Reducing Polarization in Opinion Networks in the Presence of Stubborn Leaders

Samuel Selleck

LUND UNIVERSITY

Department of Automatic Control

MSc Thesis
TFRT-6157
ISSN 0280-5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by Tryckeriet i E-huset
Lund 2022

Abstract

We study the problem of reducing polarization (variance) of opinions at stationarity in a directed weighted graph with node set divided into two groups: stubborn, initialized with a fixed opinion and regular who repeatedly update their opinion to the average of their out-neighbors, known as the DeGroot model with stubborn nodes. We show how the polarization can be minimized for a number of simple constraints, but that the problem in general is not convex. Theory is developed for the change in opinions at stationarity and the polarization measure for a rank-1 update of the network (encompassing both addition of a directed and undirected link in the network). An algorithm for gradient approximation is presented, given directly by the analytical gradient formulation and method of matrix-vector product estimation. Lastly variations of the algorithm together with other trivial methods of recommending a link are compared for a number of random and real networks.

Acknowledgments

I would like to thank Giacomo Como my supervisor for his patience, guidance and the many times he's steered me back on the right track when I've felt lost, my family and friends for their encouraging words and support during the last months, and life the universe and everything for me and your existence.

Contributions

We develop a model of measuring polarization in networks. The problem was shown to in general be non-convex and belong to the class of invex functions. The corresponding discrete link problem was shown to be NP-hard to minimize. The effect of a rank-1 update of the edges was characterized (giving rise to the concept of opinion kernels) and an efficient vectorized iterative method for computing the gradient of the problem was introduced. Finally, estimates of the reduction in polarization possible under a single link as well as link-cost constraints in a number of synthetic and real world networks was studied.

Contents

1. Introduction **11**

2. Preliminaries **13**
- 2.1 Networks . . . 13
- 2.2 DeGroot Model of Opinion Dynamics . . . 15
- 2.3 The Addition of Stubborn Agents . . . 16
- 2.4 Relation to Other Problems . . . 17

3. Reducing Polarization **19**
- 3.1 Polarization in Networks . . . 19
- 3.2 Problem Formulation . . . 20
- 3.3 Invariant Properties . . . 21
- 3.4 Characterization of Minima . . . 22
- 3.5 General Problem Properties . . . 23
- 3.6 Approximation Methods . . . 28

4. Examples **35**
- 4.1 Path graph . . . 35
- 4.2 Barbell Graph . . . 36

5. Algorithms **39**
- 5.1 Opinions at Stationarity . . . 39
- 5.2 Fast Computation of $(I-P_{RR})^{-1}v$ . . . 39
- 5.3 Computing the Maximum Edge Derivative . . . 40

6. Simulations **41**
- 6.1 Networks . . . 41
- 6.2 Single Link Addition . . . 41
- 6.3 Gradient Descent . . . 42
- 6.4 Results . . . 43

7. Conclusion **45**

Bibliography **46**

1

Introduction

Opinion Dynamics studies how the opinions of agents in a network changes given a model of their interaction, and is commonly applied to the case of understanding the evolution of opinions of individuals in social networks. One interesting metric to observe in opinion networks is the polarization of opinions and how it varies between different networks and models. The polarization metric could be assumed to correlate with levels of unrest and violence in a larger society or how close a room full of politicians are to finding common ground [Jilani and Smith, 2019]. In a political context it is well known that polarization when measured in a multitude of ways is considered increasing (see Figure 1.1), the primary cause being disputed.

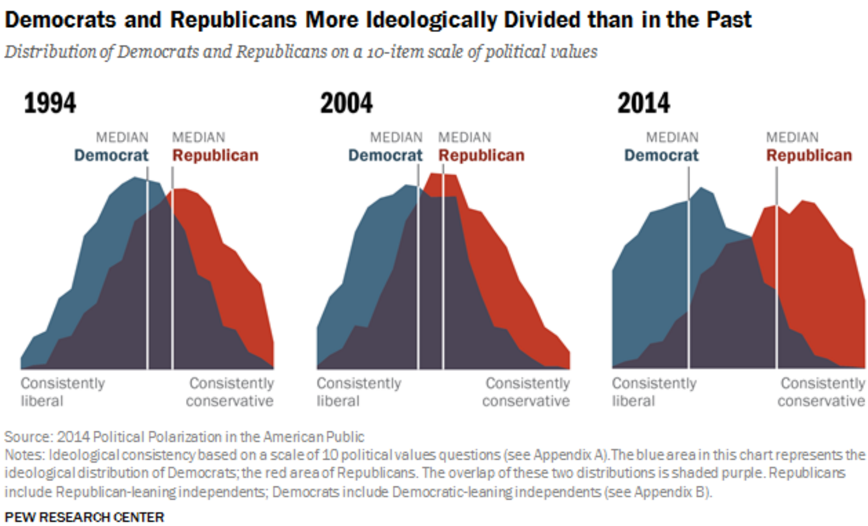

Figure 1.1 Increasing Polarization [Pew Research Center, 2014]

A wide variety of models in opinion dynamics converge under the influence of a group of individuals with fixed opinions to the same state. The choice to add these individuals with unchanging opinion is backed by observations of how group polarization arises [Cass, 1999]. Studying how this steady state changes after a network intervention and finding a way to measure the polarization before and after this action could give rise to real ways of lowering the level of polarization in media and society.

Related Work

Studying ways of reducing polarization in networks is not new, most closely related to this work is Garimella et al. (2017) who explores ways of reducing controversy by link recommendation in a setting where no opinions are known. Differences are measured by partitioning of the graph into two subsets and studying the ratio between the number of random walks that starts on one side and ends in the other and the number that stays in the same set, a measure by the name RWC (random-walk controversy). Link recommendation is also done by Amelkin and Singh [Amelkin and Singh, 2019] in the context of restoring the mean opinion of the network. Mimimizing a combined polarization and disagreement measure is studied by Musco et al. [Musco et al., 2017] in the context of the closely related Friedkin-Johnsen model, where global graph structure is considered instead of individual links.

Research Question

This work aims to contribute a strategy of single link recommendation for reducing polarization when only the opinion of key agents are known, the advantage compared to the other models presented being that cases were nodes in different clusters doesn't necessarily have differing opinions can be detected, without assuming that the opinion of all agents are known. Given a network, estimation of the optimal way to modify the network to reduce polarization can then be done by only surveying the most influential agents in the network, the so called opinion leaders.

Thesis Outline

In Chapter 2 we present the theory of networks and common terminology. In Chapter 3 we formulate the problem of opinion minimization, present a number of cases for when the optimal solution can be found directly, show why the problem in general is not convex, and present methods of gradient approximation. We use these approximation methods to construct algorithms in Chapter 5 and explore polarization reduction of real and random networks computationally in Chapter 6.

2

Preliminaries

2.1 Networks

Networks/graphs can represent a vast number of systems describing the interconnections (edges) between entities (nodes), a number of representation examples are presented in Table 2.1 and a simple network is depicted in Figure 2.1.

System	Nodes	Edges
Transport	Destinations	Roads
The Brain	Neurons	Dendrites
Research	Publications	References
Social Network	People	Relationships

Table 2.1 Network Examples

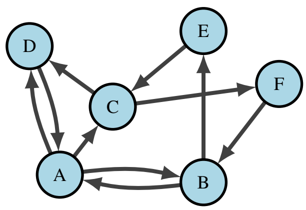

Figure 2.1 A directed unweighted network

Graphs can be categorized based on their properties. They are said to be

- **Undirected** if the edges have no inherent order of the nodes belonging to an edge. Links from one website to another clearly point **from** one **to** another, while friendships are **undirected** relationships (if x is a friend of y, y is also a friend of x).

- **Weighted** if there for each edge is an associated number w. This number can for example can represent the throughput in the context of roads or the strength of a friendship in the context of a social network.

Mathematically, any directed weighted network can be described as a collection $G = (V, \mathscr{E}, W)$ where $V = \{1, 2 \dots |V|\}$ is the set of nodes in the graph and $\mathscr{E} \subset V \times V$ is the set of edges containing pairs of nodes that are connected. The weights of the edges in the graph are given by an associated adjacency matrix $W \in \mathbb{R}^{|V| \times |V|}$ whose elements are equal to

$$W_{i,j} = \begin{cases} 0 & \text{if } (i,j) \notin \mathscr{E} \\ w_{i,j} & \text{if } (i,j) \in \mathscr{E}. \end{cases}$$

If a graph is **undirected**, its adjacency matrix is symmetric. If a graph is **unweighted**, all non-zero entries $w_{i,j} = 1$. We present another list of properties

- The **out-degree** of a node is the total weight of the links pointing from the node to other nodes in the network. Equivalently, the **in-degree** is the total weight of all links pointing towards the node. If a graph is regular, nodes in-degree and out-degree are the same. The out-degree of a node is equal to the row-sum of the nodes corresponding row in the adjacency matrix, and its in-degree is the column sum.
- A **walk** on a graph is a sequence of edges $(\mathscr{E}_1, \mathscr{E}_2 \dots \mathscr{E}_n)$ where the tail of each link is equal to the head of the last one. A **trail** is a walk where all edges in the sequence are unique and a **cycle** is a trail that starts and ends in the same node but otherwise have never visits the same node twice.
- A set of nodes U in a graph G is said to be **globally reachable** if there for each node in the network exists a walk that starts in the node and ends in a node in the set U. In the graph in Figure 2.1, $\{D\}$ is globally reachable, but $\{A, C\}$ is not, since there exists no path from D to A or C.
- A graph is **connected** if the set of all nodes V is globally reachable.
- A graph is said to be **aperiodic** if there exists at least one pair cycles in the network whose lengths are relatively prime, ie. their greatest common divisor is 1. The graph in Figure 2.1 is aperiodic, since the two cycles (A, C, D, A) and (A, B, A) have length 3 and 2 respectively.

