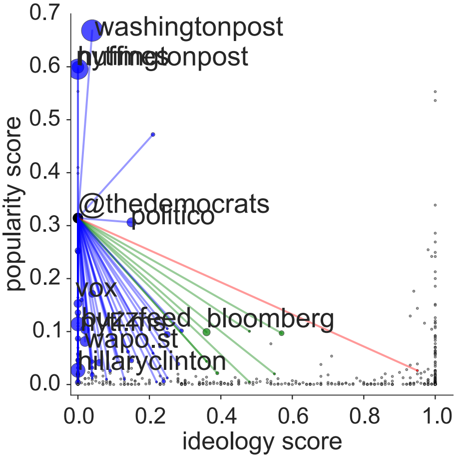
correlation with ground-truth scores 0.82

audience ideology scores

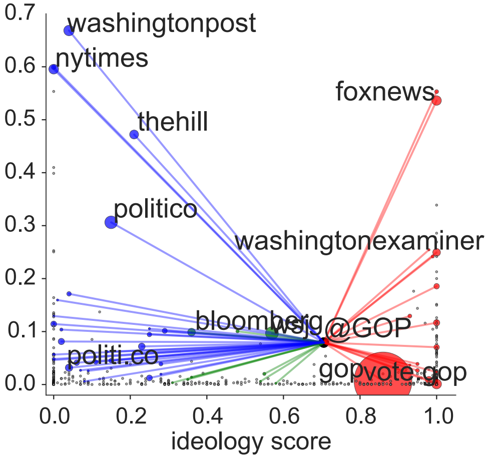


correlation of user ideology scores with ground-truth 0.90

visualizing the information bubble



@thedemocrats



@gop

maximizing diversity in social networks

[Garimella et al., 2017, Matakos et al., 2020b, Matakos et al., 2020a]

maximizing diversity

- goal : make recommendations to maximize diversity
- what is diversity and how to measure it?
- user level : recommend diverse content
- network level : make recommendations so that friends see different content
  - motivation : friends can discuss / debate
- combinations
- another consideration : propagation effects, or not

maximizing diversity in social networks

- one way to achieve this is via recommendations
- important question : what is diversity and how to measure it?
- what kind of recommendations ?
  - friend recommendation : recommend an interesting account to follow
  - content recommendations : recommend interesting content
- considerations :
  - make minimal number of recommendations (k)

    motivation : intervene as little as possible in the “organic” operation of the network
  - make recommendations so that friends see different content

    motivation : friends can discuss / debate
  - consider network-cascade effects

background knowledge

influence maximization in social networks

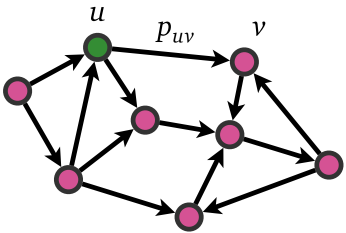
[Kempe et al., 2003]

influence in social networks

- people living in a society are influencing each other
  - behavior they exhibit
  - opinions they form
  - decisions they make
  - products they buy
  - clothes they wear
- viral marketing
  - select few initial adopters in a social network
    to maximize the spread of adoption of a product / action /

modeling influence in a social network

- independent-cascadel: a popular model in the literature
  - introduced by [Kempe et al., 2003] (about 9K citations in Google scholar)
- it assumes a social network, directed graph, $G = (V, E, p)$
  with edge probabilities p_{uv} for each edge $(u, v) \in E$
- basic model considers a single action adopted (spread) in the network
- at time 0: a set of initial nodes (seeds) adopt
- at time t: a node u adopts
- at time $t + 1$: a neighbor v of u adopts with prob p_{uv} -
  - one-time opportunity



the influence-maximization problem

informally

- select k initial adopters to maximize the total number of adopters in the network

