Empathetic Social Choice on Social Networks

Amirali Salehi-Abari and Craig Boutilier
Department of Computer Science
University of Toronto
{*abari,cebly*}*@cs.toronto.edu*

Abstract

Social and economic networks play a fundamental role in facilitating interactions and behaviors between individuals, businesses, and organizations. It is widely recognized that such networks can correlate behaviors (and arguably preferences) among connected agents. We introduce a model for social choice—specifically, consensus decision making—on such networks that reflects certain interdependencies among agent utilities. Specifically, we define an *empathetic social choice framework* in which agents derive utility based on both their own intrinsic preferences and the satisfaction of their neighbors. We show how this problem translates into a weighted form of classical preference aggregation (e.g., social welfare maximization or certain forms of voting), and develop effective algorithms for consensus decision making that we believe should scale to large-scale (online) social or economic networks. Preliminary experiments validate the effectiveness of our proposed algorithms.

1 Introduction

Social networks play a central role in individual interactions and decision making. Indeed, it is widely acknowledged that the behaviors [7], and to a lesser extent the preferences, of individuals connected in a social network are correlated in ways that can be explained, in part, by network structure [10, 14]. Because of this, and the increasing availability of data that allows one to infer such relationships, the study of social choice problems on social networks is one of tremendous practical import. In fact, arguably most group decision problems, whether social, corporate, or policy-oriented, involve people at least some of whom are linked via myriad social ties. However, social choice in the context of social networks is something that has received, until recently, relatively little attention. Recent work has examined, for example, the formation of (hedonic) coalitions on social networks [6, 5], and stable matching on social networks [3, 20], in which the network captures one's affinity for potential partners. The influence of social networks on voting behavior has received considerable attention in the social sciences (e.g., [1, 16, 18]), and the emergence of online social network has even spawned computational research on the mechanisms to support delegation of votes in an online network [4].

In this paper, we consider the problem of *consensus decision making on social networks*, for example, in the form of voting over some option space. Specifically, we consider the problem of selecting a single option from a set of alternatives, for some group connected by a social network—e.g., a local constituency electing a political representative, or colleagues selecting a venue for a corporate retreat. While individuals have, as usual, personal *intrinsic utility* over the option space, we also incorporate a novel form of *empathetic utility* on social networks: in our model, the utility (or satisfaction) of an individual with a winning alternative a is a function of both her intrinsic utility for a and her *empathetic utility* for the "happiness" of her neighbors in the network. This use of empathetic utility can be seen as reflecting recent findings that suggest a person's happiness is influenced by the happiness of others with whom they are connected [11].

We consider two varieties of empathetic preference. In the first, the *local empathetic* *model*, the utility of individual i for alternative a combines her intrinsic preference for a with the *intrinsic* preference of i's neighbors for a, where the weight given to the preference of any neighbor j depends on the strength of the relationship between i and j. For instance, in selecting a restaurant, i may be willing to sacrifice some of her own intrinsic preference for the chosen restaurant if her colleagues are happier with the cuisine, and she defers more strongly to her closest friends. In the second, *global empathetic model*, i's utility for a depends on her intrinsic preference and the *total utility* that her neighbors have for a (not just their intrinsic preference). In other words, she doesn't just want her neighbors to be satisfied with a, she wants them to have high utility, which depends on the utility of *their* neighbors, and so on. For example, in voting for a political candidate, i may have a mild preference for a over b, but if b is strongly preferred by not only her closest neighbors, but also by their neighbors and many others in the community, she might prefer to see b elected so she won't have to interact with grumpy neighbors for the next five years.

Our main contributions in this paper are to develop a model for preference aggregation (e.g., certain forms of voting) that select consensus alternatives in a way that is sensitive to both intrinsic and empathetic preferences. Of course, we don't expect voters to actually compute such combined preferences; indeed, they may not have direct knowledge of the preferences of their neighbors. Instead voters specify their preferences for options and for the satisfaction of their neighbors (the latter could be inferred or estimated directly from the social network in some settings). We then propose methods for computing optimal alternatives under both the local and the global models. The former, unsurprisingly, corresponds to a simple form of weighted preference aggregation or weighted voting in which each voter implicitly "delegates" a portion of her vote to her neighbors. The latter, because individual utilities are co-dependent—indeed, utility spreads throughout the network much like PageRank values—requires the solution of a linear system to determine the optimal (fixed-point) option for the group. We describe (mild) conditions under which a fixed point is guaranteed to exist, and show that it too results in a form of weighted voting, where the weights assigned to each voter's intrinsic preference is readily derived from the solution to this linear system. Experiments explore various properties of our model and algorithms.

2 Social Empathetic Model

We begin by outlining our basic social choice model, motivating two notions of empathetic preference on social networks, and then defining socially optimal outcomes within this model. We also briefly discuss related work.

2.1 The Social Choice Setting

Apart from empathetic preferences on a social network, which we specify below, the choice framework we adopt is standard. We assume a set of alternatives $\mathcal{A} = \{a_1, \ldots, a_m\}$ and a set of agents $\mathcal{N} = \{1, \ldots, n\}$. Each agent j has *intrinsic preferences* over $\mathcal{A}$ in the form of either a (strict) preference ranking $\succ_j^I$ or a utility function u_j^I. For ease of presentation, we describe preferences in terms of utility functions, but discuss below on how to interpret voting procedures within our model. For example, in our experiments we use simple utility functions based on rankings of alternatives and score-based voting rules (specifically, Borda and plurality) to define "utility" for alternatives.

Our goal is to select a single consensus alternative $a^* \in \mathcal{A}$ that implements some social choice function f relative to the preferences of $\mathcal{N}$. For example, if agents' utilities were dictated solely by intrinsic preference and f were (utilitarian) social welfare, we would select $a^* = \arg\max \sum_j u_j^I(a)$. If preferences were given by intrinsic preference rankings, f

would typically be represented by some voting rule (e.g., plurality or Borda).[1]

2.2 Empathetic Preference on Social Networks

We depart now from the typical social choice framework by considering *empathetic preferences*, in which the preferences of one agent are dependent on those of others. We consider the specific case in which these influences are induced by connections in a social network (though the notion of empathetic preference need not be confined to networks). We focus on agent utility functions rather than preference rankings, since these allow the straightforward expression of quantitative tradeoffs between intrinsic and empathetic preference.[2]

Before discussing additional motivation, we introduce our model and notation. We assume a directed weighted graph $G = (\mathcal{N}, E)$ over agents, with an edge jk indicating that j's utility is dependent (in a way to be specified below) on its neighbor k's preference, the strength of this dependence given by edge weight w_{jk}. Naturally j's utility will usually depend on its own intrinsic preferences, so loops jj will usually be present. We assume that $w_{jk} \geq 0$ for any edge jk, and that $\sum_k w_{jk} = 1$ for any j (though allowing variable weightings to reflect, say, weighted voting schemes is also possible). For convenience, we treat missing edges as if they had weight zero (and vice versa). Thus, we represent the graph with a weight matrix $\mathbf{W} = [w_{ij}]$. We generally think of these edges as corresponding directly to some relationship in a social network, or possibly induced from such relationships. See Fig. 1(a) for an illustration.

We take j's utility for a to be a linear combination of it's own intrinsic preference for a and the empathetic preference derived from each of its neighbors—recall that we consider pure consensus/single-winner voting scenarios in which a single option a is selected for all $j \in \mathcal{N}$—where network weights determine the relative importance of each.[3] Letting $e_{jk}(a)$ denote the *empathetic utility derived by j from k*, we define j's utility $u_j(a)$ to be

$$u_j(a) = w_{jj} u_j^I(a) + \sum_{k \neq j} w_{jk} e_{jk}(a).$$

The ratio of w_{jj} to $\sum_{k \neq j} w_{jk}$ captures the relative importance of intrinsic and empathetic utility to j.

We consider two ways in which to define empathetic preferences e_{jk}. In the *local empathetic model*, we simply define $e_{jk}(a) = u_k^I(a)$; in other words, j's utility for a is simply a linear combination of intrinsic utilities of j's neighbor (including it's own):

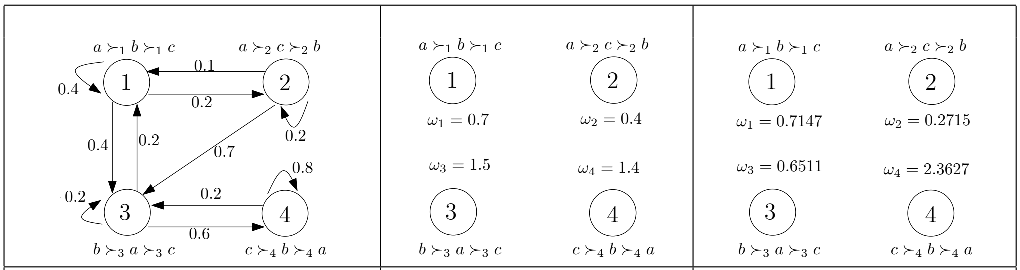
$$u_j(a) = \sum_k w_{jk} u_k^I(a). \tag{1}$$

This model captures the fact that an agent j is concerned about the "direct" preference of a neighbor k for alternative a; but the fact that k's utility may depend on k's *own* neighbors does not impact j. For instance, consider a family or a group of friends deciding on a movie (or restaurant or outing): the preferences of certain family members (e.g., parents) for a specific film may depend on the preferences of others (e.g., children, whom they want to be entertained by the choice of film).