A adjacency matrix of size $n \times n$ uniquely describes the interconnections of a graph with n nodes. From the adjacency matrix we derive a number of related variables

Diagonal Degree Matrix A diagonal matrix $D = \text{diag}(W\mathbb{1})$ whose diagonal entries are the out-degree of the corresponding node. Used in the construction of the below matrices.

Normalized Weight Matrix The matrix $P = D^{-1}W$ is he normalized weight matrix. Often used in the context of Markov chains (then known as the stochastic matrix or probability transition matrix), each entry in P describes the probability of moving from one node to another during a random walk, if the chance of following each link from a given node is uniformly distributed. The P matrix has the property that the sum of out-degree weights are equal to 1, $P\mathbb{1} = \mathbb{1}$. Notably, the probability of transitioning between two nodes i and j in n steps is given by $(P^n)_{i,j}$. Perron-Frobenius theory gives that there exists a non-negative eigenvector π of P^T associated with the eigenvalue 1 [Como and Fagnani, 2021]. The normalized eigenvector of P^T such that $\mathbb{1}^T\pi = 1$ is called an **invariant probability distribution** of the network. π can be interpreted for a connected network as the probability of ending up in each node when performing a random walk on the graph as time approaches infinity.

Laplacian Matrix The Laplacian matrix $L = D - W$ has properties itself direcly related to the graph structure, it also shows up in calculations of distributed averaging dynamics on the graph, as will be seen in sections below. One such property is that the second smallest eigenvalue is an approximation of the sparsest cut of a graph. Similarily to the normalized weight matrix, the Laplacian has related properties (many of them given directly by the fact that $L = D^{-1}(I - P)$). All row sums of L are zero ($L\mathbb{1} = 0$). The diagonal of L is always positive and all off-diagonal entries are negative.

Now suppose the nodes in a network represent any kind of entity/agent that can form an opinion on a subject that can be represented as a real number. This could be a sensor measuring temperature in a room, the amount of opposition or agreement of a new law being passed or the level of perceived climate change urgency by a population. One could imagine the opinions of agents change in some way over time depending on the network structure.

2.2 DeGroot Model of Opinion Dynamics

A common model for the change of opinions in a graph is the DeGroot learning model [DeGroot, 1974]. It has been shown to be a good predictor of the real opinion formation process empirically in the case of binary opinion formation [Chandrasekhar et al., 2015]. The model states that agents update their opinion by a simple weighted average of the opinion of its out-neighbors (outwardly going directed links). The next value of node i given the previous is

$$x_i(t+1) = \frac{\sum_{j \in N(i)} w_{i,j} x_j(t)}{\sum_{j \in N(i)} w_{i,j}}$$

Where $w_{i,j}$ is the element in row i and column j, $N(i)$ the out neighborhood of node i (set of nodes that can be reached by following a link fron i). Given a fixed network G

with node set V and an initial opinion distribution vector $x(0) \in \mathbb{R}^{|V|}$ the stationary opinion distribution in the network is described by the following theorem:

THEOREM 1 ([COMO AND FAGNANI, 2016] DEGROOT MODEL) For any connected aperiodic graph G where every node i has an associated initial opinion $x_i(0)$, and the opinion of each node is updated each step by the weighted average of the opinion of its out-neighbors according to (2.2). The stationary opinions are equal to

$$x = \pi^T x(0)\mathbb{1}$$

where π is the invariant probability distribution of the network.

2.3 The Addition of Stubborn Agents

Now divide the set of nodes V into two parts. One set of **regular** nodes R and one set of **stubborn** nodes S. Let the regular nodes be updated by DeGroot averaging dynamics and the stubborn agents have a fixed opinions given by the constant vector $x_S \in \mathbb{R}^{|S|}$. The following theorem describes the stationary opinions of the system

THEOREM 2 ([COMO AND FAGNANI, 2016] DG MODEL + STUBBORN NODES) The stationary opinions of a network G with regular nodes R and a globally reachable set of stubborn nodes S is

$$x = \begin{pmatrix} x_R \\ x_S \end{pmatrix}, \quad x_R = L_{R,R}^{-1} W_{R,S} x_S.$$

where each indexed matrix is the sub-matrix selected by only keeping the rows associated with the nodes in the first set, and the columns corresponding to the nodes in the second set of the original matrix associated with the graph G. This model will be referred to as the DGS model.

Proof The update process for each regular node is as in (2.2) for all regular nodes. This is compactly described by the system

$$x_R(t+1) = D_{R,R}^{-1}(W_{R,R} x_R(t) + W_{R,S} x_S)$$

The system has a unique stationary solution under the added assumption that the set S is globally reachable, since if there doesn't exist a walk from a node i to S in G, its equilibrium will not depend on the opinion of the stubborn nodes (instead on the initial opinion of regular nodes). This equilibrium is given by inserting $x_R = x_R(t+1) = x_R(t)$ and we get

$$\begin{aligned} & x_R = D_{R,R}^{-1}(W_{R,R} x_R + W_{R,S} x_S) \implies \\ & (D_{R,R} - W_{R,R}) x_R = W_{R,S} x_S \implies \\ & x_R = L_{R,R}^{-1} W_{R,S} x_S \end{aligned}$$

□

This is the opinion distribution analyzed in the context of polarization in the following sections.

REMARK The stationary opinion distribution can be equivalently we-written as

$$
\begin{aligned}
L_{RR}^{-1} W_{RS} x_S &= \\
(D_R - W_{RR})^{-1} W_{RS} x_S &= \\
(I - P_{RR})^{-1} D_R^{-1} W_{RS} x_S &= \\
(I - P_{RR})^{-1} P_{RS} x_S & \qquad (2.1)
\end{aligned}
$$

This last expression is useful in the context of computing the stationary opinions in the network by a iterative method, using the property that $\|I - P_{R,R}\|_2 < 1$.

Notably, one could define a unique opinion distribution even in the case when the set S is not globally reachable (equivalent to $L_{R,R}$ being singular) by combining the above model of DeGroot averaging for the connected components where S is not reachable. This increases the set of graphs that have a defined opinion equilibrium, but also introduces dependence on the initial opinion distribution in the network and increases the complexity of the problem since the equilibrium now needs to be handled individually for each graph sink, hence we in this work restrict ourselves to polarization of networks where this property holds.

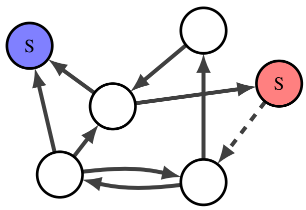

Figure 2.2 A Directed Network Where two nodes have been placed in the stubborn set, the rest of the nodes being updated by averaging dynamics. The dashed link doesn't effect the equilibrium opinions, since stubborn nodes are not updated. All links from a stubborn node doesn't effect the equilibrium opinion.

2.4 Relation to Other Problems

The Electrical Network Interpretation

The stubborn agent network model is strongly related to other problems in engineering and science. In the case of a undirected network, one can interpret the link

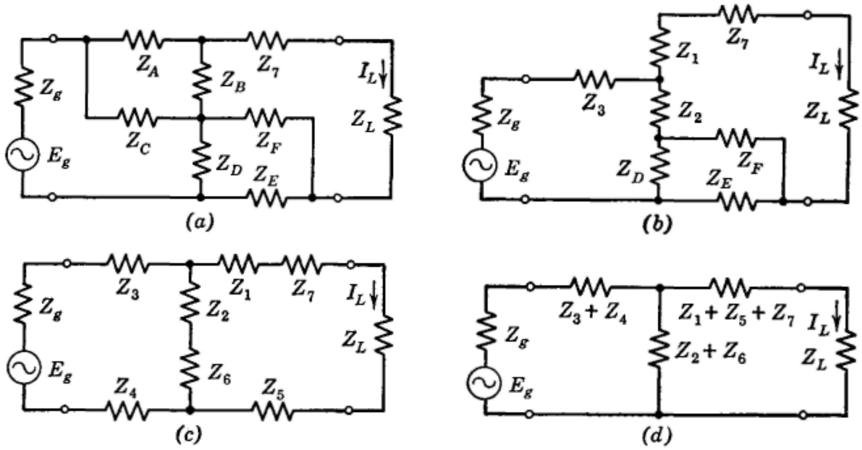

Figure 2.3 [Albert, 2011] Reduction of electrical network

weights as electrical conductance in a circuit (and thus $1/w_{i,j}$ as the electrical resistance). The potential in each (regular) node given a constant voltage in a subset of nodes (our stubborn nodes) is given by the same equilibrium as described above.

Many concepts of how electrical networks can be transformed can then - at least in the undirected case - be used in the context of opinion dynamics with stubborn nodes. One such concept is the simplification/reduction of networks, see Figure 2.3.

If we have a network G where we only care about the effect of the network when varying a handful of connections. We can then reduce the network drastically by using the rules of series and parallel resistors. However, during this process, the opinion distribution described above would change if nodes are removed as is normally done in circuit reduction.

Heat Conductance

Well known is also that the phenomenon of heat conductance exhibits averaging properties. Given Dirichlet boundary conditions on a subset of nodes (stubborn), the temperature at each regular node is the equilibrium solution found above.
Heat conductance problems are widely studied, and many methods have been created to efficiently compute stationary temperature distributions in the form of finite element methods. The defining characteristic of most of these problems is that they are locally spatially bounded, and such are not guaranteed to work well on networks where connections can be far reaching (could be compared to a heat tunnels/portals in the heat problem definition).