a bit more formally

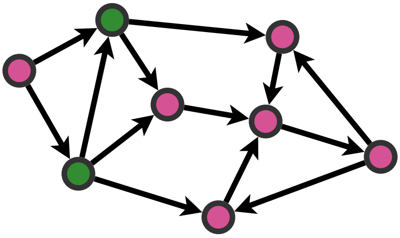
- consider social network $G = (V, E, p)$ with edge probabilities p
- let $\sigma(A)$ be the total number of adopters, for a set of initial adopters $A \subseteq V$

the influence-maximization problem

- select $A \subseteq V$ with $|A| = k$, so as to maximize $\sigma(A)$
- but notice that $\sigma(A)$ is a random variable,
  - so, we are interested to maximize the expectation $\mathbb{E}[\sigma(A)]$

computing $\sigma(A)$

- given a set of seeds A how to compute $\sigma(A)$?
- how to deal with edge probabilities p?

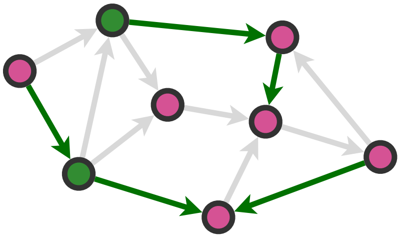


computing $\sigma(A)$

- given a set of seeds A how to compute $\sigma(A)$?
- how to deal with edge probabilities p?

idea:

- assume that edge coin-flips are drawn before the cascade
- sampled edges yield a sample graph X

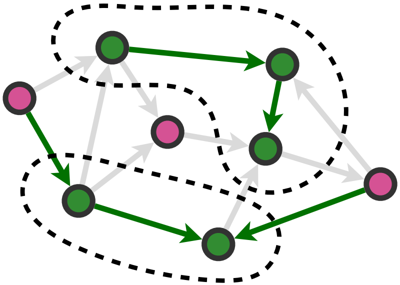


computing $\sigma(A)$

- given a set of seeds A how to compute $\sigma(A)$?
- how to deal with edge probabilities p?

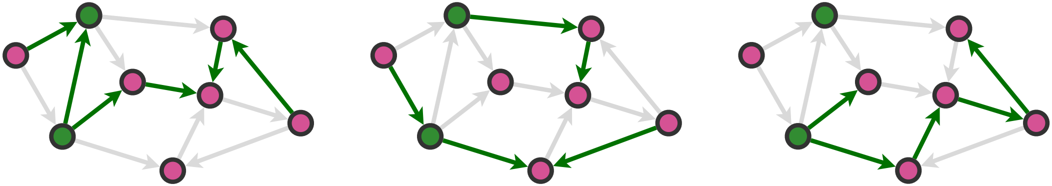
idea:

- assume that edge coin-flips are drawn before the cascade
- sampled edges yield a sample graph X
- $\sigma(A)$ in sample X is the number of nodes reachable from A



computing $\sigma(A)$

- generate a large number of sample graphs X



- compute $\sigma_X(A)$ in each X
  approximate expectation $\mathbb{E}[\sigma(A)]$ via sample mean of all $\sigma_X(A)$
  - Monte-Carlo simulation

- this works OK for approximating $\mathbb{E}[\sigma(A)]$, given A ...
  ... but how to find the optimal A that maximizes $\mathbb{E}[\sigma(A)]$?

the influence-maximization problem

- given $G = (V, E, p)$, find $A \subseteq V$ with $|A| = k$, so as to maximize $\mathbb{E}[\sigma(A)]$

bad news

- the influence-maximization problem is **NP**-hard
  - in fact, it is **NP**-hard to maximize $\sigma_X(A)$ for a sample graph X

good news

- the function $\sigma_X(\cdot)$ is monotonically non-decreasing and submodular

solving the influence-maximization problem

- the function $\sigma_X(\cdot)$ is monotonically non-decreasing and submodular
- it follows that $\mathbb{E}[\sigma(A)]$ is monotonically non-decreasing and submodular
- the Greedy algorithm provides a $\left(1 - \frac{1}{e} - \epsilon\right)$-approximation guarantee
  - (where does ϵ come from?)