[1] Our model below applies directly to more general social choice problems, such as assignment/segmentation problems with network externalities (where individuals may be assigned different alternatives), matching problems, and so on, without difficulty. Our algorithms are, however, specific to the "single-choice" assumption.

[2] Suitable *qualititative* expression of such tradeoffs is an important ongoing research direction.

[3] More general non-linear models are possible as well.



(a) social network with preferences | (b) local empathetic weights | (c) global empathetic weights

Figure 1: Social network and ranked preferences, with weights under the local and global empathetic model. Using Borda or plurality-based utility, the consensus winner is different in each model: a under intrinsic; b under local empathetic; c under global empathetic.

In the *global empathetic model*, we define $e_{jk}(a) = u_k(a)$, so that k's complete utility for a—which *may depend on k's own neighbors*—influences j's utility for a, giving rise to

$$u_j(a) = w_{jj}u_j^I(a) + \sum_{k \neq j} w_{jk}u_k(a). \tag{2}$$

In this model, j's utility for a depends on the utility (not just intrinsic preference) of its neighbors for a. For example, a voter may care about the overall level of satisfaction of her neighbors when voting for a political representative, but recognize that there is a larger societal effect at work, where their satisfaction also depends on their neighbors, etc. More concretely, companies linked in complex supply chain may well care about the overall success of their suppliers and customers, and consider adopting industry-specific or economic policies in that light. In the global model, the circular dependence of utilities requires a fixed point solution to the linear system defined by Eq. 2 (see below).

Correlations of behavior and/or preferences among agents connected in social network is widely accepted, and can be explained by a variety of mechanisms [10, 14]. Among these are: *technology/information diffusion*, in which agents become aware of opportunities or innovations from connections to their neighbors; *network externalities*, in which the benefits of adopting some behavior increase with the number of neighbors doing the same; or *homophily*, in which people with similar characteristics (say, preferences) more readily form social ties. Our empathetic model is somewhat different in that a person's intrinsic preferences over options $\mathcal{A}$ are not presumed to be correlated with their neighbors, but their revealed preferences for $\mathcal{A}$ might be: their choices (or stated utilities) will generally reflect some consideration, however estimated, of their neighbors' preferences as well.

2.3 Social Welfare as Weighted Intrinsic Utilities

In realistic social choice situations, agents with empathetic preferences must often perform sophisticated reasoning not only about their intrinsic preferences for alternatives, but also about those of their neighbors. Thus, even in the local empathetic setting, expressing preferences (e.g., voting) is difficult since agents usually have incomplete (and in some cases, no) information about the preferences of their friends, neighbors, or colleagues. The global empathetic setting is even more complex, since an agent is further required to reason about her neighbors' network connections as well as their intrinsic/empathetic tradeoffs.

In our models, preference aggregation and optimization can be performed by simply having agents specify their *intrinsic preferences*, as is standard in social choice, and the *weights* they assign to neighbors in their local network. In social scenarios, this can remove a considerable informational and cognitive burden from agents who might otherwise be required to determine their total utility for alternatives. In other situations, agents might not wish to reveal their preferences to their neighbors, but might still want their neighbors to obtain a favorable result (consider, for example, a collection of companies, voting over some economic policy alternatives, that are linked together in complex supply chain relationships which correlates their stability or profitability). It turns out that, given a known network G, the problem of consensus decision making with empathetic preferences can be recast as a *weighted preference aggregation problem over intrinsic preferences alone.* Not only does this ease the burden on agents, it also allows one to recast the problem as one of simple weighted voting, or of weighted (utilitarian) social welfare maximization, rendering the decision making process itself fully transparent. Here we focus on social welfare maximization.

For the local model, determining the weights associated with each agents' intrinsic preference is straightforward. Assume network weights $\mathbf{W}$. Let $\mathbf{u}(a)$ be the n-vector of agent utilities to be computed as a function of the corresponding vector $\mathbf{u}^I(a)$ of intrinsic utilities for some fixed alternative a. By Eq. 1, we have $\mathbf{u}(a) = \mathbf{W}\mathbf{u}^I(a)$. Then letting $\boldsymbol{\omega} = \mathbf{e}^\top \mathbf{W}$ (where $\mathbf{e}$ is a vector of ones), the social welfare of any alternative a under the local empathetic model is given by

$$sw_l(a, \mathbf{u}^I) = \boldsymbol{\omega}^\top \mathbf{u}^I(a). \tag{3}$$

Thus social welfare maximization under local empathetic utility is simply weighted maximization of intrinsic preference, where the weight of j's intrinsic utility ω_j is simply the sum of the weights of its incoming edges.

Fig. 1(b) illustrates the local model in action. The derived weights for each agent are shown. We assume preference rankings, and suppose utilities are dervied from these using either Borda or plurality scores. We see that the decision can be different under the local model than using voting based on intrinsic preferences along (a wins in the intrinsic model, while b wins in the local model). Indeed, using score-based voting rules, we can readily interpret this model as a form of *empathetic voting*, where the weight one assigns to a neighbor can be interpreted as the extent to which one would sacrifice one's own preferences to improve a neighbor's intrinsic satisfaction with the winning alternative.

Things are slightly more subtle in the global empathetic model. Computing the utility vector $\mathbf{u}(a)$ for alternative a requires solving a linear system to compute the fixed point of Eq. 2. Unfortunately, a unique solution is not guaranteed to exist.[4] However, in addition to our assumptions above of *non-negativity* (i.e., $\mathbf{W} \geq \mathbf{0}$) and *normalization* (i.e., $\sum_k w_{jk} = 1$ for all j), a third mild condition on the social network (weight matrix $\mathbf{W}$) is sufficient to ensure a unique fixed point solution, namely, *positive self-loop*: $w_{jj} > 0$ for all j. Let $\mathbf{D}$ be the $n \times n$ diagonal matrix with $d_{jj} = w_{jj}$. We can write Eq. 2 as

$$\mathbf{u}(a) = (\mathbf{W} - \mathbf{D})\mathbf{u}(a) + \mathbf{D}\mathbf{u}^I(a). \tag{4}$$

As a consequence,

Theorem 2.1 (Fixed-point Utility) *Assuming nonnegativity, normalizaton, and positive self-loop, Eq. 4 has a unique fixed-point solution* $\mathbf{u}(a) = (\mathbf{I} - \mathbf{W} + \mathbf{D})^{-1}\mathbf{D}\mathbf{u}^I(a)$.

(Proofs of all results are included in the appendix.) As in the local model, social welfare maximization in the global model can be interpreted as weighted maximization of intrinsic preference (though with a less straightforward interpretation):

Corollary 2.1 *In the global empathetic model, social welfare of alternative a is given by* $sw(a, \mathbf{u}^I) = \boldsymbol{\omega}^\top \mathbf{u}^I$ *where* $\boldsymbol{\omega}^\top = \mathbf{e}^\top(\mathbf{I} - \mathbf{W} + \mathbf{D})^{-1}\mathbf{D}$.

[4]Consider two individuals j and k, with $w_{jj} = w_{kk} = 0$, $w_{jk} = w_{kj} = 1$, $u^I_j(a) = 0.1$, and $u^I_k(a) = 1$. The induced system does not have a unique fixed-point solution.

Once again, in (score-rule based) voting contexts, one can interpret the global empathetic model as trading off one's own satisfaction with a winning alternative with the "overall" (not just intrinsic) satisfaction of one's neighbors: see Fig. 1(c) for an illustration. We discuss weight computation in Sec. 3.

2.4 Related Work

We are unaware of other formal models which consider the dependency between agent utilities in a social network using the type of empathetic utility we introduce above. However, empathetic utilities might be viewed as a form of network externality in an agent's utility function, though unlike typical models of externalities, an agent's utility depends on the (latent) utility of its neighbors for the chosen alternative rather than the behavior of, or allocation made to (at least directly), her neighbors (or others).

Decision making in the presence of network externalities has recently attracted attention. Bodine-Baron *et al.* [3] study stable matchings (e.g., of students to residences) with *peer effects*: these local network externalities reflect the fact that students prefer to be assigned to the same residence as their friends in a social network. Brânzei and Larson address coalition formation on social networks in two different settings: (a) agent utility for a coalition depends on its affinity weights with others in the coalition [5]; and (b) agent utility depends on her distance to others on the induced social network [6]. The problem of auction design in social networks with positive network externalities is studied in [13].