3

Reducing Polarization

3.1 Polarization in Networks

Polarization is defined as “Division into two sharply contrasting groups or sets of opinions or beliefs” [Lexico, n.d.] This intuitive sense of what polarization is could in a mathematical context be interpreted in many different ways. In the context of the DGS opinion dynamics model, the polarization should naturally be dependent of the stationary opinions $x \in \mathbb{R}^{|R|}$, as well as optionally on the structure of the network itself. For ease of analysis, we restrict our polarization measure to have a co-domain of $\mathbb{R}$, where high values represent high polarization and zero none.

Since the concept of polarization is not well defined as a measurement, we can only hope to find a measurement that exhibits properties that are associated with the underlying concept, while at the same time behaving nicely from a mathematical perspective. The most important properties where considered to be:

- The measurement should be continuous function $\mathscr{P} : \mathbb{R}^{|R|} \to \mathbb{R}+$.
- If all opinions are equal, the polarization should be zero.
- The measurement should take on its largest value when the opinions are devided into two sharply contrasting groups.

Some examples of opinion distributions are shown in Figure 3.1. A number of possible polarization measures for which the preferred properties hold in some context are are presented below.

Polarization Measures

One family of candidates is powers of the ℓ_p norms of the mean centered opinions. Let $\hat{x}_R = x_R - \frac{\mathbb{1}^T x_R}{|V|}\mathbb{1}$, a group of polarization measures is then $\|\hat{x}_R\|_p^n$ for any real number $p > 1, n > 0$. We only choose two cases, $p = n = 2$ and $p = \infty, n = 1$.

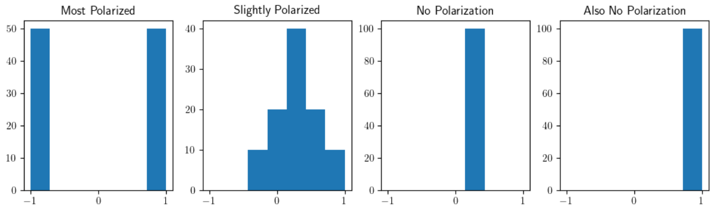

Figure 3.1 Different opinion distributions and their level of polarization. Note that the level of polarization is still minimal (0) when all opinions are centered around an extreme value.

DEFINITION 1 (VARIANCE, $p = 2$)

$$\mathscr{P}_{avg} = V[x] = \|\hat{x}_R\|_2^2 \quad \text{where} \quad \hat{x}_R = x_R - m\mathbb{1} = x_R - \frac{\mathbb{1}^T x_R}{|R|}\mathbb{1}$$

Worthy of note is that this is equivalent to the sum of differences between all opinions in the network up to a constant.

DEFINITION 2 (MAXIMUM DIFFERENCE, $p = \infty$)

$$\mathscr{P}_{max} = \max(x) - \min(x)$$

This norm maximizes the impact of a single opinion deviating from a unanimous opinion of the others.

DEFINITION 3 (DISSAGREEMENT)

$$\mathscr{D} = \sum_i \sum_j w_{i,j}(x_i - x_j)^2$$

The stationary opinions of DGS minimize the disagreement measure.

Choice of Metric

The main polarization metric used throughout will be the variance of the opinions $\mathscr{P}_{avg}$. The disagreement measure has value zero when two distinct connected components in a network have different unanimous opinions, arguably the most polarizing state possible. The maximum difference metric has the flaw that any opinion state where there exists two nodes maximally divided in opinion, the opinion of the rest of the nodes does not effect the polarization measure.

3.2 Problem Formulation

In the most general sense, we aim to minimize the polarization of opinions at stationarity in a network with a fixed set of stubborn (with known opinions) and regular nodes, given constraints on links. We choose to only study the case where the network structure is such that the set of stubborn nodes are globally reachable, as the equilibrium opinions of the regular nodes then does not depend on the initial opinion distribution.

$$\min_{W} \quad \mathscr{P}(W)$$
$$\text{subject to constraints on } W$$

where W is the weight matrix of the graph. A number of constraints are considered in the following sections. First, we present invariant properties of the problem. Then we show a number of scenarios for when the minimization problem is easy after which we show why the problem for general constraint choices is hard. Lastly, we explore approximation methods in the form of rank-1/link updates and analytic expression for the gradient for use in these harder problems.

3.3 Invariant Properties

Below follows a number of transformations for which the opinion equilibrium or polarization measure is invariant.

Of the opinion distribution

Multiplication of Rows Multiplying any row of W doesn't change the opinion equlibirum, since the opinion of a node is a weighted average of the opinions of its out-neighbors. This is easily seen in equation (2.1), since P itself is constructed by normalizing the rows.

Addition of Self Loops Adding a self loop (any constant a to a diagonal entry of W) doesn't change the stationary opinion distribution. This can be motivated both by the averaging dynamics, but can also be seen by noting that the equilibrium opinions depend on $L_{R,R}$, whose diagonal entries does not change when self loops are added.

Linear Transforms Every link from a node i to j, its weight can be linearly exchanged for a change in weight on a connection to another node $\hat{j}$. By this process, any network equilibrium can be expressed by another network where no regular nodes are connected, ie $W_{RR} = 0$ and all equilibirum opinions arise from the elements of W_{RS}. Notably the stationary opinions of the original network need to be known to do this transformation. This is analogous to the electrical network interpretation. The usefulness of this in the case of reducing polarization is only as discussed above in the case where we are only locally reducing polarization (given n links to modify) while the rest of the network is fixed. Then this fixed portion of the network can be simplified.

Of The Polarization Measure

Shift of opinions If all opinions in the network shift, the polarization doesn't change. This is equivalent to elements in W_{RS} changing by a matrix $\hat{W}_{RS}$ where $\hat{W}_{RS}x_S = a\mathbb{1}$.

3.4 Characterization of Minima

We explore for a number of constraints what the minimized solution looks like.

Without Constraints

In a complete network with equal weights, the opinion of all regular nodes is the mean of all the stubborn nodes. This follows from graph symmetry (regular nodes must have the same value) and that:

$$x_r = \frac{|R|x_r + \sum_{i\in S} x_i}{|V|} \implies (|V| - |R|)x_r = \sum_{i\in S} x_i \implies x_r = \frac{\sum_{i\in S} x_i}{|S|}$$

The entire network mean is then $m = x_r$, and the network polarization is

$$\mathscr{P}_{avg} = \sum_{i\in V} (x_i - m)^2 = 0$$

If in a graph, the only regular nodes connected to the stubborn nodes are so to all of them equally, the equilibrium is the exact same one as above. We show a more general result

PROPOSITION 1 The polarization $\mathscr{P}_{avg}$ of any graph can be minimized by at most adding $|\mathscr{E}_S|$ links, the number of links pointing to a stubborn node.

Proof Given a graph G and its interaction matrix between regular and stubborn nodes W_{RS}, we can for each row with non-zero elements (of which there exist $|\mathscr{E}_S|$) add a weight in a column such that $W_{RS}x_S = a\mathbb{1}$ for any constant a in the range of stubborn values. □

Given Stubborn-Regular Node Interaction

If the connections between the stubborn and regular nodes are fixed, the following proposition shows that if the network is not already at a minimum, a solution does not exist.

PROPOSITION 2 For a directed weighted graph G. If

- There exists at least two regular nodes whose opinion equilibrium if only connected to stubborn nodes are different.

- All edges between regular and stubborn nodes are fixed, and no restrictions apply to the edges between regular nodes.

The problem of minimizing polarization has no solution.

Proof Suppose there exists minimal polarization value for which the average opinion is m. To bring the stationary opinion of a node i to a target value m, we can only increase the link weight to another regular node. If a regular node exists with a stationary value on the other side of m, then this node is not the last node whose opinion we need to bring to m. If this is the last node, there exists no other node we can choose to connect to to bring the opinion of the node to m. □

Fixing a Larger Set of Links

The following proposition descibes a special case for which a graph problem with very specific constraints can be turned into a version of the unconstrained problem.

PROPOSITION 3 Given a network G, for any set $V_n \subset V$ such that there exists no path from nodes in V_n to nodes in $V \setminus V_n$ for which all outgoing links are fixed, the optimal network structure is found by applying proposition 1 assuming all nodes V_n stubborn.

Proof If there exists no path from nodes in V_n to nodes in $V \setminus V_n$, the opinions of nodes in V_n can not be changed by modifying links in $V \setminus V_n$, and so we can only hope to minimize the opinions within that set. This can be done the same way as if V_n where stubborn nodes. □

3.5 General Problem Properties

A Useful Relationship

To start with, we prove a useful property between the $\mathscr{P}_{avg}$ metric and the $\mathscr{P}_{zero}$ metric used in the proofs following.

DEFINITION 4 (DISTANCE TO ZERO)

$$\mathscr{P}_{zero} = \|x_R\|_2^2$$

In some cases the difference in variance between two opinion distributions can be approximated by the difference between distance to a fixed average, in the case of large networks where a small subset of the edges are being altered.

PROPOSITION 4 Given any graph $G = (V, \mathscr{E}, W)$ and a real positive number ε, set of stubborn nodes S for which a convex combination of opinions can reach the value 0, and a modification of the edges of this graph giving $\hat{G}$, it is always possible to find a graph H that contains G as a sub-graph such that when performing the same modification of the edges in the graph as in G (giving $\hat{H}$)

$$\left|\left[\mathscr{P}_{zero}(\hat{G})-\mathscr{P}_{zero}(G)\right]-\left[\mathscr{P}_{avg}(\hat{H})-\mathscr{P}_{avg}(H)\right]\right|<\varepsilon$$

Proof Create H from G in the following way: Add a regular node and attach it to the stubborn nodes in the graph in such a way that its stationary opinion is 0. Then add a chain of M nodes to this node (see Figure 3.2). As this chain gets longer and longer, the average of the regular nodes in the graph are less and less effected by changes of opinions in the original subgraph G.

$$\left[\sum_{i\in V_G}\hat{x}_i^2-\sum_{i\in V_G}x_i^2\right]-\left[\sum_{i\in V_H}(\hat{x}_i-\hat{m})^2-\sum_{i\in V_H}(x_i-m)^2\right] \qquad \square$$

We can always choose a length M such that m and $\hat{m}$ is arbitrarily close to zero, giving the above holds for any ε.