Greedy

1. $A \leftarrow \emptyset$
2. **while** $(|A| < k)$ **do**
3. $\quad v \leftarrow \arg\max_{u \in V} \{\sigma(A \cup \{u\}) - \sigma(A)\}$
4. $\quad A \leftarrow A \cup \{v\}$

the influence-maximization problem — extensions and variants

- the influence-maximization problem is widely-studied problem, many extensions and variants have been proposed
  - inferring the network structure and the influence probabilities
  - scaling up the method and avoiding Monte-Carlo simulation
  - multiple actions and topic-aware influence
  - competitive and game-theoretic formulation
  - time-varying networks
  - reinforcement-learning formulation
  - ...

balancing information exposure

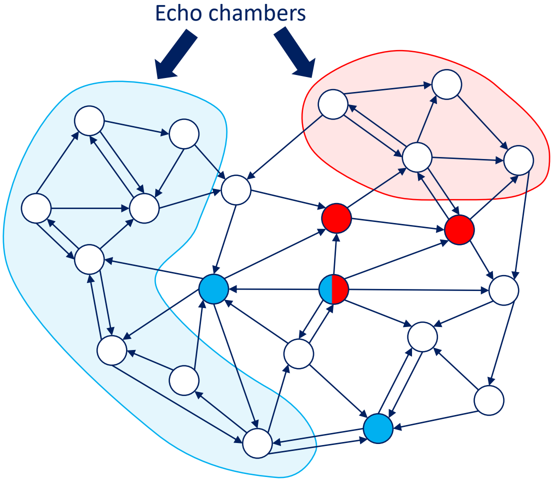
[Garimella et al., 2017]

balancing information exposure

- setting inspired by influence maximization and viral marketing
  - a social network and two campaigns
  - seed nodes $\mathcal{I}_1$ and $\mathcal{I}_2$ for the two campaigns
  - a model of information propagation
  - for each edge e, we assume two probabilities, $p_1(e)$ and $p_2(e)$, one for each campaign
- the problem of balancing information exposure
  - find additional seeds S_1 and S_2, with $|S_1| + |S_2| \leq k$
  - s.t. minimize # of users who see only one campaign or maximize # of users who see both or none

illustration

social discussion on fracking



balancing information exposure — results

- optimization problem is **NP**-hard
- minimization problem is **NP**-hard to approximate
- maximization problem: objective function non monotone and non submodular
  - but there is some structure that we can exploit
- different models of how the two campaigns propagate
- approximation guarantee $\frac{1}{2}(1 - \frac{1}{e})$

adapted greedy algorithm

Algorithm 2: Hedge, greedy algorithm, where each step is as good as adding the best common seed

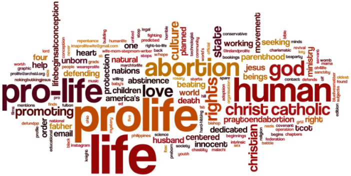
$S_1 \leftarrow S_2 \leftarrow \emptyset$;
while $|S_1| + |S_2| \leq k$ **do**
  $c \leftarrow \arg\max_c \Phi(S_1 \cup \{c\}, S_2 \cup \{c\})$;
  $s_1 \leftarrow \arg\max_s \Phi(S_1, S_2 \cup \{s\})$;
  $s_2 \leftarrow \arg\max_s \Phi(S_1 \cup \{s\}, S_2)$;
  add the best option among $\langle c, c\rangle$, $\langle \emptyset, s_1\rangle$, $\langle s_2, \emptyset\rangle$, $\langle s_2, s_1\rangle$, to $\langle S_1, S_2\rangle$ while respecting the budget.

