Boundedly Rational Voters in Large(r) Networks

Alan Tsang*
National University of Singapore
Singapore
akhtsang@nus.edu.sg

Amirali Salehi-Abari
University of Ontario Institute of Technology
Oshawa, ON, Canada
abari@uoit.ca

Kate Larson
University of Waterloo
Waterloo, ON, Canada
klarson@uwaterloo.ca

ABSTRACT

In Iterative Voting, voters first cast their ballots but may change their minds upon observing the ballots of others. Previous models have extended Iterative Voting to the incomplete information domain of social networks, where voters only observe the ballots of their friends. However, these models are based on computationally-intensive calculations of expected utilities. We propose a framework of bounded rationality for voters situated in social networks. Using this framework, we propose and test a number of heuristics that reduce the computation required for optimal strategic reasoning by several orders of magnitude compared to previous work, while retaining similar qualitative behaviors. These heuristics enable us to conduct simulations on how the size of the voting population affects strategic behavior. To illustrate the effectiveness of our approach, we apply our heuristics to explore the Micromega rule — an observation in political science that large political parties favor small assemblies. We find that the size of electoral districts is a contributing factor to the Micromega rule in some networks. Fringe candidates retain more support in smaller districts, while larger parties dominate in larger districts.

KEYWORDS

Social Choice Theory; Social Networks; Iterative Voting; Simulations

ACM Reference Format:
Alan Tsang, Amirali Salehi-Abari, and Kate Larson. 2018. Boundedly Rational Voters in Large(r) Networks. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 8 pages.

1 INTRODUCTION

Social networks are mathematical descriptions of how individuals interact in a community. They capture how news and information flow between people, and what social structures are present in different communities. As the Internet matures as a technology, more and more information about these social networks are captured as "big data". At the same time, these online social structures wield ever increasing influence over our lives at all scales – from the minutiae of our day-to-day moods [14], to turnout at congressional elections [6]. It is therefore of paramount importance that we understand the mechanisms by which social networks affect decision making.

*This author was affiliated with the University of Waterloo at the time this paper was submitted.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

One natural area to focus our attention is the problem of group decision making, i.e. social choice. In particular, we examine the problem of voting. Voting is a method by which a community elicits the opinions of its members so it may make a collective decision. Savvy voters may choose to vote strategically; that is, when a voter's favorite candidate is unlikely to win, she may choose to vote for a more promising alternative instead. One way of modeling this behavior is *Iterative Voting*, where voters may choose to change their ballots after observing the interim outcome of the election [17]. This process terminates when no voter wishes to change their ballot anymore. In the basic model, voters only change their ballot when doing so will alter the outcome favorably. Subsequent work incorporates voters who are truth biased (who prefer voting sincerely if they cannot otherwise affect the outcome), lazy (who prefer abstaining, all else being equal) [21], or optimistic (who assume up to a fixed number of voters may be swayed to their cause) [20].

Recently, there have been growing research in studying social choice problems on social networks (see, for example, [4, 11, 23, 24]) Iterative Voting has also been extended to social networks [25]. In this model, voters no longer have complete information on the ballots in the election. Instead, voters are embedded in a social network and only observe the ballots of their neighbors. Each voter must use this observation to compute the likely outcomes of the election, and respond accordingly. This calculation is computationally intensive; it scales poorly in larger networks and elections with many candidates. In this paper, we propose a number of heuristic models that greatly reduce the computational and cognitive requirements on the voters. We argue that these heuristics represent natural models of boundedly rational human behavior. Simultaneously, our heuristics speed up the computation of strategic response by up to 2 orders of magnitude, allowing us to explore the strategic behavior of voters in larger populations. In particular, we examine the Micromega rule, the tendency for large political parties to favor small assemblies with large electoral districts, and vice versa. We show that population is a contributing factor to the Micromega rule in some networks.

2 FRAMEWORK

To establish a framework for voter behavior for elections with large populations, we consider several desirable criteria for these models. We base our framework on the desiderata presented for voters in general populations [16], and adapt them to our domain: voters in large populations embedded in social networks. The social network naturally restricts the availability of information to voters in an asymmetric way; one voter has different information about the election than another. Moreover, we emphasize that the voters in our framework are boundedly rational, and therefore computationally limited. This both reflects the human nature of real world voters, and also ensures our heuristics can be computed in a timely manner.