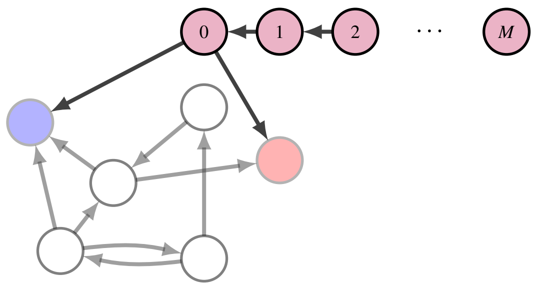

Figure 3.2 The network H, transparent section is the original network G

And so without loss of generality, we may consider $\mathscr{P}_{zero}$ instead of $\mathscr{P}_{avg}$ when showing properties of their differences.

In General Non-Convex

We remember the definition of convexity

DEFINITION 5 A function $f: X \to \mathbb{R}$ is convex if and only if $\forall x_1, x_2 \in X, t \in [0,1]$

$$f(tx_1+(1-t)x_2) \geq tf(x_1)+(1-t)f(x_2)$$

PROPOSITION 5 In a directed or undirected weighted graph, the polarization metric $\mathscr{P}_{avg}$ is not in general convex as a function of the link weights.

Proof See Figure 3.3 for choices for x_1 and x_2 for which this does not hold. Using the property described in proposition 4 the same result holds for the $\mathscr{P}_{avg}$ metric.□

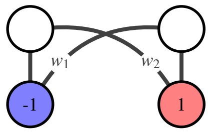

Figure 3.3 Let $w_1 = t$ and $w_2 = 100t$. Then $\mathscr{P}_{zero}$ sharply decreases until the right node passes opinion zero when increasing t. After which it increases again. Eventually when $t = 1$ the left node passes the zero opinion point, creating a second local minimum. $\mathscr{P}_{zero}$ as a function of t is $\left(\frac{100t-1}{100t+1}\right)^2 + \left(\frac{1-t}{1+t}\right)^2$, containing more than one local minimum (see Figure 3.4).

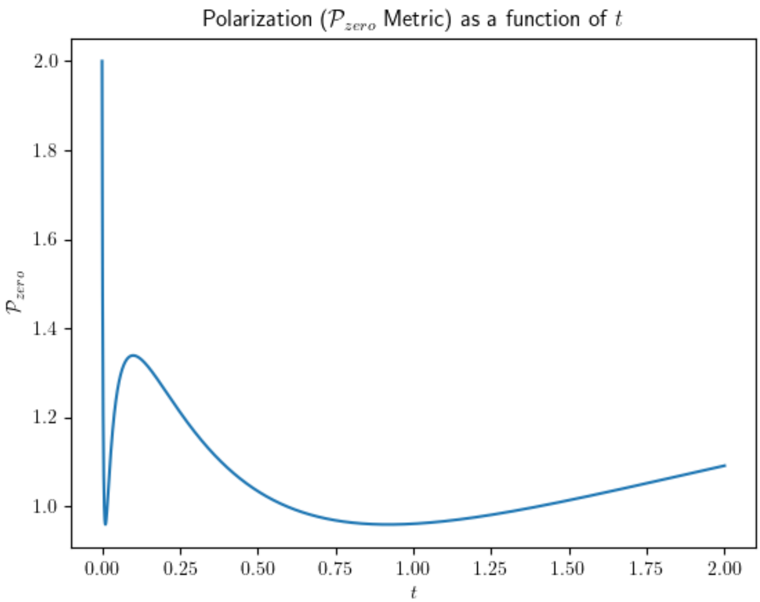

Figure 3.4 The polarization measure clearly has two local minima for the chosen parametrization, and so is not convex.

We continue by showing that this is also in general the case even if we are not allowed to touch the connections between regular and stubborn nodes.

PROPOSITION 6 In a undirected or directed weighted graph where we are only allowed to modify the links between regular nodes, the polarization metric $\mathscr{P}_{avg}$ is not in general convex.

Proof See Figure 3.5. Using the property described above the same result holds for the $\mathscr{P}_{avg}$ metric. □

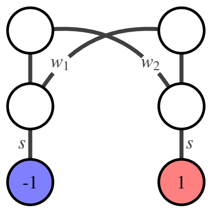

Figure 3.5 As $s \to \infty$, the change in polarization as a function of t approaches the one described in Figure 3.4.

Discrete Link Addition is NP-hard

PROPOSITION 7 Finding $m \leq k$ discrete edges that minimizes $\mathscr{P}_{avg}$ is in general NP-hard.

Proof We can construct a network for which the solution solves the well-known NP-hard subset sum problem, we use the following variation:

LEMMA 1 (SUBSET SUM PROBLEM, SSP[KLEINBERG AND TARDOS, 2006]) Let S be a set of integers. Showing it is possible to choose elements of a subset $\mathscr{E} \in S$ such that the sum equals 0 is NP-hard.

Let S be the set of stubborn nodes. Finding the set of edges that minimizes the polarization to zero metric in the network presented in Figure 3.6 solves the subset sum problem. □

Non-Submodular

The discrete equivalent of function convexity is submodularity. Some problems, even if NP-hard, have properties that give approximation guarantees. Such guarantees can be found for example for the problem of maximizing the sum of opinions [Hunter and Zaman, 2019]. Submodularity can be defined in many ways, of which the below definition is most useful to construct a intuitive example.

DEFINITION 6 (SUBMODULAR FUNCTION) The set function $f : 2^{\omega} \to \mathbb{R}$ is submodular iff for every $X \subseteq \Omega, x, y \in \Omega \setminus X$ such that $x \neq y$. $f(X \cup \{x\}) + f(X \cup \{y\}) \geq f(X \cup \{x, y\}) + f(X)$

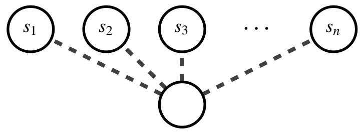

Figure 3.6 Minimizing the polarization distance to zero metric in the network gives the set of edges that places the value of the regular node closest to zero. If it could could find a solution where the opinion is zero, this is clearly the optimal. The subset sum problem is then reduced to the minimization problem by simply checking if the minimum value of the regular node is zero or not.

PROPOSITION 8 The polarization measure $\mathscr{P}_{avg}$ is not in general a submodular function.

Proof See the example in figure 3.7 and the definition of submodularity.

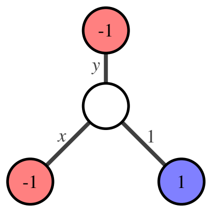

Figure 3.7 Let f be our polarization measure in the above definition of submodularity and x, y the elements not in the original set of edges. We have that $\mathscr{P}_{avg}(W \cup \{x\}) + \mathscr{P}_{avg}(W \cup \{y\}) \leq \mathscr{P}_{avg}(W \cup \{x,y\}) + \mathscr{P}_{avg}(W)$.

Invexity

A class of functions our polarization measure is part of is the class of invex functions, well summarized by Israel and Mond [Ben-Israel and Mond, 1986]. We define invexity.

DEFINITION 7 A function f from $\mathbb{R}^n$ to $\mathbb{R}$ is invex if there exist a vector-valued function $\eta(x,u)$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ for which

$$f(x) - f(u) \geq \eta(x,u) \cdot \Delta f(u)$$

We say that f is invex with respect to $\eta(x,u)$ for any $\eta(x,u)$ for which the above holds.

A very useful result is shown by [Ben-Israel and Mond, 1986].

THEOREM 3 A function is invex if and only if all its stationary points are global minima.

This we can directly apply to our polarization measure.

PROPOSITION 9 The unconstrained function $\mathscr{P}_{avg}$ is invex.

Proof Since we for any network configuration where we are not at a minimum (all opinions equal) can add weight to a link connected to one of the stubborn nodes to bring us closer to the current mean. And so the only stationary points (and minimums) are ones where all opinions are equal. □

Invexity gives guarantees that the Karush–Kuhn–Tucker conditions are sufficient for a global minimum under the condition that the constraints are also invex with respect to the same invexity function $\eta(x,u)$ [Hanson, 1999]. What this set of constraints would look like could certainly be interesting to know, but is unfortunately outside the scope of this thesis.

3.6 Approximation Methods

We start by computing the resulting change in polarization from a rank-1 perturbation of the network links (addition of a rank-1 matrix to the adjacency matrix, any two columns/rows in a rank-1 matrix are linearly dependent). This is useful in the context of gradient descent methods and single link addition for both the undirected and directed case, the complete set of derived quantities is found in Table 3.1.

Opinions After Rank-1 Pertubation

REMARK In a graph context, a rank-1 update represents adding links with weight $u_i v_j$ from node i to j, $\forall(i,j) \in \binom{A}{B}$ where A and B are two node sets.