balancing information exposure — example

Side 1
Pro-Choice

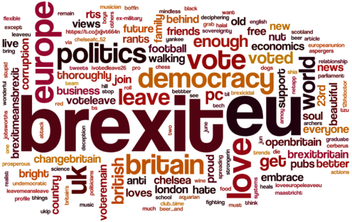
Side 2
Pro-Life

Hedge



Pro-Remain

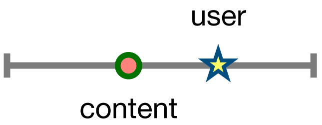
Pro-Leave



maximizing content diversity in social networks

[Matakos et al., 2020a]

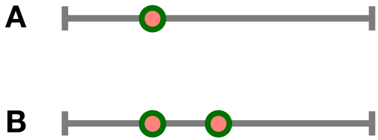
a simple model



a simple model:

represent user and content leanings as values in the $[0, 1]$ interval

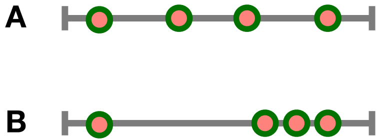
a simple model



for diversity :

B is preferable to **A**

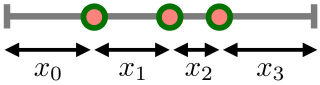
a simple model



for diversity :

A is preferable to **B**

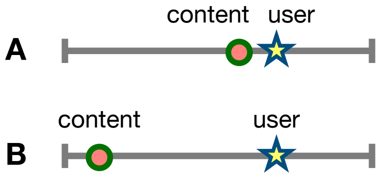
a simple model



user v diversity score for exposure to set of items $\mathcal{I}$

$$d_v(\mathcal{I}) = 1 - \sum_i x_i^2$$

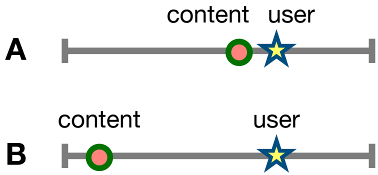
a simple model



for a cascade effect :

A is preferable to **B**

a simple model



for modeling cascades :

- if user u shares item i
- a user v, friend of u, may reshare item i
- reshare probability depends on leanings $\ell(u)$, $\ell(v)$, and $\ell(i)$

maximizing diversity of exposure

- given a set of items with different leanings, and possibly different viralities
- goal : recommendations to maximize user diversity
- consideration : recommended content may be shared among users, creating possible cascades
- make a small number of recommendations
  - why ? intervene as little as possible in the organic operation of the network

[Matakos et al., 2020a]

maximizing diversity of exposure

- problem formulation inspired by influence maximization
- item propagation modeled by independent cascade
  - influence prob. depend on user and item leanings
- recommend a (small) number of items to users so as to maximize the expected diversity score

$$E\left[\sum_{v \in V} d_v(\mathcal{I}(v))\right]$$

$\mathcal{I}(v)$: items that v is exposed, considering also cascades

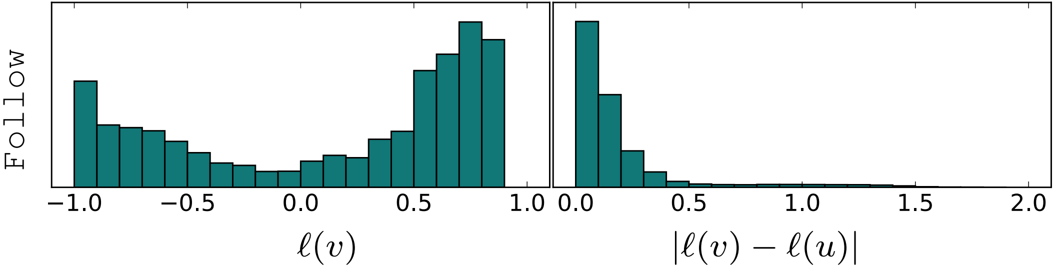
- small number of recommendations
  - at most k recommendations in total
  - at most k_v recommendations to user v