Boldi *et al.* [4] consider voting on social networks, describing a form of *delegative democracy* in which an individual can either express her preferences directly, or to delegate her vote to a proxy from among her neighbors. In our model, individuals are not asked to delegate their votes or preferences: we simply consider the dependency of their preferences on those of others, though this can be viewed loosely as *implicit, partial* delegation of preferences.

3 Computing Winners in the Empathetic Models

We now consider the question of computing the social welfare maximizing alternative in both the local and global empathetic models. In Sec. 2.3, we observed that—for both the local and global empathetic models—social welfare can be expressed as $sw(a, \mathbf{u}^I) = \boldsymbol{\omega}^\top \mathbf{u}^I(a)$ for an appropriate weight vector $\boldsymbol{\omega}$. Given the vectors $\mathbf{u}^I(a)$ for any $a \in \mathcal{A}$, we can readily compute the optimal alternative $a^* = \arg\max_{a \in A} \boldsymbol{\omega}^\top \mathbf{u}^I(a)$, requiring $O(nm)$ time. Of course, this presupposes access to $\boldsymbol{\omega}$, which has different meanings in each model, and hence requires different approaches for its computation. In the global model, this suggests a different method for computing a^* as well, without (necessarily) requiring the full computation of $\boldsymbol{\omega}$.

We first consider the local model, where $\boldsymbol{\omega}^\top$ can be calculated easily with a single vector-matrix multiplication, $\boldsymbol{\omega}^\top = \mathbf{e}^\top \mathbf{W}$, in time $O(n^2)$. However, social networks are generally extremely sparse, with the number of outgoing edges associated with any node j in the graph bounded by some small constant c which is independent of the network size (generally, social networks, while potentially locally dense, are sparse in a global sense). In sparse networks, $\boldsymbol{\omega}$ can be computed much more efficiently: ω_j is simply the sum of j's outgoing edges weights. If the outgoing neighbors of any node are bounded by a constant, $\boldsymbol{\omega}$ can be computed in $O(n)$ time and a^* can be determined in the straightforward fashion mentioned above in $O(nm)$ time. Thus the complexity of computing optimal alternatives in the local empathetic model is no different than that of straightforward social welfare maximization of straightforward (e.g., scoring rule-based) voting.

In the global model, $\boldsymbol{\omega}^\top$ has a more complicated expression, $\boldsymbol{\omega}^\top = \mathbf{e}^\top \mathbf{A}^{-1} \mathbf{D}$ where $\mathbf{A} = \mathbf{I} - \mathbf{W} + \mathbf{D}$ (see Cor. 2.1). The difficulty lies largely in matrix inversion: $\mathbf{A}^{-1}$ can be computed via Gauss-Jordan elimination, which has complexity $O(n^3)$. This implies that straightforward computation of a^* requires $O(n^3 + nm)$ time. In general, matrix inversion is no harder than matrix multiplication (see, e.g., [9, Thm. 28.2]). Although efficient matrix multiplication is the topic of ongoing research (e.g., [8]), its complexity cannot be less than $O(n^2)$ since all n^2 entries must be computed. Therefore, straightforward computation of a^* in the global model cannot have complexity less than $O(n^2 + nm)$.

We expect n to be extremely large in at least some social choice problems on social networks, e.g., in the tens of thousands (number of people in a small town), the millions (large big cities), or hundreds of millions (large country, number of Facebook or Twitter users). This makes algorithms that scale more than linearly in n problematic, both from the perspective of time and memory. Of course, many iterative methods have been proposed for matrix inversion and solving linear systems (e.g., Jacobi, Gauss-Siedel) which have $O(n)$ complexity (in non-sparse systems) per iteration and tend to converge very quickly in practice. We now briefly describe the use of a standard Jacobi method for computing a^* in the global model. We first show how to compute the utility vector $\mathbf{u}(a)$ for each alternative a, and then propose an algorithm called *iterated candidate elimination (ICE)* that will compute the optimal a^* without (necessarily) computing each $\mathbf{u}(a)$ fully.

Consider first a simple iterative method for computing $\mathbf{u}(a)$. Let $\mathbf{u}^{(t)}(a)$ be the estimated utilities for alternative a after t iterations.

Theorem 3.1 *Consider the following iteration:*

$$\mathbf{u}^{(t+1)}(a) = (\mathbf{W} - \mathbf{D})\mathbf{u}^{(t)}(a) + \mathbf{D}\mathbf{u}^I(a).$$

Assuming nonnegativity, normalizaton, and positive self-loop, this method converges to $\mathbf{u}(a)$*, the solution to Eq. 4.*

For each $j \in \mathcal{N}$, the method computes:

$$u_j^{(t+1)}(a) = w_{jj} u^I(a) + \sum_{k \neq j} w_{jk} u_k^{(t)}(a). \tag{5}$$

We can interpret $u_j^{(t)}(a)$ as agent j's estimated utility for alternative a after t iterations. This updating scheme has a natural interpretation in terms of agent behavior: suppose that each individual is able to repeatedly observe her friends' revealed utilities, and updates her own utility for various alternatives in response. This process will eventually converge (this is true even if the updates are "asynchronous"). One can readily bound the error in the estimated utilities at the t^{th} iteration:

Theorem 3.2 *In the iterative scheme above,*

$$\left\| \mathbf{u}(a) - \mathbf{u}^{(t)}(a) \right\|_\infty \leq (1 - \tilde{w})^t \left\| \mathbf{u}(a) - \mathbf{u}^{(0)}(a) \right\|_\infty ,$$

where $\tilde{w} = \min_{1 \leq i \leq n} w_{ii}$*.*

Hence, societies in which individuals have self-loops with relatively larger weight (i.e., less empathy) converge to fixed-point utilities faster societies with greater empathy (our empirical results below support this).

This error bound allows one to bound the error in estimated social welfare if the utilities of all alternatives are estimated in this fashion. Let $sw^{(t)}(a) = \sum_j u_j^{(t)}(a)$.

Theorem 3.3 *Assume* $u_j^I(a) \in [c, d]$ *and* $u_j^{(0)}(a) \in [c, d]$*, for all* $j \in \mathcal{N}$*. Under the conditions above, for any* t*:* $\left| sw(a) - sw^{(t)}(a) \right| \leq n(d - c)\,(1 - \tilde{w})^t$*, where* $\tilde{w} = \min_{1 \leq i \leq n} w_{ii}$*.*

As a result, we know that (under the same assumptions):

Proposition 3.4 *If* $sw^{(t)}(b) - sw^{(t)}(a) \geq 2n(d-c)\,(1-\tilde{w})^t$ *then* $sw(b) > sw(a)$.

We can exploit Prop. 3.4 in a simple iterative algorithm for computing a^* we call *iterated candidate elimination (ICE).* The intuition behind ICE is to iteratively update the estimated utilities of the subset $C \subset \mathcal{A}$ of candidates that are non-dominated, and gradually prune away any candidate that is dominated by another until only one, a^*, remains. Roughly, ICE first initializes $C = \mathcal{A}$ and $u_j^{(0)}(a) = c$ for all $j \in \mathcal{N}, a \in \mathcal{A}$. An iteration of ICE consists of: (1) updating estimated utilities using Eq. 5 for all j and $a \in C$; (2) computing estimated social welfare of each $a \in C$; (3) determining the maximum estimated social welfare $\hat{sw}^{(t)}$; (4) testing each $a \in C$ for domination, i.e., $\hat{sw}^{(t)} - sw^{(t)}(a) \geq 2n(d-c)\,(1-\tilde{w})^t$; and (5) eliminating all dominated candidates from C. The algorithm terminates when only one candidate (i.e., a^*) remains in C (the pseudo-code for the algorithm is provided in the appendix). The running time of ICE is at most $O(tmn^2)$ where t is the number of iterations required. More precisely, ICE runs in $O(tm|E|)$ time; and if the number of outgoing edges is bounded, $O(tmn)$. Our hope is that in practice, the methods converges in relatively few iterations, a fact indeed borne out in our preliminary experiments below. ICE also provides a natural means of approximation in large problems.

4 Empirical Results

We now describe some preliminary experiments on randomly generated networks and intrinsic preferences designed to test the differences in the decisions that result under standard non-empathetic, local empathetic and global empathetic models, the impact of these decisions on different agents, and the performance of the ICE algorithm.

Experimental Setup. Our test scenarios require generation of intrinsic preferences and a social network. We assume that individual intrinsic utilities arise from an underlying preference ordering over $\mathcal{A}$. In all experiments, we assume $m = 5$ or $m = 10$ alternatives, and draw a random preference ordering for each agent j under the impartial culture assumption (all permutations are equally likely). For simplicity, and to draw connections to voting on social networks, we assume j's utility is given by the Borda or plurality score of the alternative in its ranking. If treating these strictly as utility, they embody very different, extreme assumptions: Borda treats utility differences as smooth and linear, whereas plurality views utility in a more "all or nothing fashion."