Knowledge: A voter does not have perfect information on the actions of the populations. Rather, each voter is only able to observe the actions of a limited number of other voters. In particular, we will restrict these observations to the neighbors in the social network, which we will define in the sequel. Voters must infer the current state of the world based on this limited information.

Rationality: Subject to their observations, preferences, and beliefs, voters act to maximize their expected utility of the electoral outcome. In particular, while the chances of casting a pivotal vote in an election is very small, it is the only event of importance to rational voters. Their observations allow them to compare the likelihood of pivot conditions between different candidates and act accordingly.

Anonymity: Beyond readily available network properties, voters treat observations from their social contacts anonymously. Candidates are judged only based on the utility that they bring to the voters upon being elected.

Equilibrium: The model converges to an equilibrium outcome (according to some established solution concept), or readily shows it cannot exist.

Tractability: The computation of voter responses is computationally tractable for the voter. Real world voters are boundedly rational agents and frequently employ heuristics to simplify their cognitive load. While the computation or approximation of probabilities may be unavoidable, this computation should be fast, particularly for "easy" cases.

Optimistic: Voters act in the belief that their actions may have an impact, even when this is not guaranteed. This is in sharp contrast to the complete information setting of Iterative Voting, where voters act strategically *only* when they know they are pivotal.

Motivated by these desiderata, we make the following assumptions in our heuristic models. We do not consider these desirable criteria; rather, we consider these to be natural ways of implementing them in our voter heuristics.

Markovian Strategy: While voters have access to histories of past actions from their social contacts, we assume voter response is Markovian and computed as a function of current observations. While making use of past history may allow the voter to detect patterns and trends, doing so is computationally intensive and further compounds concerns of tractability. We argue that this simplifying assumption reasonably models human behavior, because humans are bounded rational agents and human memory is limited.

Myopic Response: Being boundedly rational, voters are not concerned about second order effects in the network. That is, they do not consider that an adjustment in their ballot may also cause others to adjust their ballots, and this knock-on effect may be detrimental to the original voter. Instead, we assume voter responses are myopic improvements to the current situation. This is a reasonable assumption to adopt because predicting these knock-on effects will be computationally intensive. However, adopting this assumption may actually make equilibria more difficult to achieve. For instance, non-myopic agents may be able to predict actions that lead to cycling behaviors within the population, which prevent the convergence to an equilibrium outcome.

2.1 Voting Model

Let the population V of n voters be situated in a social network represented by a simple, directed graph $G = (V, E)$. A directed edge $(i, j) \in E$ means voter i observes voter j, and therefore, j's actions may influence i. Let $C = \{1, 2, \ldots m\}$ denote the set of candidates. Each voter i has a preference p_i represented as an integer from the interval $[0, 100]$; each candidate supports a position drawn from the same space. If the winning candidate supports position $\hat{p}$, voter i derives utility $u_i(p_i, \hat{p}) = -|p_i - \hat{p}|^2$ from this outcome. These single peaked preferences are commonly used in social choice theory to represent the political spectrum [3, 12], and quadratic diminishing utility functions, in particular, were used by Myerson and Weber [19].

Each voter casts a ballot from the set of admissible ballots $\mathcal{B}$. A social choice function $\mathcal{F}$ maps the set of submitted ballots to a unique winner from C. We focus on the plurality voting rule, where $\mathcal{B} = C$. If n_i is the number of ballots supporting candidate i, then the plurality voting rule $\mathcal{F} = \arg\max_x n_x$ maps the set of ballots to the candidate receiving the most votes, breaking ties randomly.

Voting proceeds in rounds. In the first round, voters' ballots reflect their sincere top choice. In subsequent rounds, one by one, voters are allowed to revise their ballots based on observing the ballots of their out-neighbors. For a particular voter i, let the vector $\mathbf{s} = (s_1, s_2, \ldots s_m)$ denote the fraction of i's out-neighbors supporting each candidate, with Laplace smoothing applied by adding one to the tally of support for each candidate. Each voter i computes her revised ballot $b' \in \mathcal{B}$ according to her Voter Response Function $R_i(\mathbf{s}, u_i) = b'$. We assume voter behavior is symmetric and so omit the subscript for simplicity. We also omit the parameter u_i when it is clear from context. Voting terminates when no voter's ballot changes in a round; we say that our population has converged to an equilibrium.