maximizing diversity of exposure — theoretical results

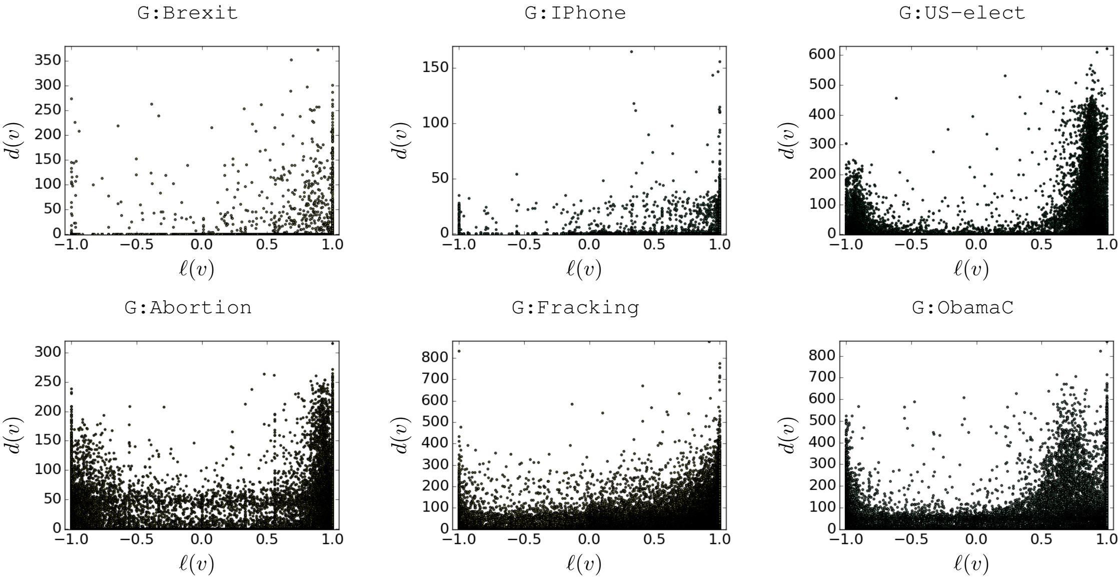
- diversification problem is **NP**-hard
- diversity objective is submodular
- greedy algorithm provides $\frac{1}{2}$ approximation
  - maximizing a submodular function under partition matroid constraints
- but computation prohibitively expensive
  - Monte-Carlo simulations
- adapt recent techniques to obtain highly scalable algorithm
  - generalize the idea of reverse-reachable sets
  - sample-size estimation for early stopping lower bound OPT, which gives bound on sample size

experiments — exploration of datasets

Histograms of node leanings (left) and leaning differences across the edges (right) of `Twitt` network.



experiments — exploration of datasets



comparison against baselines

Dataset (k, k_u)	diversity score: Myopic	diversity score: Max-Var	diversity score: Min-Var	diversity score: TDEM	memory (mb)	run time (s)
`DBLP:BSch` (5, 1)	91	38	112	132	457	2
`DBLP:CPap` (5, 1)	177	44	160	254	276	1
`DBLP:PYu` (5, 1)	535	100	706	912	285	2
`TPair:X` (5, 1)	198	58	160	289	279	1
`TPair:Y` (5, 1)	946	246	841	1 051	174	2
`TPair:Z` (5, 1)	4 147	3 020	3 243	3 997	1 658	42
`Tweet:S5` (50, 5)	816	492	1 006	1 298	5 943	24
`Tweet:S2` (50, 5)	6 127	1 500	10 673	12 386	656	9
`Tweet:M5` (50, 5)	15 509	4 273	21 189	27 659	3 100	77
`Twitt:Follow` (50, 5)	17 635	17 396	8 138	28 164	373	44
`G:Brexit` (50, 5)	542	340	477	1072	23 725	72
`G:IPhone` (50, 5)	14 758	4 129	9 567	26 633	1 803	15
`G:US-elect` (50, 5)	427	311	380	844	45 828	525
`G:Abortion` (50, 5)	640	278	520	3 608	154 588	1 275
`G:Fracking` (50, 5)	561	339	514	4 575	400 565	4 785
`G:ObamaC` (50, 5)	523	311	536	4 420	360 449	3 936
`Twitt:XL` (50, 5)	394 203	360 079	261 281	939 215	3 438	806