We generate random social networks using a *preferential attachment* model for scale-free networks [2] (this is only one of many models that can be adopted). The model works in the following iterative fashion: start with n_0 initial nodes; we repeatedly add nodes (until we have a graph with n nodes), where each new node added is connected to $k \leq n_0$ existing nodes, and an existing node i is selected as a neighbor with probability $P_i = \frac{\deg(i)}{\sum_j \deg(j)}$. We set $n_0 = 2$ and $k = 1$ or $k = 2$ in all our experiments. We then convert the resulting undirected graph to an directed graph by replacing each undirected edge with the two corresponding directed edges; add a self-loop to each node with weight α; then add normalized weights to all other edges (all outgoing edges from j excluding the self loop have equal weights that sum to $1-\alpha$). The parameter $\alpha \in (0,1]$ represents the degree of self-interest, and $1-\alpha$ the degree of empathy in the society.

Performance Metrics. To measure whether the different models result in difference decisions, we assume the agents *actual utility model* is one of intrinsic (non-empathetic), local or global. We then consider making decisions using any of these models as an *assumed utility model*, and measure the effect on actual utility (e.g., global empathetic utility) of making a decision using the assumed model (e.g., intrinsic). Since decisions might be different in

Actual Utility	Assumed Utility			WSWL
	intrinsic	local	global	
intrinsic	0%(0%)	1.45%(9.95%)	1.10%(8.00%)	5.59%(14.63%)
local	2.95%(19.28%)	0%(0%)	0.09%(3.21%)	11.22%(25.10%)
global	1.78%(12.73%)	0.07%(2.73%)	0%(0%)	9.01%(20.97%)

Table 1: Average (maximum) RSWL and WSWL: 2500 runs, Borda scoring, $m = 5$, $n = 1000$, $k = 1$, $\alpha = 0.25$.

Actual Utility	Assumed Utility		
	intrinsic	local	global
intrinsic	0.0%(0.0%)	28.4%(100.0%)	22.6%(100.0%)
local	28.5%(100.0%)	0.0%(0.0%)	1.2%(86.9%)
global	22.3%(100.0%)	1.1%(97.0%)	0.0%(0.0%)

Table 2: Average (maximum) NSWL: 2500 runs, Borda, $m = 5$, $n = 1000$, $k = 1$, $\alpha = 0.25$.

each case, we measure the loss in social welfare due to making a decision using the incorrect model. Let $sw^{ac}(\cdot)$ and $sw^{as}(\cdot)$ be social welfare under the actual and assumed models, respectively, and a_w and a_s be the corresponding optimal alternatives (or winners). We define *relative social welfare loss (RSWL)* to be $L(as, ac) = [sw^{ac}(a_w) - sw^{ac}(a_s)]/sw^{ac}(a_w)$ (we sometimes report it as a percentage). RSWL has a lower bound that is independent of the assumed model: let the alternative a^- have *minimum social welfare* under the actual model (so it is no better than the decision under the assumed model). *Worst-case social welfare loss (WSWL)* is defined as $W(ac) = [sw^{ac}(a_w) - sw^{ac}(a^-)]/sw^{ac}(a_w)$. Finally, it usually makes sense to normalize RSWL by considering the range of possible social welfare values actually attainable: *normalized social welfare loss (NSWL)* is simply $N(as, ac) = [sw^{ac}(a_w) - sw^{ac}(a_s)]/[sw^{ac}(a_w) - sw^{ac}(a^-)]$. This offers a more realistic picture of loss due to using an incorrect assumed utility model (by comparing it to the loss associated with making the *worst possible decision* under the actual model).

Social Welfare Loss. We first consider RSWL, WSWL and NSWL for all nine combinations of assumed and actual utility models. We fix $\alpha = 0.25$, $n = 1000$, $m = 5$, $k = 1$, and the scoring rule to Borda. We generate 50 random networks, and for each generate 50 intrinsic utility profiles (2500 problem instances), and compute RSWL and WSWL. Average (with maximum in parentheses) RSWL for various combinations of actual and assumed models is reported in Table 1 as are average (maximum) WSWL. Maximum RSWL is more than 19% and 12% when intrinsic utility is assumed but actual utility is local or global, respectively. Moreover, we can see that global vs. local and local vs. global are quite close. Notice that average differences are quite slight: this is because the impartial culture model, in essence, renders are alternatives quite close in terms of Borda or plurality score. By normalizing for the fact that most decisions are reasonably good, we get a more accurate picture of the loss incurred by using non-empathetic voting. NSWL is reported in Table 2, which shows that making the wrong assumptions can be quite damaging; e.g., the intrinsic model loses 22.3–28.5% of empathetic social welfare on average.

Since impartial culture is generally viewed as an unrealistic model of real-world preferences, we also tested our methods using preferences drawn from 2002 Irish electoral data from the Dublin West constituency, with 9 candidates and 29,989 ballots of top-t form, of which 3800 are complete rankings.[5] Generating 1000-node networks as above, we randomly assign full rankings to nodes from this set of 3800 complete rankings. Results on RSWL and WSWL for plurality scoring are shown in Table 3. As above, average RWSL is slight; but the maximum values show significant social welfare loss in certain instances, especially when using the intrinsic model to make decisions for empathetic preferences.

[5] See www.dublincountyreturningofficer.com.

Actual Utility	Assumed Utility			WSWL
	intrinsic	local	global	
intrinsic	0%(0%)	1.82%(33.62%)	1.30%(19.55%)	97.22%(99.62%)
local	2.64%(39.25%)	0%(0%)	0.01%(6.80%)	97.28%(99.85%)
global	1.53%(31.26%)	0.10%(8.4%)	0%(0%)	97.24%(99.77%)

Table 3: Average (maximum) RSWL and WSWL for West Dublin data set: 2500 runs, plurality scoring, $m = 9$, $n = 1000$, $k = 1$, $\alpha = 0.25$.

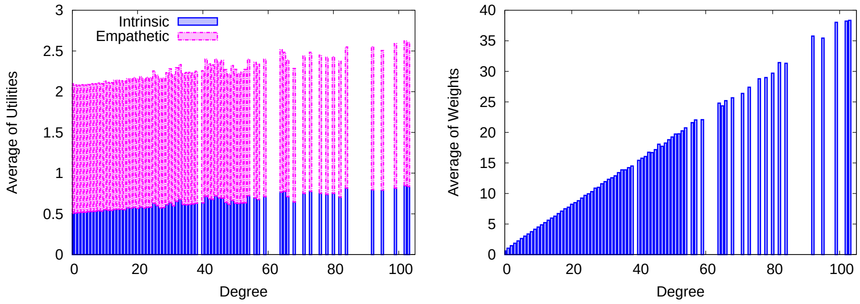


Figure 2: Average (a) instrinsic and empathetic utilities and (b) individual weights as function of node degree (global model, 2500 runs, $n = 1000$, $m = 5$, $\alpha = 0.25$, Borda scoring).

Utility and Societal Weights. We now examine how individual utility—and its intrinsic and empathetic components—and computed weights depend on their degree of nodes in the social networks in global empathetic model. Using data from the previous experiment, we show average utility and average weight in Fig. 2. From Fig. 2(a) we see that as node degree increases (each node has identical in/out-degree), overall utility tends to increase; moreover most of this increase is due to in large part to an increase in intrinsic utility. Fig. 2(b) also illustrates a strong correlation between degree and agent weight. Nodes with higher degree are more powerful and "influential" in the choice of the consensus alternative. This correlation might be an artifact of the specific social networks we generate. However, the relationship between Figs. 2(a) and (b)—which is independent of the specifics of our experiments—shows that individuals with higher weight tend to prefer the consensus winner more than individuals with lower weight.

The effect of m, k, and scoring rule. We now explore the impact on RSWL of changing the numbers of agents m, the number of initial nodes k when generated the network, and difference between Borda and plurality scoring. We set $\alpha = 0.25$, $n = 1000$, and run 2500 instances for each parameter setting (as above).

Fig. 4 shows average (and maximum, minimum) RSWL for three actual, assumed model combinations for various combinations of rule, m and k, denoted by rule(m, k) (e.g., Plura$(5, 1)$ represents $m = 5$, $k = 1$, and plurality). Comparing Borda$(5, 1)$ and Plura$(5, 1)$, and Borda$(10, 1)$ and Plura$(10, 1)$, we see plurality is more susceptible to social welfare loss than Borda. Increasing m has negligible effect on RSWL when Borda is used, but this is not true of plurality. Surprisingly, increasing k from 1 to 2 decreases RSWL (see Borda$(5, \cdot)$): this occurs because, when $k = 2$, the resulting network is denser since each node has at least two neighbors. This connectivity, causes the number of "very influential" agents to increase; but since weights are normalized (the sum of all weights sums to n), their overall influence decreases as they "share their influence," and weight variance over $\mathcal{N}$ decreases.