PROPOSITION 10 The new stationary opinions after a weighted rank-1 update $\delta uv^T = \delta \begin{bmatrix} u_R & u_S \end{bmatrix} \begin{bmatrix} v_R & v_S \end{bmatrix}^T$ of W where $\sum_i v_i = 0$ is given by

$$x_R(W + \delta uv^T) = x_R(W) + \frac{\delta a}{1 - \delta c} d_R \tag{3.1}$$

where $c = v^T L_{R,R}^{-1} u_R$, $a = v^T x(W)$ and $d_R = L_{R,R}^{-1} u_R$

Proof

$$x_R(W+\delta uv^T) =$$
$$(\mathrm{diag}((W_{R,R}+\delta u_R v_R^T)\mathbb{1}+(W_{R,S}+\delta u_R v_S^T)\mathbb{1})-W_{R,R}-u_R v_R^T)^{-1}(W_{R,S}+\delta u_R v_S^T)x_S =$$
$$(\mathrm{diag}(W_{R,R}\mathbb{1}+W_{R,S}\mathbb{1}+\delta u_R\sum_i v_i)-W_{R,R}-\delta u_R v_R^T)^{-1}(W_{R,S}+\delta u_R v_S^T)x_S =$$
$$(L_{R,R}-\delta u_R v_R^T)^{-1}(W_{R,S}+\delta u_R v_S^T)x_S$$

Using the Sherman–Morrison formula we get that

$$(L_{R,R}-\delta u_R v_R^T)^{-1} = L_{R,R}^{-1} - \frac{L_{R,R}^{-1}(-\delta)u_R v_R^T L_{R,R}^{-1}}{1+v_R^T L_{R,R}^{-1}(-\delta)u_R} =$$
$$L_{R,R}^{-1} + \frac{\delta}{1-\delta v_R^T L_{R,R}^{-1}u_R} L_{R,R}^{-1}u_R v_R^T L_{R,R}^{-1}$$

Inserting the expression above yields

$$L_{R,R}^{-1}W_{R,S}x_S + \delta L_{R,R}^{-1}u_R v_S^T x_S + \frac{\delta}{1-\delta v_R^T L_{R,R}^{-1}u_R} L_{R,R}^{-1}u_R v_R^T L_{R,R}^{-1}W_{R,S}x_S$$
$$+\frac{\delta^2}{1-\delta v_R^T L_{R,R}^{-1}u_R} L_{R,R}^{-1}u_R v_R^T L_{R,R}^{-1}u_R v_S^T x_S = L_{R,R}^{-1}W_{R,S}x_S +$$
$$\frac{\delta L_{R,R}^{-1}u_R v_S^T x_S(1-\delta v_R^T L_{R,R}^{-1}u_R)+\delta^2 L_{R,R}^{-1}u_R v_R^T L_{R,R}^{-1}u_R v_S^T x_S+\delta L_{R,R}^{-1}u_R v_R^T L_{R,R}^{-1}W_{R,S}x_S}{1-\delta v_R^T L_{R,R}^{-1}u_R} =$$
$$L_{R,R}^{-1}W_{R,S}x_S + \frac{\delta}{1-\delta v_R^T L_{R,R}^{-1}u_R} L_{R,R}^{-1}u_R(v_R^T L_{R,R}^{-1}W_{R,S}x_S + v_S^T x_S) =$$
$$x_R + \frac{\delta}{1-\delta v_R^T L_{R,R}^{-1}u_R} L_{R,R}^{-1}u_R v^T x$$

Now let $c = v_R^T L_{R,R}^{-1}u_R$, $a = v^T x$ and $d_R = L_{R,R}^{-1}u_R$ and we get

$$x_R + \frac{\delta a}{1-\delta c} d_R \qquad \square$$

COROLLARY 10.1 The new stationary opinion of node k after adding the weight δ to the directed link from i to j is

$$\hat{x}_k = x_k + \frac{\delta(x_j - x_i)}{1+\delta(l_{ii}-l_{ji})} l_{ki} \quad \text{where} \quad l_{ij} = \begin{cases} \left(L_{R,R}^{-1}\right)_{ij} & \text{if } i\in R \text{ and } j \in R \\ 0 & \text{otherwise} \end{cases}$$

Proof Insert $u = e_i, v = e_j - e_i$ $\square$

COROLLARY 10.2 The new stationary opinion of node k after adding the weight δ an undirected link between i and j is

$$\hat{x}_k = x_k + \frac{\delta(x_j - x_i)}{1 + \delta(l_{ii} - l_{ji} - l_{jj} + l_{ji})}(l_{ki} - l_{kj}) \tag{3.2}$$

Proof Insert $u = e_i - e_j, v = e_j - e_i$ $\square$

COROLLARY 10.3 The change in the stationary opinion of node k of the network when adding a infinitesimal connection from i to j is for all $i, k \in R, j \in V$

$$\frac{\partial x_k}{\partial W_{ij}} = l_{ki}(x_j - x_i) \tag{3.3}$$

Proof Take the derivative and insert $\delta = 0$. $\square$

REMARK The vector $k_i = l_{:i}$ can be interpreted as a **opinion kernel** for node i, and is found by computing $L_{R,R}^{-1} e_i$ for the regular nodes and inserting 0 for the stubborn. This kernel represents the influence of a node on its neighbors (directly interpreted from the formulas above). Examples of the opinion kernels for a path graph can be seen in Figure 3.8. The maximum height of the opinion kernel k_i is always at node i, this height is larger for nodes that are further away from stubborn nodes.

COROLLARY 10.4 The change in the stationary opinion of node k of the network when adding a infinitesimal connection between i and j is for all $i, k, j \in R$

$$\frac{\partial x_k}{\partial (W_{ij} + W_{ji})} = (l_{ki} - l_{kj})(x_j - x_i) \tag{3.4}$$

Proof Same as above $\square$

Polarization after Rank-1 Pertubation

PROPOSITION 11 The change in polarization (variance) after a rank-1 pertubation uv^T of a network with magnitude δ is

$$\mathscr{P}_{avg}(W + uv^T) - \mathscr{P}_{avg}(W) = A(\delta)^2 V(d_R) + 2A(\delta) \cdot \mathrm{cov}(x_R, d_R)$$

where $d_R = L_{RR}^{-1} u_R$, $a = v^T x$, $A(\delta) = \frac{\delta a}{1 - \delta c}$ and $c = v_R^T d_R$.

Proof

$$\begin{gathered} V(x_R + Ad_R) - V(x_R) = \\ V(x_R) + 2\mathrm{cov}(x_R, Ad_R) + V(Ad_R) - V(x_R) = \\ A^2 V(d_R) + 2A\mathrm{cov}(x_R, d_R) \end{gathered}$$ $\square$

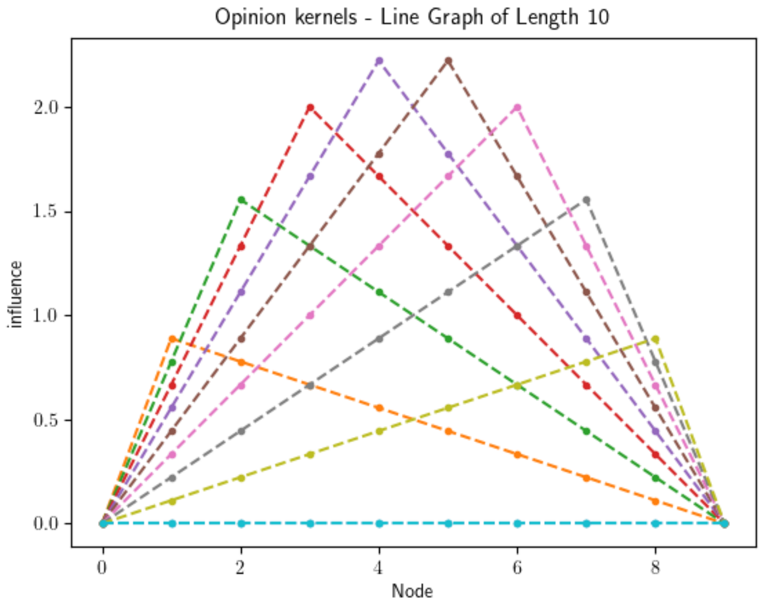

Figure 3.8 Opinion kernels of a path graph of length 10. The kernel function for a path graph is derived in the Examples chapter.

COROLLARY 11.1 For any given weighted rank-1 update, the optimal choice of δ is

$$\delta = \begin{cases} 0 & \text{if } a \cdot \mathrm{cov}(x_R, d_R) > 0 \\ \frac{-\mathrm{cov}(x_R, d_R)}{aV(d_R) - c\,\mathrm{cov}(x_R, d_R)} & \text{if } c \cdot \mathrm{cov}(x_R, d_R) < aV(d_R) \\ \infty & \text{otherwise} \end{cases}$$

for which the polarization is reduced by

$$\begin{cases} 0 & \text{if } a \cdot \mathrm{cov}(x_R, d_R) > 0 \\ \frac{a^2 \cdot \mathrm{cov}(x_R, d_R)^2}{V(d_R)} & \text{if } c \cdot \mathrm{cov}(x_R, d_R) < aV(d_R) \\ \left(\frac{a}{c}\right)^2 V(d_R) + \frac{a}{c}\mathrm{cov}(x_R, d_R) & \text{otherwise} \end{cases}$$