2.2 Fully Rational Voter

The fully rational voter computes the exact pivot probability for each pair of candidates by assuming future ballots will be distributed according to a multinomial distribution with support $\mathbf{s}$. The probability of observing a final tally of $\mathbf{b} = (b_1, b_2, \ldots b_m)$ is

$$Pr(\mathbf{b}; n-1; \mathbf{s}) = \frac{(n-1)!}{b_1! b_2! \ldots b_m!} \prod_{i=1}^{m} s_i^{b_i} \tag{1}$$

The voter then computes the probability $T(y, x)$ that the any two given candidates x and y are in a pivot condition. That is, by adding one ballot supporting x, the winner changes from y to x. This is calculated by enumerating all possible such pivot outcomes, and summing the probability of each outcome. For example, under lexicographic tie breaking, and ignoring multi-way ties, the pivot outcomes for candidates 1 and 2, when $n = 10$, $m = 4$, are $(5, 5, 0, 0)$, $(4, 4, 2, 0)$, $(4, 4, 1, 1)$, $(4, 4, 0, 2)$, and $(3, 3, 2, 2)$; each would be associated with a probability of occurrence based on the multinomial distribution. Then, she calculates the Prospect Rating for that candidate:

$$C_x = \sum_{y=1}^{m} T(x,y)(u_x - u_y) \quad (2)$$

The Voter Response Function is $R_{\text{FULL}}(\mathbf{s}) = \arg\max_x C_x$. We will refer to this model as the `Full` Voter model.

The consideration of probable outcomes is crucial in two ways: to allow our voter to act strategically even when she does not know for certain that she will be pivotal, and to allow strategic actions to be sensitive to the size of the voter population. To illustrate this, suppose our voter observes a ballot ratio of (3,5,4) in her neighbors, in order of her preferences. If the election is small, say with only 12 others, the voter could make a convincing argument to stay true to her top choice. But if the election is larger, she might prefer securing her second choice against a more serious threat from her last choice, because the prospect her favorite candidate will make up such a large gap seems much more remote. This model captures the stochasticity of this situation, so that voters act rationally despite the incomplete information available to them.

2.3 Voter Heuristics

As we illustrate above, the `Full` voter model must enumerate all possible pairwise pivot conditions, calculating R_{FULL} is computationally intensive, and scales poorly as n or m increases. To establish a crude upper-bound on the complexity of this computation, we examine a related counting problem, the classic Stars and Bars Problem (SB) by William Feller [10]:

> Given positive integers n and m, the number of distinct m-tuples of *non-negative* integers that sum to n is given by the multiset function $\left(\!\binom{n+1}{m-1}\!\right)$.

The SB number counts the number of possible outcomes of the voting process, which is a gross upper-bound on the number pivotal outcomes. We further refine this estimate by enumerating the number of ballots received by a co-winner, which is bounded between $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{m} + 1 \rceil$, and utilizing the SB number to count the number of ways to distribute the remaining ballots between the remaining candidates. Using this technique, each best response calculation requires $O(m^2 n^{m-2})$ queries to the multinomial distribution.

While this bound is very loose, empirical experimentation reinforces this bound. The `Full` Voter Model scales very poorly as either n or m increases. Since one of our desiderata is Tractability, both for the purpose of scalability, and to more accurately model human bounded rationality, we propose a number of voter heuristics that reduces the computational and cognitive load on the voters. In each of these models, the pivot probabilities are simplified to the chance that x and y are exactly tied as winners (i.e. discounting the cases where x has one fewer ballot than y).

`Top-k` Voter: An intuitive way of easing the voters' cognitive burden is for them to ignore unpromising candidates. This is observed in the political science literature. For example, Meffert and Gschwend [15] conducted studies in the laboratory on strategic voting behavior in coalitional governments. They used fictional parties with monetary incentives based on the elected outcome, and found that participants used a number heuristics when considering their ballot. One of these heuristics was to avoid parties that did not enjoy enough popular support to contribute meaningfully to the result. Similar to Reijngoud and Endriss [22], we allow voters to disregard all but the top k candidates when they consider whom to support. We note that, due to our incomplete information setting, one voter's top k candidates may differ from those of another voter. We model this type of behavior as the `Top-k` voter. Here, voters consider only the $k \leq m$ candidates with the most popular support according to $\mathbf{s}$, breaking ties in favor of utility of victory. The voter treats the election as if only these top k candidates were participating, and computes $R_{\text{Top-k}}$ based on $\binom{k}{2}$ pivot probabilities, rather than $\binom{m}{2}$. The resulting algorithm requires only $O(k^2 n^{k-2})$ queries to the multinomial distribution, though determining the top-k supporters and permuting the entries adds a small overhead to the computation that scales with m.