Self-loop weight α. When we vary the self-loop weight α, it has a significant effect on RSWL when the actual utility model is global but the intrinsic utility model is assumed. We fix $n = 1000$, $m = 5$, $k = 1$ and vary α over $\{0.05, 0.1, 0.25, 0.5, 0.75\}$ (2500 instances for each setting). Table 4 shows that, for both Borda and plurality, increasing α (i.e., decreasing

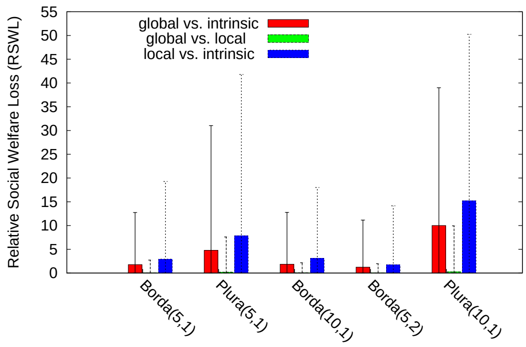


Figure 3: The average (maximum, minimum) RSWL (2500 runs).

	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
Borda	15.84%	14.82%	12.73%	9.79%	5.42%	0%
Plurality	39.12%	36.29%	31.02%	22.46%	13.09%	0%

Table 4: Maximum values of RSWL, global vs. intrinsic models.

overall degree of empathy) decreases RSWL.

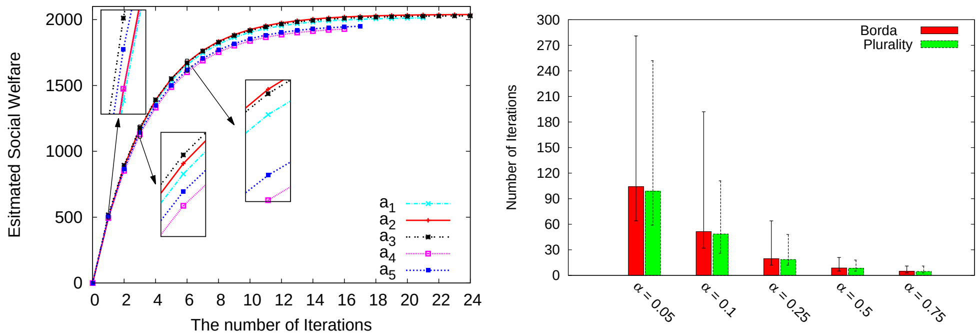
Number of Iterations of ICE. Finally we examine how the self-loop weight α and Borda/plurality utilities affect the expected number of iterations by the iterated candidate elimination algorithm. We fix $n = 1000$ and $m = 5$, and vary α over $\{0.05, 0.1, 0.25, 0.5, 0.75\}$ (2500 instances). Fig. 4(a) illustrates estimated social welfare for each alternative in one representative instance, with $\alpha = 0.25$ and Borda scoring (this instance of ICE converges in under 2 milliseconds). It converges completely in 24 iterations (n.b. $n = 1000$). Alternatives a_4 and a_5 are eliminated at iterations 16 and 17, respectively; a_1 after 20 iterations; and a_2 after 24 iterations; hence a_3 is optimal. We note that the relative ordering of the alternatives is fixed after 6 iterations (in this instance), which might suggest new methods for early termination.

Fig. 4(b) shows the average (and max, min) number of iterations of ICE for various α, for both Borda and plurality. In all cases, the number of iterations is small compared to the size of the network. ICE is relatively insensitive to the scoring rule, and convergence time decreases dramatically with increasing α, as is typical for iterative algorithms (e.g., for Markov chains). (i.e., for a specific α, the average required iterations is almost the same for Borda and plurality).

5 Concluding Remarks and Future Work

We have presented a new model for social choice situations in which an individual's intrinsic preference for alternatives is combined with their *empathetic* preferences, reflecting their desire to see others satisfied with the selected alternative. Treating a social network as one possible measure of strength of empathetic preference, we developed models and algorithms, for both local and global empathetic settings, that allow one to compute social welfare maximizing outcomes efficiently by weighting the contribution of each agent. Our models have a natural interpretation as empathetic voting models when scoring rules are used. Critically, we require only that individuals specify their intrinsic preferences (and network weights): they need not reason about their neighbor's preferences.

This model, while novel, is merely a starting point for a broader investigation into the role of empathetic preferences in social choice. We are currently exploring more realistic processes for simultaneous generation of networks and preferences that are even better suited



(a) Estimated social welfare over iterations (b) Number of iterations for different α

Figure 4: (a) Estimated social welfare over iterations of ICE for 1 run and (b) average (with maximum and minimum) number of iterations of ICE.

to empathetic voting than preferential attachment networks. While our focus has been the choice of a single alternative/winner, our model can also be applied to matching, assignment, and other group decision problems; each will require its own analysis and algorithmic developments. More importantly is the question of the prevalence and strength of empathetic preferences, the extent to which social network structure is indicative of such preferences, and how one can best discover these preferences in practical settings without an excessive burden on users. Two other important directions are: voting schemes in which agents can specify their tradeoffs between intrinsic and empathetic preference in a more qualitative fashion; and considering the possibility that agents are not truthful and fully aware of their utility functions. These questions require both social scientific and computational insight.

References

[1] S. L. Althaus. Information effects in collective preferences. *American Political Science Review*, 92(3):545–558, 1998.

[2] A. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286:509–512, 1999.

[3] E. Bodine-Baron, C. Lee, A. Chong, B. Hassibi, and A. Wierman. Peer effects and stability in matching markets. *SAGT-10*, pp.117–129, 2011.

[4] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna. Voting in Social Networks. *CIKM'09*, pp.777–786, 2009.

[5] S. Brânzei and K. Larson. Coalitional affinity games and the stability gap. *IJCAI'09*, pp.79–84, 2009.

[6] S. Brânzei and K. Larson. Social distance games. *IJCAI'11*, pp.91–96, 2011.

[7] N. Christakis and J. Fowler. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives.* Back Bay Books, 2011.

[8] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990.

[9] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms.* 1990.

[10] D. A. Easley and J. M. Kleinberg. *Networks, Crowds, and Markets - Reasoning About a Highly Connected World.* Cambridge, 2010.

[11] J. H. Fowler and N. A. Christakis. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, 337:a2338, 2008.

[12] G. H. Golub and C. F. Van Loan. *Matrix Computations*, Vol. 3. JHU Press, 1996.

[13] N. Haghpanah, N. Immorlica, V. Mirrokni, K. Munagala. Optimal auctions with positive network externalities. *EC'11*, pp.11–20, 2011.

[14] M. O. Jackson. *Social and Economic Networks.* Princeton, Princeton, NJ, 2008.

[15] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*, Vol. 2. SIAM, 2000.

[16] D. C. Mutz. Cross-cutting social networks: Testing democratic theory in practice. *American Political Science Review*, 96(1):111–126, 2002.

[17] Y. Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2003.

[18] A. E. Sokhey and S. D. McClurg. Social networks and correct voting. Preprint, 2008.

[19] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 14:354–356, 1969.

[20] A. Wilson, M. Baccara, A. Imrohoroglu, and L. Yariv. A field study on matching with network externalities. *Behavioral and Quantitative Game Theory (BQGT'10)*, 2010.

A Linear Algebra Background

We here first explain some definitions and theorems used in our proofs.

Definition Spectrum For $n \times n$ matrix $\mathbf{A}$, the set of its distinct eigenvalues, denoted by $\sigma(\mathbf{A})$, is called the spectrum of $\mathbf{A}$.

Definition Spectral Radius $\rho(\mathbf{A})$. Let $\mathbf{A}$ be an $n \times n$ matrix with real or complex eigenvalues which belong to $\sigma(\mathbf{A})$. Then the spectral radius of $\mathbf{A}$ is $\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$.

Definition M-matrix . A matrix $\mathbf{A}$ in the form of $\mathbf{A} = s\mathbf{I} - \mathbf{B}$ is M-matrix if $s \geq \rho(\mathbf{B})$ and $\mathbf{B} \geq \mathbf{0}$.

Proposition A.1 (Nonsingular M-matrix [15]) *If $s > \rho(\mathbf{B})$ in M-matrix $\mathbf{A} = s\mathbf{I} - \mathbf{B}$, then $\mathbf{A}$ is nonsingular and $\mathbf{A}^{-1} \geq \mathbf{0}$*

Note that M-matrix as defined in Definition A can be either singular or non-singular. Therefore, the condition $s > \rho(\mathbf{B})$ as stated in Proposition A.1 is necessary to guarantee the nonsingularity of M-matrix. However in some references, for example in [15], M-matrices defined originally with $s > \rho(\mathbf{B})$ and as a result nonsingularity is the property of all M-matrices due to that definition.