Proof If $a \cdot \mathrm{cov}(x_R, d_R) > 0$, the change in polarization will always be positive (since $V(d_R)$ and $A(\delta)$ are always positive), and so we should choose the link weight to be 0. If $a \cdot \mathrm{cov}(x_R, d_R) < 0$, the minimum polarization change is attained when

$$A = \frac{-\mathrm{cov}(x_R, d_R)}{V(d_R)}$$

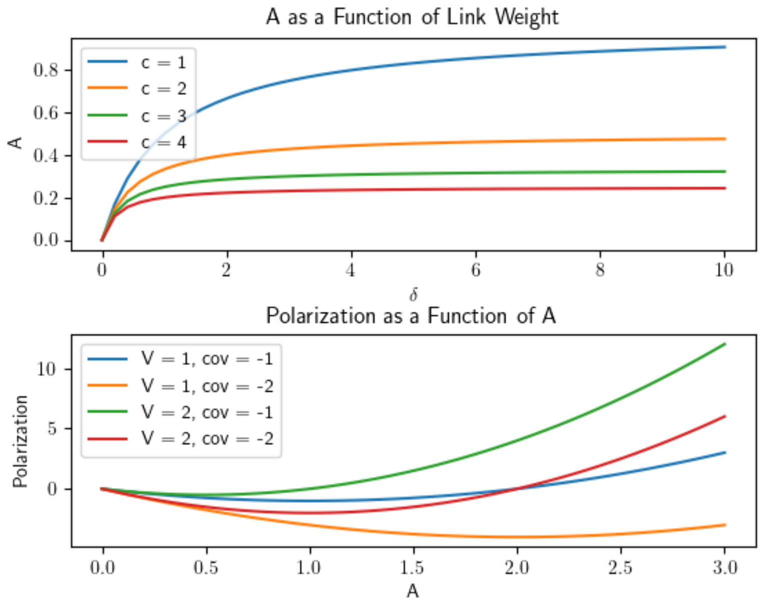

Figure 3.9 A as a function of δ and the difference in polarization as a function of A

If there exists a positive δ such that A attains this value, this is the optimum link weight.

$$\frac{\delta a}{1-\delta c} = \frac{-\mathrm{cov}(x_R, d_R)}{V(d_R)}$$
$$\delta a V(d_R) = \mathrm{cov}(x_R, d_R)(\delta c - 1)$$
$$\delta(V(d_R)a - \mathrm{cov}(x_R, d_R)c) = -\mathrm{cov}(x_R, c_R)$$
$$\delta = \frac{-\mathrm{cov}(x_R, d_R)}{V(d_R)a - \mathrm{cov}(x_R, d_R)c}$$

This quantity is positive if $V(d_R)a - \mathrm{cov}(x_R, d_R)c > 0 \implies V(d_R)a > \mathrm{cov}(x_R, d_R)c$.

If such δ does not exist, it is optimal to choose A and therefore δ as large as possible. The behavior of the quantities referenced above is shown in Figure 3.9. □

PROPOSITION 12 The derivative of $\mathscr{P}_{avg}(W)$ with respect to a rank-1 update of W is $2a \cdot \mathrm{cov}(x_R, d_R)$ where $a = v^T x$, $x_R = L_{RR}^{-1} W_{RS} x_S$ and $d_R = L_{RR}^{-1} u_R$. This can equivalently be written as

$$\frac{2}{|R|}v^T x \cdot u_R^T r \quad \text{where} \quad r = L_{R,R}^{-T} x_C$$

where $x_C = \left(I - \frac{11^T}{|R|}\right) x_R$ is the stationary opinions of the regular nodes centered around their mean.

Proof Using that $A'(0) = a$ and $A(0) = 0$ gives the first result. Rewriting the covariance gives us

$$\mathrm{cov}(x_R, d_R) = \mathrm{cov}(x_R, L_{R,R}^{-1} u_R) =$$
$$\left(\frac{x_R^T L_{R,R}^{-1} u_R}{|R|} - \frac{(\mathbb{1}^T x_R)(\mathbb{1}^T L_{R,R}^{-1} u_R)}{|R|^2}\right) =$$
$$\left(\frac{(L_{R,R}^{-T} x_R)^T u_R}{|R|} - \frac{(\mathbb{1}^T x_R)((L_{R,R}^{-T}\mathbb{1})^T u_R)}{|R|^2}\right) =$$
$$\left(\frac{(L_{R,R}^{-T} x_R)^T}{|R|} - \frac{(\mathbb{1}^T x_R)(L_{R,R}^{-T}\mathbb{1})^T}{|R|^2}\right) u_R = \left[\frac{L_{R,R}^{-T}}{|R|}\left(x_R - \frac{\mathbb{1}^T x_R}{|R|}\mathbb{1}\right)\right]^T u_R =$$
$$\frac{1}{|R|}\left[L_{R,R}^{-T} x_C\right]^T u_R = \frac{1}{|R|} u_R^T L_{R,R}^{-T} x_C$$

□

COROLLARY 12.1

$$\frac{\partial \mathscr{P}_{avg}}{\partial W_{ij}} = 2 x_C^T l_{*i}(x_j - x_i)$$

Proof Follows directly □

All the perturbations analyzed above are presented in summary in Table 3.1.

Change in opinions - rank-1 update	$x_R(W+\delta uv^T) = x_R(W) + \frac{\delta a}{1-\delta c} d_R$
Change in opinions - directed link	$\hat{x}_k = x_k + \frac{\delta(x_j - x_i)}{1+\delta(l_{ii}-l_{ji})} l_{ki}$
Change in opinions - undirected link	$\hat{x}_k = x_k + \frac{\delta(x_j - x_i)}{1+\delta(l_{ii}-l_{ji}-l_{jj}+l_{ji})}(l_{ki}-l_{kj})$
Opinion derivative - directed link	$\frac{\partial x_k}{\partial W_{ij}} = l_{ki}(x_j - x_i)$
Opinion derivative - undirected link	$\frac{\partial x_k}{\partial(W_{ij}+W_{ji})} = (l_{ki}-l_{kj})(x_j - x_i)$
Change in polarization - rank-1 update	$A(\delta)^2 V(d_R) + 2A(\delta)\cdot \mathrm{cov}(x_R, d_R)$
Optimal choice of link weight	$\delta = \begin{cases} 0 & \text{if } a\cdot\mathrm{cov}(x_R,d_R) > 0 \\ \frac{\mathrm{cov}(x_R,d_R)}{aV(d_R) - c\mathrm{cov}(x_R,d_R)} & \text{if } c\cdot\mathrm{cov}(x_R,d_R) < aV(d_R) \\ \infty & \text{otherwise} \end{cases}$
Polarization derivative - rank-1 update	$\frac{2}{\lvert R\rvert} v^T x \cdot u_R^T r$
Polarization derivative - directed link	$\frac{\partial \mathscr{P}_{avg}}{\partial W_{ij}} = 2x_C^T l_{*i}(x_j - x_i)$

Table 3.1 The derived quantities in the following chapter

4

Examples

4.1 Path graph

Kernels

The kernel of node i in a undirected unweighted path graph of length $|V|$ where each node is indexed by 0 to n can be found by solving the following system:

$$k_i = \frac{k_{i-1} + k_{i+1} + 1}{2}$$

$$k_m = \begin{cases} \frac{k_i}{i} m & \text{if } m \leq i \\ \frac{-k_i}{n-i} + k_i & \text{if } m \geq i \end{cases}$$

Derived from expanding $L_{R,R} k_R = e_i$ as described in the kernel remark in section 3.6 and using the fact that averaging dynamics on a path graph where the opinion is known on each side has a linear solution in between. The solution is

$$k_m = \begin{cases} \frac{(n-i)m}{n} & \text{if } m \leq i \\ \frac{(n-m)i}{n} & \text{if } m \geq i \end{cases}$$

Directed Link Addition

The mean of the kernel of node i corresponding to regular nodes can be computed

$$E[k_i] = \frac{1}{n-1} \left(\sum_{m=1}^{i-1} \frac{(n-i)m}{n} + \sum_{m=i}^{n-1} \frac{(n-m)i}{n} \right) = \frac{i(n-i)}{2(n-1)}$$

The mean of the elementwise squared kernel is

$$E[k_i^2] = \frac{1}{n-1}\left(\sum_{m=1}^{i-1}\left(\frac{(n-i)m}{n}\right)^2 + \sum_{m=i}^{n-1}\left(\frac{(n-m)i}{n}\right)^2\right) = \frac{(i-n)i(2i^2-2ni-1)}{6n(n-1)}$$

Using this we get that the variance of kernel i is

$$E[k_i^2] - E[k_i]^2 = \frac{i(n-i)(n^2i-(i^2+4i-2)n+4i^2-2)}{12(n-1)^2n}$$

To compute the covariance $\mathrm{cov}(x_R, k_i)$, we need the stationary opinions of a path graph. Using the same linearity property and giving the stubborn nodes on either side the value -1 and 1, we get that

$$x_m = \frac{2m}{n} - 1$$

And finally we can compute the covariance

$$\mathrm{cov}(x_R, k_i) = \frac{1}{n-1}\left(\sum_{m=1}^{i-1}\left(\frac{(n-i)m}{n} - E[k_i]\right)\left(\frac{2m}{n}-1\right) + \sum_{m=i}^{n-1}\left(\frac{(n-m)i}{n} - E[k_i]\right)\left(\frac{2m}{n}-1\right)\right) = \frac{(i-n)i(2i-n)}{6n(n-1)}$$

We restrict ourselves to add a link of weight 1 from i to j where $j < i$ without loss of generality because of line symmetry. This gives

$$A = \frac{a}{1-c} = \frac{\frac{2j}{n} - \frac{2i}{n}}{1-(\frac{(n-i)j}{n} - \frac{(n-i)i}{n})} = \frac{2(i-j)}{i^2-(j+n)i+n(j-1)}$$

Combined they give us the total change in polarization $A^2V(k_i) + A\mathrm{cov}(x_R, k_i)$ as a function of i, j and n, Sadly not simplifying to that nice of an expression. In Figure 4.1 we can see the polarization for a handful of values of n. This formula has been verified with a computational method.