diversity-aware friend recommendation

[Matakos et al., 2020b]

problem setting and modeling assumptions

- a social graph $G = (V, E, w)$, with adjacency matrix A
- w_{ij} denotes the connection strength between people i and j
- s_i denotes person's i content exposure on some topic, say, $s_i \in [-1, 1]$
- an echo chamber is when people around you are exposed to similar content
- diversity index : $\eta = \sum_{(i,j)\in E} w_{ij}(s_i - s_j)^2 = \mathbf{s}^\top \mathbf{L}\,\mathbf{s}$
  - high diversity : many individuals have different exposure than social connections
- a recommendation to person i may lead to a change in their exposure value s_i
- probability of accepting the recommendation
- problem : make k recommendations to maximize the diversity index η

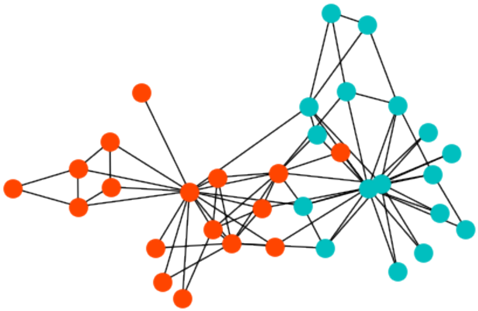
[Matakos et al., 2020b]

a simple problem abstraction

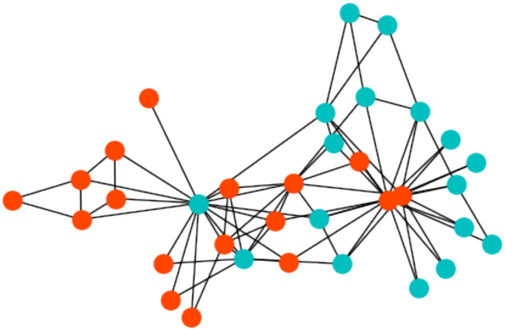
- simplify the problem to understand its complexity
- assume "extreme" exposure values $s_i \in \{-1, 1\}$
- assume that a recommendation reverses a user's exposure value $-1 \longleftrightarrow 1$
- our problem becomes: given a graph whose vertices are painted in two colors (say, `red` or `blue`), select k vertices and swap their color, so as to maximize the number of bi-chromatic edges
- in math form

$$\begin{array}{ll} \max & \mathbf{x}^\mathsf{T} \mathbf{P}\, \mathbf{x} \\ \text{subject to} & \mathbf{b}^\mathsf{T} \mathbf{x} \leq k \\ & \mathbf{x} \in \{0,1\}^n \end{array}$$

toy example



(a) Echo-chamber graph



(b) Graph with diversified exposure

optimal solution for $k = 4$

problem complexity

- problem **NP**-hard (generalization of MAX-CUT)
  - also **NP**-hard to approximate
- problem is non-convex
  - matrix $\mathbf{P}$ is not positive-semidefinite
  - an instance of quadratic knapsack (QK)

solution I — SDP relaxation

- inspired by MAX-CUT and QK solutions
- "lift" the program to the space of square matrices
  - set $\mathbf{X} = \mathbf{x}\,\mathbf{x}^\intercal$ (not convex)
  - relax $\mathbf{X} - \mathbf{x}\,\mathbf{x}^\intercal \succeq 0$ (it implies $\mathbf{X} \succeq 0$)
- objective and constraints are relaxed to corresponding convex ones
- solve the resulting SDP relaxation for $Opt = (\mathbf{X}^*, \mathbf{x}^*)$
- sample fractional $\mathbf{z}$ from a certain distribution specified by $(\mathbf{X}^*, \mathbf{x}^*)$
- round $\mathbf{z}$ to binary $\bar{\mathbf{x}}$
- repeat until constraints are satisfied