Theorem A.2 (Gerschgorin Circles [15]) *The eigenvalues of matrix $\mathbf{A} \in \mathcal{C}^{n \times n}$ are contained in $\cup_{i=1}^{n} \mathcal{G}_i$ where $\mathcal{G}_i$ is the Gerschgorin circle defined by:*

$$\mathcal{G}_i = \left\{ c \in \mathcal{C} \middle| |c - a_{ii}| \leq R_i \right\} \textit{ where } R_i = \sum_{\substack{0 \leq j \leq n \\ j \neq i}} |a_{ij}|$$

We here explain induced matrix norms which have been used in our analysis for convergence rate of our iterative method for fixed-point utilities. For a given vector norm $\|.\|$, the induced norm for $n \times m$ matrix $\mathbf{A} \in \mathcal{C}^{n \times m}$ is defined by:

$$\begin{aligned} \|\mathbf{A}\| &= \max\left\{\|\mathbf{A}\mathbf{x}\| : \mathbf{x} \in \mathcal{C}^m \text{ and } \|\mathbf{x}\| = 1\right\} \\ &= \max\left\{\frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} : \mathbf{x} \in \mathcal{C}^m \text{ and } \mathbf{x} \neq \mathbf{0}\right\} \end{aligned}$$

We here mostly focus on the p-norm matrix norm $\|.\|_p$ which are induced by p-norm in vector spaces. More precisely, the p-norm of matrix $\mathbf{A}$ is

$$\|\mathbf{A}\|_p = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$$

where p-norm of vector $\mathbf{x} \in \mathcal{C}^n$ is denoted by $\|\mathbf{x}\|_p$ and defined by:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

There are some important properties for p-norm matrix norms: (1) they are *consistent* or *submultiplicative*: $\|\mathbf{AB}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p$. A consequence of this consistency property is that, for any square matrix $\mathbf{A}$, $\|\mathbf{A}^k\|_p \leq \|\mathbf{A}\|_p^k$. (2) By definition, they are *compatible*: $\|\mathbf{A}\mathbf{x}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{x}\|_p$ where $\mathbf{A} \in \mathcal{C}^{n \times m}$ and $\mathbf{x} \in \mathcal{C}^m$.

For two cases of $p = 1$ or $p = \infty$, the matrix p-norm can be simply calculated. When $p = 1$, the 1-norm for matrix $\mathbf{A}$ is simply the maximum absolute column sum of $\mathbf{A}$, which is formally calculated by:

$$\|\mathbf{A}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}| \tag{6}$$

The $\infty - norm$ for matrix $\mathbf{A}$, denoted by $\|\mathbf{A}\|_\infty$, is simply the maximum absolute row sum of $\mathbf{A}$, which is formally calculated by

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}| \tag{7}$$

We review the Jacobi iterative method and its convergence criteria and rate. Iterative methods have shown practical advantages for solving linear systems [17]. The linear system problem can be formally defined as follows: Given an $n \times n$ real-valued matrix $\mathbf{A}$ and a real n-vector $\mathbf{b}$, the problem is to find n-vector $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} = \mathbf{b}$.

Jacobi method [17] is an iterative method for solving linear systems. Consider this decomposition $\mathbf{A} = \mathbf{\Lambda} - \mathbf{E} - \mathbf{F}$ where $\mathbf{\Lambda}$ is a diagonal matrix of $\mathbf{A}$, $\mathbf{E}$ is strictly lower triangular matrix of $-\mathbf{A}$, and $\mathbf{F}$ is strictly upper triangular matrix of $-\mathbf{A}$. Note that we assume that the diagonal entries of $\mathbf{A}$ are all non-zero. We can write the Jacobi iteration in the vector form as follows:

$$\mathbf{x}^{(t+1)} = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})\mathbf{x}^{(t)} + \mathbf{\Lambda}^{-1}\mathbf{b} \tag{8}$$

Theorem A.3 (General Convergence Theorem for Iterative Methods [17]) *Given an iterative method in the form of* $\mathbf{x}_{t+1} = \mathbf{G}\mathbf{x}_t + \mathbf{f}$ *where* $\mathbf{G}$ *is an* $n \times n$ *iteration matrix and* $\mathbf{f}$ *is an n-vector. The iterative method converges if and only if* $\rho(\mathbf{G}) < 1$.

Corollary A.1 (Jacobi Method Convergence) *Given the Jacobi iterative method as presented in Equation 8, it converges to the solution of linear system* $\mathbf{Ax} = \mathbf{b}$ *if* $\rho(\mathbf{G}) < 1$ *where* $\mathbf{G} = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})$.

Proof The proof of convergence is trivial and immediately follows form the Theorem A.3 by letting $\mathbf{G} = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})$ and $\mathbf{f} = \mathbf{\Lambda b}$. Now, we prove that Jacobi method converges to the solution of linear system. Since the Jacobi method converges, let $\mathbf{x}^* = \lim_{t \to \infty} \mathbf{x}^{(t)}$. From Jacobi iterative formula presented in Equation 8, we can write:

$$\begin{aligned}
&\lim_{t\to\infty} \mathbf{x}^{(t+1)} = \lim_{t\to\infty} \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})\mathbf{x}^{(t)} + \mathbf{\Lambda}^{-1}\mathbf{b} \\
\implies &\lim_{t\to\infty} \mathbf{x}^{(t+1)} = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F}) \left(\lim_{t\to\infty} \mathbf{x}^{(t)}\right) + \mathbf{\Lambda}^{-1}\mathbf{b} \\
\implies &\mathbf{x}^* = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})\mathbf{x}^* + \mathbf{\Lambda}^{-1}\mathbf{b} \\
\implies &\mathbf{\Lambda}\mathbf{x}^* = (\mathbf{E} + \mathbf{F})\mathbf{x}^* + \mathbf{b} \\
\implies &(\mathbf{\Lambda} - \mathbf{E} - \mathbf{F})\mathbf{x}^* = \mathbf{b} \\
\implies &\mathbf{A}\mathbf{x}^* = \mathbf{b}
\end{aligned}$$

So $\mathbf{x}^*$ is the solution for our linear system. ∎

B Proofs

To prove Theorem 2.1, we first show

Lemma B.1 *Assuming nonnegativity, normalizaton, and positive self-loop,* $\rho(\mathbf{B}) < 1$ *where* $\mathbf{B} = \mathbf{W} - \mathbf{D}$.

Proof By the definition of $\mathbf{W}$ and $\mathbf{D}$, it can be seen that $\mathbf{B} = \mathbf{W} - \mathbf{D}$ is a matrix with $b_{ii} = 0$ and $b_{ij} = w_{ij}$ for all $i, j \in \mathcal{N}$ and $i \neq j$. Using Gerschgorin Circle Theorem, stated in Theorem A.2, we have $\sigma(\mathbf{B}) \subset \cup_{i=1}^{n} \mathcal{G}_i$ where

$$\mathcal{G}_i = \big\{c \in \mathcal{C} \big| |c - b_{ii}| \leq R_i\big\} \text{ and } R_i = \sum_{\substack{0 \leq j \leq n \\ j \neq i}} |b_{ij}|$$

As $b_{ii} = 0$ and $b_{ij} = w_{ij}$ for $i \neq j$, we have:

$$\mathcal{G}_i = \big\{c \in \mathcal{C} \big| |c| \leq R_i\big\} \text{ where } R_i = \sum_{\substack{0 \leq j \leq n \\ j \neq i}} |w_{ij}|$$

Note that each $\mathcal{G}_i$ is a closed disk in $\mathcal{C}$ which is centered at 0. So $\cup_{i=1}^{n} \mathcal{G}_i$ is the union of closed disks with various radii but the same center of 0. Since the number of these disks is finite, we can cover all these closed disks with a closed covering disk defined by $\big\{c \in \mathcal{C} \big| |c| \leq R_{max}\big\}$ where $R_{max} = \max_{i=1}^{n} R_i$. Without loss of generality, let $l = \arg\max_i R_i$. So, we have

$$\sigma(\mathbf{B}) \subset \cup_{i=1}^{n} \big\{c \in \mathcal{C} \big| |c| \leq R_i\big\} \subseteq \big\{c \in \mathcal{C} \big| |c| \leq R_l\big\}$$

From this, it follows that:

$$|\lambda| \leq R_l, \forall \lambda \in \sigma(\mathbf{B}) \implies \max_{\lambda \in \sigma(\mathbf{B})} |\lambda| \leq R_l \implies \rho(\mathbf{B}) \leq R_l$$

Using $R_l = \sum_{j \neq l} |w_{lj}|$ and normalization assumption of $\sum_j w_{lj} = 1$, we have $\rho(\mathbf{B}) \leq 1 - w_{ll}$. Since $w_{ll} > 0$ based on self-loop positivity, we have $\rho(\mathbf{B}) \leq 1 - w_{ll} < 1$. ∎

Proof of Theorem 2.1 Based on Eq. 2, we can write:

$$\begin{aligned} \mathbf{u}(a) = (\mathbf{W} - \mathbf{D})\mathbf{u}(a) + \mathbf{D}\mathbf{u}^I(a) &\implies \mathbf{u}(a) - (\mathbf{W} - \mathbf{D})\mathbf{u}(a) = \mathbf{D}\mathbf{u}^I(a) \\ &\implies (\mathbf{I} - (\mathbf{W} - \mathbf{D}))\mathbf{u}(a) = \mathbf{D}\mathbf{u}^I(a) \end{aligned}$$

So it is only sufficient to show that $(\mathbf{I} - (\mathbf{W} - \mathbf{D}))^{-1}$ always exists to prove that $\mathbf{u}(a) = (\mathbf{I} - \mathbf{W} + \mathbf{D})^{-1}\mathbf{D}\mathbf{u}(a)$ always exists and is unique. Since $(\mathbf{I} - (\mathbf{W} - \mathbf{D}))$ is $n \times n$ matrix, we only need to show that it is nonsingular to guarantee the existence of $(\mathbf{I} - (\mathbf{W} - \mathbf{D}))^{-1}$.