4.2 Barbell Graph

Kernels

The opinion kernels for a barbell graph can be computed by observing that the only two cases with unique distributions is where the kernel node is in a side-group, or

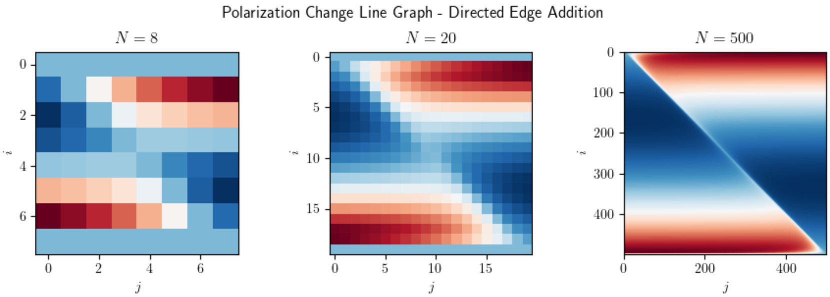

Figure 4.1 The polarization in a number of path graphs for different choices of directed link addition. Red is decreasing polarization and blue is increasing.

the node connecting to the other side (all other cases are given by symmetry). We first analyze the first case. Let x_l be the opinion of the regular nodes on the left side of the barbell who are not connected to the right, x_{lr} be the opinion of of the node on the left who is connected to the right, and x_{rl} and x_r be the equivalent opinions on the right side. Finally, choose one node i on the right side for which we compute the kernel, with stationary opinion x_i. This gives rise to the following system of equations

$$\begin{aligned}
2x_l \quad - x_{lr} \qquad\qquad\qquad\qquad\qquad\qquad &= 0 \\
(2-n)x_l + nx_{lr} + x_{rl} \qquad\qquad\qquad\qquad &= 0 \\
- x_{lr} + nx_{rl} + (3-n)x_r \qquad - x_i &= 0 \\
- x_{rl} \qquad + 3x_r \qquad - x_i &= 0 \\
- x_{rl} + (3-n)x_r + (n-1)x_i &= 1
\end{aligned}$$

With solutions

$$x_l = \frac{1}{n(n+4)}, \quad x_{lr} = \frac{2}{n(n+4)}, \quad x_{rl} = \frac{n+2}{n(n+4)} \quad x_r = \frac{n+3}{n(n+4)} \quad x_i = \frac{2n+7}{n(n+4)}$$

Now slighly modify the system, choosing i to be the node connected to the left side. We get the system

$$\begin{aligned}
2x_l \quad - x_{lr} \qquad\qquad\qquad &= 0 \\
(2-n)x_l + nx_{lr} + x_i \qquad\qquad &= 0 \\
- x_{lr} + nx_i + (4-n)x_r &= 1 \\
- x_i \qquad + 4x_r &= 0
\end{aligned}$$

With solutions

$$x_l = \frac{2}{n(n+4)}, \quad x_{lr} = \frac{4}{n(n+4)}, \quad x_i = \frac{2(n+2)}{n(n+4)}, \quad x_r = \frac{n+2}{n(n+4)}$$

The mean of the first case kernel is

$$E[k_{\text{in}}] = \frac{(n-2)x_l + x_{lr} + x_{rl} + (n-3)x_r + x_i}{2(n-1)} = \frac{1}{2(n-1)}$$

And the second happens to be equal

$$E[k_{\text{out}}] = \frac{(n-2)x_l + x_{lr} + x_i + (n-2)x_r}{2(n-1)} = \frac{1}{2(n-1)}$$

The mean of the squares in each respective case, gives us the kernel variance in each case:

$$E[k_{\text{in}}^2] = \frac{n^3+8n^2+24n+28}{2(n-1)n^2(n+4)^2} \implies V[k_{\text{in}}] = \frac{n^4+6n^3+16n^2+8n-56}{4(n-1)^2n^2(n+4)^2}$$
$$E[k_{\text{out}}^2] = \frac{(n+2)(n^2+4n+8)}{2(n-1)n^2(n+4)^2} \implies V[k_{\text{out}}] = \frac{n^4+2n^3+4n^2-32}{4(n-1)^2n^2(n+4)^2}$$

Computing the stationary opinion distribution can either be done with the kernels, or by solving a new system. We get

$$x_l = -\frac{n+2}{(n+4)}, \quad x_{lr} = -\frac{n}{(n+4)}, \quad x_{rl} = \frac{n}{(n+4)}, \quad x_r = \frac{n+2}{(n+4)}$$

The expression of covariance in the first and second case is

$$\operatorname{cov}(x_R, k_{in}) = \frac{(n^2+4n+2)n}{n(n+4)^2},$$
$$\operatorname{cov}(x_R, k_{out}) = \frac{(n-3)n(n+4)+2n^4-3n^3-7n^2+20n-12}{n(n+4)(n-1)(n+4)}$$

Finally, A can be computed, illustrated below in Table 4.2 for two cases.

l-sr	$A = \frac{a}{1-c} = \frac{1+(n+2)/(n+4)}{1-(0-(2n+7)/(n(n+4)))} = \frac{(2n+6)n}{n(n+6)+7}$
l-rl	$A = \frac{a}{1-c} = \frac{n/(n+4)+(n+2)/(n+4)}{1-(1/(n(n+4))-(2n+7)/(n(n+4)))} = \frac{2n(n+1)}{(n+6)n+6}$

5

Algorithms

In this chapter, we introduce efficient ways to compute the quantities used for choosing an optimal link recommendation.

5.1 Opinions at Stationarity

To compute the stationary opinions of the network, we need to evaluate $L_{RR}^{-1} W_{RS} x_S$. The by far most computationally expensive operation is the inversion of $L_{R,R}^{-1}$. Thankfully, we can use a different approach, namely a convenient Taylor series expansion.

As mentioned in the introducing chapters, another way to express the stationary opinions is $(I - P_{RR})^{-1} P_{RS} x_S$.

5.2 Fast Computation of $(I - P_{RR})^{-1} v$

The series

$$I - P_{RR} + P_{RR}^2 - P_{RR}^3 + P_{RR}^4 \ \dots$$

converges to $(I - P_{RR})^{-1}$ [Guth, 2017]. Iteratively computed as

$$M_0 = I, \quad M_{i+1} = I - P_{RR} M_i$$

For large sparse matrices, the inverse of $(I - P_{RR})$ will inconveniently not converge to a sparse matrix, and the matrix inverse is therefore unfeasible to both compute but also store directly (a network with 1000 nodes would already require 1 million elements). Thankfully, the product of $(I - P_{RR})^{-1}$ with any vector v can be computed iteratively as follows

$$m_0 = v, \quad m_{i+1} = v - P_{RR} m_i$$

requiring significantly less computational power and memory, with the drawback that the iteration process needs to take place for every time this product needs to be computed. If we have a batch of n vectors, we can efficiently vectorize the computation by simply replacing the vector m with a matrix M of size $|R| \times n$, and perform the same iteration, achieving a compromise between computing the entire matrix inverse and the product with a single vector at a time.

5.3 Computing the Maximum Edge Derivative

To find the largest first derivative we can use proposition 18. Notably the optimal node j to point to can be chosen purely based on the opinion of node i. The difference $(x_i - x_j)$ will always need to be either minimized or maximized, depending of the sign of the factor $2x_R^T(I - \frac{\mathbb{1}\mathbb{1}^T}{|R|})l_{*i}$. The minimum or maximum is found when j is the node with either the minimum or maximum opinion in the network. This is always one of the two stubborn nodes with minimal and maximal opinion respectively (the regular nodes' opinions are upper and lower bounded by the opinions of the stubborn nodes). We arrive at the following process:

- Compute node the weight vector $p = 2L_{R,R}^{-T}x_C$. (The matrix products are approximated by the technique mentioned in section 5.2 with β iterations)
- For each node, choose either the minimal or maximal stubborn node (based on the sign of p_i), and assemble them into the vector m.
- The largest difference in polarization when perpetuating a link is given by the largest value in the component wise product of p and m. The index of this value gives both our optimal i and j.

6

Simulations

In this chapter, we will compare two of methods of reducing polarization in the network. First we will compare a number of strategies of single link recommendation. Then we will for smaller networks perform gradient descent with a number of different link costs.

6.1 Networks

The list of networks used in the simulations are presented in Table 6.1. Some are to large to compute the optimal solutions for, but serve as a test for the speed and scalability of approximation methods. The effect of the choice of stubborn node placement, or for that matter the number of stubborn nodes was not studied.

6.2 Single Link Addition

We first explore methods for performing single discrete link addition. The following heuristics will be compared:

PROP choose a regular node based on three metrics- its distance to the average opinion (how much its opinion should be changed), its out degree (how hard the opinion is to change) and its in degree (how many opinions it effects). Connect this node to the node with an opinion the furthest away towards the network average.
ITR-β Iteratively compute the largest first derivative with respect to a link addition in the network, using β iterations.
OPT Compute the true reduction for each link.

Name	Description	Nodes	Edges	Stubborn
LINE	Undirected path graph.	20	19	Ends.
BARB	Undirected barbell graph.	20	110	Sides.
POLIT	A undirected weighted network representing the number of collaborations between Swedish parties after the 2018 election. [Nyqvist, 2018]	9	62	Most left/right leaning parties.
KARATE	A undirected weighted network containing the friendships between people in a karate club before it split into two different clubs because of ideological reasons. [Zachary, 1976]	34	78	The two dividing leaders, Mr. Hi and John A.
EMAILS	Connections between people in a large European research institution, where the connection strength is approximated by email frequency. [Leskovec et al., 2007].	1005	25571	Low out /high in degree.
EPIN	A undirected weighted network representing trust between users on a consumer review site. [Richardson et al., 2003].	75879	508837	Low out /high in degree.

Table 6.1 Network Datasets. The choice of stubborn nodes is not always obvious.