solution II — greedy

- SDP solvers use expensive interior-point methods
- efficient methods are needed for large-scale data
- consider simple greedy method that select the most cost-effective item
  - linear-time complexity

experimental results

Table 2. Solution quality and bounds from the relaxations

Dataset	k	IQP	SDP-Relax	Glover	I-Greedy	S-Greedy
Karate	$0.1\,n$	184	184 (185.72)	184 (209.12)	184	184
	$0.2\,n$	224	224 (236.52)	216 (276.2)	224	204
	n	244	244 (253.92)	208 (312.00)	228	204
Karate-D	$0.1\,n$	200	200 (215.84)	196 (263.4)	200	200
	$0.2\,n$	220	220 (242.96)	200 (319.36)	212	208
	n	244	244 (253.92)	200 (372.00)	220	192
Books	$0.1\,n$	828	828 (831.24)	828 (943.6)	828	828
	$0.2\,n$	1056	1048 (1089.04)	996 (1320.04)	1056	992
	n	1236	1224 (1273.72)	1068 (1788.00)	1192	1012
Books-D	$0.1\,n$	1060	1048 (1091.8)	996 (1313.28)	1052	1008
	$0.2\,n$	1140	1124 (1193.92)	1120 (1555.28)	1136	1016
	n	1236	1228 (1273.76)	1128 (1990.0)	1144	972
Twitter100	$0.1\,n$	1700	1700 (1700.76)	1696 (1918.08)	1700	1696
	$0.2\,n$	2396	2396 (2405.52)	2356 (3166.2)	2396	2368
	n	–	3172 (3216.84)	2932 (5 624.00)	3160	2588
Twitter100-D	$0.1\,n$	2972	2968 (3011.76)	2888 (3671.96)	2968	2860
	$0.2\,n$	–	3.028 (3102.12)	2916 (4 285.28)	3044	2860
	n	–	3172 (3216.84)	3100 (5 984.00)	3064	2948
Blogs	$0.1\,n$	–	–	39 512 (50 636.24)	39 516	35 556
Elections	$0.01n$	–	–	–	471 800	469 928
Twitter	$0.001n$	–	–	–	–	6 715 012

summary and discussion

- different ideas to mitigate polarization and information bottlenecks in social networks
  - improve awareness for opposing viewpoints
  - balance information exposure, maximize diversity of content
  - friend recommendations
- an important aspect that was not discussed:
  - how to give control to users over these choices

references I

Bakshy, E., Messing, S., and Adamic, L. A. (2015).
Exposure to ideologically diverse news and opinion on facebook.
Science, 348(6239):1130–1132.

Barberá, P. (2015).
Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data.
Political analysis, 23(1):76–91.

Garimella, K., Gionis, A., Parotsidis, N., and Tatti, N. (2017).
Balancing information exposure in social networks.
Advances in Neural Information Processing Systems, 30.

Kempe, D., Kleinberg, J., and Tardos, E. (2003).
Maximizing the spread of influence through a social network.
In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM Press.

Lahoti, P., Garimella, K., and Gionis, A. (2018).
Joint non-negative matrix factorization for learning ideological leaning on twitter.
In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 351–359.

references II

Matakos, A., Aslay, C., Galbrun, E., and Gionis, A. (2020a).
Maximizing the diversity of exposure in a social network.
IEEE Transactions on Knowledge and Data Engineering.

Matakos, A., Tu, S., and Gionis, A. (2020b).
Tell me something my friends do not know: Diversity maximization in social networks.
Knowledge and Information Systems, 62(9):3697–3726.