Let $\mathbf{B} = \mathbf{W} - \mathbf{D}$. By the definition of $\mathbf{W}$ and $\mathbf{D}$, the matrix $\mathbf{B}$ has $b_{ii} = 0$ and $b_{ij} = w_{ij}$ for all $i, j \in \mathcal{N}$ and $i \neq j$. Based on nonnegativity assumption, we have $w_{ij} \geq 0$. So we have $\mathbf{B} \geq \mathbf{0}$. By setting $s = 1$, $(\mathbf{I} - (\mathbf{W} - \mathbf{D})) = (s\mathbf{I} - \mathbf{B})$ which is an M-matrix (See Definition A). Using Lemma B.1, we have $\rho(\mathbf{B}) < 1$. Since $s = 1$, then $\rho(\mathbf{B}) < s$. From Proposition A.1, it follows that $(\mathbf{I} - (\mathbf{W} - \mathbf{D}))$ is nonsingular and $(\mathbf{I} - (\mathbf{W} - \mathbf{D}))^{-1} \geq \mathbf{0}$. ∎

Proof of Theorem 3.1 From Eq. 2, we observe that $\mathbf{u}(a)$ is the solution to the linear system equation of $\mathbf{A}\mathbf{u}(a) = \mathbf{b}$ with $\mathbf{A} = \mathbf{I} - (\mathbf{W} - \mathbf{D})$ and $\mathbf{b} = \mathbf{D}\mathbf{u}^I(a)$. The Jacobi method (as presented in Eq. 8) for solving this linear system is

$$\mathbf{u}(a)^{(t+1)} = \mathbf{\Lambda}^{-1}(\mathbf{E} + \mathbf{F})\mathbf{u}(a)^{(t)} + \mathbf{\Lambda}^{-1}\mathbf{b}$$

Since $\mathbf{A} = \mathbf{I} - (\mathbf{W} - \mathbf{D})$, we have that $\mathbf{\Lambda} = \mathbf{I}$ and $\mathbf{E} + \mathbf{F} = \mathbf{W} - \mathbf{D}$ based on the definitions. As $\mathbf{b} = \mathbf{D}\mathbf{u}^I(a)$, we have:

$$\begin{aligned} &\mathbf{u}(a)^{(t+1)} = \mathbf{I}^{-1}(\mathbf{W} - \mathbf{D})\mathbf{u}(a)^{(t)} + \mathbf{I}^{-1}\mathbf{D}\mathbf{u}^I(a) \\ \implies &\mathbf{u}(a)^{(t+1)} = (\mathbf{W} - \mathbf{D})\mathbf{u}(a)^{(t)} + \mathbf{D}\mathbf{u}^I(a) \end{aligned}$$

From Lemma B.1, we have $\rho(\mathbf{W}-\mathbf{D})<1$. Then, using Corollary A.1, we have shown that $\mathbf{u}^{(t+1)}(a)=(\mathbf{W}-\mathbf{D})\mathbf{u}^{(t)}(a)+\mathbf{D}\mathbf{u}^{I}(a)$ converges to $\mathbf{u}(a)$ which is the solution to the linear system equation of $\mathbf{A}\mathbf{u}(a)=\mathbf{b}$ with $\mathbf{A}=\mathbf{I}-(\mathbf{W}-\mathbf{D})$ and $\mathbf{b}=\mathbf{D}\mathbf{u}^{I}(a)$. ∎

Proof of Theorem 3.2 Using Eq. 2 and $\mathbf{u}^{(t)}(a)=(\mathbf{W}-\mathbf{D})\mathbf{u}^{(t-1)}(a)+\mathbf{D}\mathbf{u}^{I}(a)$, we can write $\mathbf{u}(a)-\mathbf{u}^{(t)}(a)=(\mathbf{W}-\mathbf{D})(\mathbf{u}(a)-\mathbf{u}^{(t-1)}(a))$. By induction on t, we have $\mathbf{u}(a)-\mathbf{u}^{(t)}(a)=(\mathbf{W}-\mathbf{D})^{t}\left(\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right)$. Thus, we have

$$
\begin{aligned}
\left\|\mathbf{u}(a)-\mathbf{u}^{(t)}(a)\right\|_{\infty} &= \left\|(\mathbf{W}-\mathbf{D})^{t}\left(\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right)\right\|_{\infty} \\
&\leq \left\|(\mathbf{W}-\mathbf{D})^{t}\right\|_{\infty}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using compatibility property of p-norm} \\
&\leq \|\mathbf{W}-\mathbf{D}\|_{\infty}^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using consistency property of p-norm} \\
&= \left(\max_{1\leq i\leq n}\sum_{j=1}^{n}|w_{ij}-d_{ij}|\right)^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using } \infty\text{-norm formula in Equation 7} \\
&= \left(\max_{1\leq i\leq n}\sum_{\substack{j=1\\ j\neq i}}^{n}|w_{ij}|\right)^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using the definition of matrix } \mathbf{D} \\
&= \left(\max_{1\leq i\leq n}\sum_{\substack{j=1\\ j\neq i}}^{n}w_{ij}\right)^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using the nonnegativity assumption} \\
&= \left(1-\min_{1\leq i\leq n}w_{ii}\right)^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty} && \text{using the normalization assumption}
\end{aligned}
$$

Let $\tilde{w}=\min_{1\leq i\leq n}w_{ii}$. So $\left\|\mathbf{u}(a)-\mathbf{u}^{(t)}(a)\right\|_{\infty}\leq(1-\tilde{w})^{t}\left\|\mathbf{u}(a)-\mathbf{u}^{(0)}(a)\right\|_{\infty}$. ∎

Lemma B.2 *Assuming nonnegativity and normalization, and the iterative updating scheme of* $\mathbf{u}^{(t)}(a)=(\mathbf{W}-\mathbf{D})\mathbf{u}^{(t-1)}(a)+\mathbf{D}\mathbf{u}^{I}(a)$*, If* $\forall i\in\mathcal{N}$*,* $u_i^{I}(a)\in[c,d]$ *and* $u_i^{(0)}(a)\in[c,d]$*, then* $u_i(a)^{(t)}\in[c,d]$*,* $\forall i\in\mathcal{N}$ *and* $\forall t\in\mathbb{N}$*. Moreover, we have* $u_i(a)\in[c,d]$*,* $\forall i\in\mathcal{N}$*.*

Proof We first prove the first part of the lemma by induction on t. The base case is $t=0$ for which it is given that $u_i^{(0)}(a)\in[c,d]$, $\forall i\in\mathcal{N}$. Now, we prove the inductive step. The induction hypothesis is that $u_i^{(t)}(a)\in[c,d]$ for all $\forall i\in\mathcal{N}$. There are two useful inequalities which follows immediately from induction hypothesis: $\max_{i\in\mathcal{N}}u_i^{(t)}(a)\leq d$ and $\min_{i\in\mathcal{N}}u_i^{(t)}(a)\geq c$. We can write the updating scheme for each individual $i\in\mathcal{N}$ and the alternative $a\in\mathcal{A}$ as follows:

$$u_i^{(t+1)}(a)=w_{ii}u_i^{I}(a)+\sum_{k\neq i}w_{ik}u_k^{(t)}(a) \tag{9}$$

where $u_i^{(t)}(a)$ denote the utility of individual i for alternative a after t iterations. Using this equation and two inequalities mentioned above, we will first show the upper bound of d and then the lower found of c for $u_i^{(t+1)}(a)$, $\forall i\in\mathcal{N}$, and fixed a. For the upper bound, we can write the following:

$$
\begin{aligned}
u_i^{(t+1)}(a) \leq & \max_{i\in\mathcal{N}} \left\{u_i^{(t+1)}(a)\right\} \\
= & \max_{i\in\mathcal{N}} \left\{ w_{ii}u_i^I(a) + \sum_{k\neq i} w_{ik}u_k^{(t)}(a) \right\} && \text{using Equation 9} \\
\leq & \max_{i\in\mathcal{N}} \left\{w_{ii}u_i^I(a)\right\} + \sum_{k\neq i} \max_{i\in\mathcal{N}} \left\{w_{ik}u_k^{(t)}(a)\right\} && \text{triangle inequality} \\
\leq & w_{ii} \max_{i\in\mathcal{N}} \left\{u_i^I(a)\right\} + \sum_{k\neq i} w_{ik} \max_{i\in\mathcal{N}} \left\{u_k^{(t)}(a)\right\} \\
\leq & w_{ii}d + \sum_{k\neq i} w_{ik} \max_{i\in\mathcal{N}} \left\{u_k^{(t)}(a)\right\} && \text{based on the given assumption} \\
\leq & w_{ii}d + \sum_{k\neq i} w_{ik}d && \text{using induction hypothesis} \\
= & d\sum_{k} w_{ik} = d && \text{using normalization assumption}
\end{aligned}
$$