6.3 Gradient Descent

We consider classical methods of gradient descent with added link change costs. For smaller networks, this is a natural choice for finding an approximation to the minimum given soft and/or hard constraints. For large networks, it is unfeasible to consider the entire parameter space of $|V| \times |V|$ edge weights. In this case, one could instead consider optimization with or without cost with where the hard constraints are limiting enough to reduce the number of variables considerably.

ABS -α a positive or negative link change δ adds a cost of $\delta\alpha$.

SQUARE a positive or negative link change δ adds a cost of δ^2.

6.4 Results

Method	LINE	BARB	POLIT	KARATE	EMAILS	EPIN
PROP-α 0.5	-72.2%	-16.7%	2.5%	2.3%	-34.4%	-3.4%
ITR-β	-59.8%	-16.7%	14.5%	-20.6%	-34.4%	-3.4%
OPT	-73.6%	-16.7%	-22.9%	-20.6%	-	-

Method	LINE	BARB	POLIT	KARATE	EMAILS	EPIN
PROP-α 0.5	0.00523	0.00523	0.00497	0.00505	0.02771	0.94883
ITR-β	0.01008	0.00952	0.00896	0.00911	0.05874	1.26213
OPT	2.06227	2.05199	0.39599	5.88018	-	-

Table 6.2 Change in polarization when adding a single link of weight 1 for different networks and methods in the upper table, time of computation in seconds in lower. A dash signifies a too long computation time.

Method	LINE	BARB	POLIT	KARATE
ABS 1	-0.0%	0.0%	0.0%	0.0%
ABS 0.5	-25.5%	0.0%	0.0%	0.0%
ABS 0.1	-84.3%	-25.9%	0.0%	0.0%
SQUARE	-85.9%	-51.5%	-1.5%	-51.3%

Method	LINE	BARB	POLIT	KARATE
ABS 1	0.21070	0.21336	0.19985	0.23065
ABS 0.5	0.14322	0.21352	0.19581	0.22980
ABS 0.1	0.25741	0.15156	0.19254	0.23258
SQUARE	0.11313	0.73286	0.05005	0.09050

Table 6.3 Decrease in polarization achieved and time in seconds for a simple gradient descent method for a number of smaller graphs.

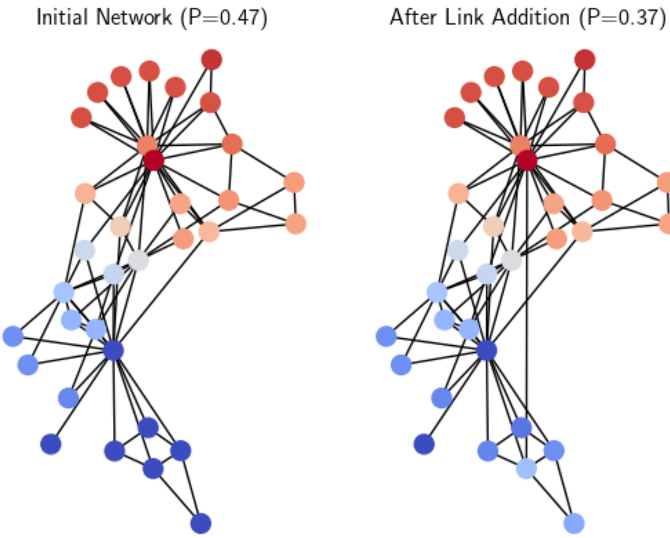

Figure 6.1 Change in polarization when adding a link using the ITR-β method in the KARATE network.

7

Conclusion

We have formalized the problem of reducing polarization in opinion networks in the context of variance on the stationary opinions of regular nodes on the DeGroot model with stubborn nodes. The problem was shown to be simple for a number of cases with basic constraints, but in general non-convex. The problem was shown to instead belong to the class of invex functions.

An iterative vectorized method to compute the approximate gradient has been proposed. This method was evaluated for adding a directed link to a network and shown to perform well for small link additions, but expectedly worse for larger ones, something further work could address.

There is an endless number of ways one could want to quantify polarization in a network context. Not only could other opinion models and measures be studied, but also different contexts altogether, some examples found in the literature mentioned in the introduction.

Within the bounds of this model, one could explore different constraints (limited out-degree of nodes, costs, total weight change) and optimization methods (specifically line search could be interesting, since we've characterized the behavior of the function along any line). A direction for future research would be to make use of (combinatorial) optimization theory to explore further properties of the problem.

Reducing polarization in a social network setting is furthermore only useful if the polarization measure itself is strongly positively or negatively correlated with other properties of the network (such as violence, individual well being and the spread of fake news). Future work could include exploring how well this and other polarization measures correlate with other societal phenomenon, giving a more clear picture of what a reduction of polarization by x percent entails.

Bibliography

Albert, A. L. (2011). *Simplification of networks*. http://www.vias.org/albert_ecomm/aec05_electric_networks_007.html. Accessed: 2022-01-14.

Amelkin, V. and A. K. Singh (2019). "Fighting opinion control in social networks via link recommendation". In: *Proc. of ACM SIGKDD Conference of Knowledge Discovery and Data Mining (KDD'19)*. ACM. Anchorage, AK, US, pp. 677–685. DOI: 10.1145/3292500.3330960.

Ben-Israel, A. and B. Mond (1986). "What is invexity?" *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics* **28**:1, pp. 1–9. DOI: 10.1017/S0334270000005142.

Cass, S. R. (1999). "The law of group polarization". *Chicago Unbound*. John M. Olin Program in Law and Economics Working Paper No. 91.

Chandrasekhar, A., H. Larreguy, and J. Xandri Antuña (2015). *Testing Models of Social Learning on Networks: Evidence from a Lab Experiment in the Field*. NBER Working Papers 21468. National Bureau of Economic Research, Inc. URL: https://EconPapers.repec.org/RePEc:nbr:nberwo:21468.

Como, G. and F. Fagnani (2016). "From local averaging to emergent global behaviors: the fundamental role of network interconnections". English. *Systems and Control Letters* **96**, pp. 70–76. ISSN: 0167-6911. DOI: 10.1016/j.sysconle.2016.02.003.

Como, G. and F. Fagnani (2021). *Lecture notes on Network Dynamics*.

DeGroot, M. H. (1974). "Reaching a consensus". *Journal of the American Statistical Association* **69**:345, pp. 118–121. ISSN: 01621459. URL: http://www.jstor.org/stable/2285509.

Guth, J. (2017). *On distributed maximization of influence in social networks*. eng. Student Paper.

Hanson, M. A. (1999). "Invexity and the kuhn–tucker theorem". *Journal of Mathematical Analysis and Applications* **236**, pp. 594–604.

Hunter, D. S. and T. Zaman (2019). *Optimizing opinions with stubborn agents under time-varying dynamics*. arXiv: 1806.11253 [cs.SI].

Jilani, Z. and J. A. Smith (2019). *What is the true cost of polarization in america?* https://greatergood.berkeley.edu/article/item/what_is_the_true_cost_of_polarization_in_america. Accessed: 2022-01-03.

Kleinberg, J. and É. Tardos (2006). *Algorithm Design*. 2nd ed., p. 491. ISBN: 0-321-37291-3.

Leskovec, J., J. Kleinberg, and C. Faloutsos (2007). *Graph evolution: densification and shrinking diameters*. URL: https://snap.stanford.edu/data/email-Eu-core.html.

Lexico (n.d.). https://www.lexico.com/definition/polarization. Accessed: 2021-10-28.

Musco, C., C. Musco, and C. E. Tsourakakis (2017). "Minimizing polarization and disagreement in social networks". *CoRR* **abs/1712.09948**. arXiv: 1712.09948. URL: http://arxiv.org/abs/1712.09948.

Nyqvist, O. (2018). *Vilka partier samarbetar med varandra?* https://www.svt.se/special/partisamarbete-i-kommunerna/. Accessed: 2021-12-29.

Pew Research Center (2014). *Political polarization in the american public*. https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public/pp-2014-06-12-polarization-0-01/. Accessed: 2022-01-13.

Richardson, M., R. Agrawal, and P. Domingos (2003). *Trust management for the semantic web*. URL: https://snap.stanford.edu/data/soc-Epinions1.html.

Zachary, W. (1976). "An information flow model for conflict and fission in small groups1". *Journal of anthropological research* **33**. DOI: 10.1086/jar.33.4.3629752.

Lund University
Department of Automatic Control
Box 118
SE-221 00 Lund Sweden

Document name
MASTER'S THESIS

Date of issue
February 2022

Document Number
TFRT-6157

Author(s)
Samuel Selleck

Supervisor
Giacomo Como, Dept. of Automatic Control, Lund University, Sweden
Emma Tegling, Dept. of Automatic Control, Lund University, Sweden (examiner)

Title and subtitle
Reducing Polarization in Opinion Networks in the Presence of Stubborn Leaders

Abstract

We study the problem of reducing polarization (variance) of opinions at stationarity in a directed weighted graph with node set divided into two groups: stubborn, initialized with a fixed opinion and regular who repeatedly update their opinion to the average of their out-neighbors, known as the DeGroot model with stubborn nodes. We show how the polarization can be minimized for a number of simple constraints, but that the problem in general is not convex. Theory is developed for the change in opinions at stationarity and the polarization measure for a rank-1 update of the network (encompassing both addition of a directed and undirected link in the network). An algorithm for gradient approximation is presented, given directly by the analytical gradient formulation and method of matrix-vector product estimation. Lastly variations of the algorithm together with other trivial methods of recommending a link are compared for a number of random and real networks.

Keywords

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280-5316

ISBN

Language
English

Number of pages
1-47

Recipient's notes

Security classification

http://www.control.lth.se/publications/