Similarly, we can show the lower bound of c for $u_i^{(t+1)}(a)$:

$$
\begin{aligned}
u_i^{(t+1)}(a) \geq & \min_{i\in\mathcal{N}} \left\{u_i^{(t+1)}(a)\right\} \\
= & \min_{i\in\mathcal{N}} \left\{ w_{ii}u_i^I(a) + \sum_{k\neq i} w_{ik}u_k^{(t)}(a) \right\} && \text{using Equation 9} \\
\geq & \min_{i\in\mathcal{N}} \left\{w_{ii}u_i^I(a)\right\} + \sum_{k\neq i} \min_{i\in\mathcal{N}} \left\{w_{ik}u_k^{(t)}(a)\right\} && \text{triangle inequality} \\
\geq & w_{ii} \min_{i\in\mathcal{N}} \left\{u_i^I(a)\right\} + \sum_{k\neq i} w_{ik} \min_{i\in\mathcal{N}} \left\{u_k^{(t)}(a)\right\} \\
\geq & w_{ii}c + \sum_{k\neq i} w_{ik} \left\{u_k^{(t)}(a)\right\} && \text{based on the given assumption} \\
\geq & w_{ii}c + \sum_{k\neq i} w_{ik}c && \text{using induction hypothesis} \\
= & c\sum_{k} w_{ik} = c && \text{using normalization assumption}
\end{aligned}
$$

So we have shown that $c \leq u_i^{(t+1)}(a) \leq d$, $\forall i \in \mathcal{N}$ and given $a \in \mathcal{A}$. This implies $u_i^{(t+1)}(a) \in [c,d]$, $\forall i \in \mathcal{N}$ and given $a \in \mathcal{A}$. So the induction follows and we have proved the first part of the lemma.

Now, we will prove the second part of lemma by showing that $u_i(a) \in [c,d]$, $\forall i \in \mathcal{N}$ and given $a \in \mathcal{A}$. Fix an arbitrary $i \in \mathcal{N}$. The sequence $u_i^{(t)}(a)$ with $t = 0,1,2,\dots$ is a convergent sequence which converges to $u_i(a) = lim_{t\to\infty} u_i^{(t)}(a)$ (based on Theorem 3.1). Note that, from the first part of this lemma, we have $u_i^{(t)}(a) \in [c,d]$ for any $t \in \mathbb{N} \cup \{0\}$. So we can see $u_i^{(t)}(a)$ is a convergent sequence on the closed set $[c,d]$. As $[c,d]$ is closed, the limit point of $u_i^{(t)}(a)$ sequence which is $u_i(a)$ should belong to $[c,d]$. As i and a are chosen arbitrarily, we have $u_i(a) \in [c,d]$, $\forall i \in \mathcal{N}$ and given $a \in \mathcal{A}$. ∎

Proof of Theorem 3.3 Let $\tilde{w} = \min_{1 \leq i \leq n} w_{ii}$. Using Theorem 3.2 and the definition of $\infty-$norm for vector space, we can write

$$\begin{aligned}
&\left\|\mathbf{u}(a) - \mathbf{u}^{(t)}(a)\right\|_\infty \leq (1-\tilde{w})^t \left\|\mathbf{u}(a) - \mathbf{u}^{(0)}(a)\right\|_\infty \implies \\
&\left|u_i(a) - u_i^{(t)}(a)\right| \leq \max_i \left|u_i(a) - u_i^{(t)}(a)\right| \leq (1-\tilde{w})^t \left\|\mathbf{u}(a) - \mathbf{u}^{(0)}(a)\right\|_\infty \implies \\
&\sum_{i=1}^{n} \left|u_i(a) - u_i^{(t)}(a)\right| \leq n\,(1-\tilde{w})^t \left\|\mathbf{u}(a) - \mathbf{u}^{(0)}(a)\right\|_\infty
\end{aligned} \tag{10}$$

From lemma B.2, we know that $u_i(a) \in [c, d]$. Based on this and the assumption that $u_i^{(0)}(a) \in [c, d]$, it follows that $|u_i(a) - u_i^{(0)}(a)| \leq d - c$. So we can continue inequality 10 as follows:

$$\sum_{i=1}^{n} \left|u_i(a) - u^{(t)}(a)\right| \leq n\,(1-\tilde{w})^t \left\|\mathbf{u}(a) - \mathbf{u}^{(0)}(a)\right\|_\infty \leq n(d-c)\,(1-\tilde{w})^t \tag{11}$$

From lemma B.2, we know that $u_i(a) \in [c, d]$ and $u_i^{(t)}(a) \in [c, d]$, $\forall t \in \mathbb{N} \cup \{0\}$. Using triangle inequality, we can write:

$$\sum_{i=1}^{n} \left|u_i(a) - u_i^{(t)}(a)\right| \geq \left|\sum_{i=1}^{n} \left(u_i(a) - u_i^{(t)}(a)\right)\right| = \left|\sum_{i=1}^{n} u_i(a) - \sum_{i=1}^{n} u_i^{(t)}(a)\right| = \left|sw(a) - sw^{(t)}(a)\right| \tag{12}$$

From inequalities 11 and 12, we can conclude that

$$\left|sw(a) - sw^{(t)}(a)\right| \leq n(d-c)\,(1-\tilde{w})^t$$

where $\tilde{w} = \min_{1 \leq i \leq n} w_{ii}$. ∎

Proof of Proposition 3.4 Using the triangle inequality and the inequality presented in Theorem 3.3, we can write:

$$\begin{aligned}
sw^{(t)}(b) - sw^{(t)}(a) &= sw^{(t)}(b) - sw(b) + sw(b) - sw^{(t)}(a) + sw(a) - sw(a) \\
&\leq |sw^{(t)}(b) - sw(b)| + sw(b) + |sw(a) - sw^{(t)}(a)| - sw(a) \\
&\leq n(d-c)(1-\tilde{w})^t + sw(b) + n(d-c)(1-\tilde{w})^t - sw(a) \\
&= 2n(d-c)(1-\tilde{w})^t + sw(b) - sw(a)
\end{aligned}$$

Using this and $sw^{(t)}(b) - sw^{(t)}(a) \geq 2n(d-c)(1-\tilde{w})^t$, we have $2n(d-c)(1-\tilde{w})^t \leq 2n(d-c)(1-\tilde{w})^t + sw(b) - sw(a)$. This implies $sw(b) \geq sw(a)$. ∎

C The ICE Algorithm

Algorithm 1: Iterated Candidate Elimination (ICE)

input : Social graph G, intrinsic utilities $u_i^I(a) \in [c, d]$, $\forall i \in \mathcal{N}$ and $\forall a \in \mathcal{A}$.
output: Consensus winner a^*.

Initialize $u_i^{(0)}(a) \leftarrow c$, $\forall i \in \mathcal{N}$ and $\forall a \in \mathcal{A}$;

```
// C is the possible winner candidate set
```

$C \leftarrow \mathcal{A}$;
$\tilde{w} = \min_{1 \le i \le n} w_{ii}$;
$t \leftarrow 0$;
while ***size***$(C) > 1$ **do**
- $t \leftarrow t + 1$;
- **foreach** $a \in C$ **do**
  - $sw^{(t)}(a) \leftarrow 0$;
  - **foreach** $j \in \mathcal{N}$ **do**
    - $u_j^{(t)}(a) \leftarrow w_{jj} u_j^I(a) + \sum_{k: e_{jk} \in E, j \neq k} w_{jk} u_k^{(t-1)}(a)$;
    - $sw^{(t)}(a) \leftarrow sw^{(t)}(a) + u_j^{(t)}(a)$;
- $\hat{sw}^{(t)} \leftarrow \max_{a \in C} sw^{(t)}(a)$;
- **foreach** $a \in C$ **do**
  - **if** $\hat{sw}^{(t)} - sw^{(t)}(a) \ge 2n(d-c)\left(1 - \tilde{w}\right)^t$ **then**
    - $C \leftarrow C - \{a\}$

return $a^* \in C$