`Max-M` Voter: The `Full` Voter considers the expected utility gained by supporting a candidate over all other candidates. A boundedly rational voter may employ a different measure, supporting the candidate offering the maximum marginal gain over a rival candidate. That is, the voter focuses on pairwise comparisons between candidates, picking the candidate who offers the most compelling position, and has the best chance of beating the most serious threat. Formally, consider the following utility computation in place of prospect rating C_x (Equation 2):

$$\mathcal{D}_x = \max_{y \neq x} T(y,x)(u_x - u_y)$$

$R_{\text{Max-M}} = \arg\max_x \mathcal{D}_x$ selects the candidate maximizing this marginal utility over some other candidate. The motivation behind using this alternative utility function is that it approximates the behavior of the `Full` Voter, while offering mathematical optimizations that greatly reduce computational load. Such pairwise comparisons may also be more natural for human voters to process. Rather than considering the marginal gains over every other candidate, this calculation emphasizes the candidate's merits against the most salient of opposition. Indeed, political campaigns often focus on demonizing particular opponents to bolster the merits of the favored candidate.

Tie Sampling Heuristics. Rather than calculating the exact probability of a pivot condition between x and y, we may utilize sampling techniques to estimate this probability. Here, the voter may be thought of as sampling from the outcome space to consider specific pivot scenarios, and acting based only on these imagined, plausible outcomes. This gives us the `TieU`, `TieR` and `TieH` models below, each is based on the `Full` voter model.

`TieU`: We first enumerate all pivot outcomes using a finite state machine. But rather than querying the multinomial distribution for each outcome, we sample l outcomes from this space, and calculate the probabilities of each of these outcomes according to Equation 1, and approximate $T(y,x)$ as the sum of these l probabilities. This reduces the number of queries to l, though the algorithm must still iterate over the entire pivot space, which is still an $O(m^2 n^{m-2})$ operation.

`TieH`: Rather than sampling uniformly from the space of pivot outcomes, which requires enumerating that space fully, we use the following heuristic for sampling non-uniformly from this space. The number of ballots b_w for our tied winners is drawn uniformly from the interval $[\lceil \frac{n-2}{m}+1\rceil, \lfloor \frac{n}{2} \rfloor]$.[1] To allocate the remaining ballots r, the algorithm iterates through the other candidates in random order. For each candidate, the space of admissible allocations is the interval $[\max(0, r - m' \times \min(r, b_w - 1)), \min(r, b_w - 1)]$, where m' is the number of unallocated candidates. The algorithm draws uniformly from this interval to allocate the number of ballots for the current candidate, and updates r and m' before moving to the next candidate. The result is one possible pivot outcome. Its probability is calculated according to Equation 1, and $T(y,x)$ is approximated as the sum of l such probabilities. This requires exactly l queries, with negligible overhead.

TieR: We estimate $T(y,x)$ by using a Monte Carlo algorithm in the space of pivot outcomes using rejection sampling. We generate l outcomes of the election by sampling from the multinomial distribution. $T(y,x)$ is estimated as the proportion of those outcomes which result in a 2-way pivot between y and x. This requires exactly l queries, with negligible overhead.

Poisson Voter: Myerson proposed an alternative model for elections, treating them as large Poisson Games [18]. In this interpretation, the number of voters is *uncertain* and follows a Poisson distribution with mean n_e. That is, the probability that there are k voters is

$$\text{Poisson}(k|n_e) = e^{-n_e} n_e^k / k!\,.$$

If a given voter has probability s_b of casting a particular ballot b, then the number of voters casting b is also a Poisson distribution with mean $s_b n_e$. Crucially, this means the number of voters casting one type of ballot is independent of the number of voters casting another ballot. He focuses on the convergence behavior of the probability that a 2-candidate election results in a tie, allowing for abstention where voting may be costly. While Myerson's Poisson Game models a fundamentally different voting process, we will propose a voter heuristic that extends this model to a multi-candidate election to create a voter model with behavior similar to the Full Voter.

Let n_i be the random variable representing the number of votes that candidate i receives, and s_i denote the probability that a given voter casts a ballot supporting candidate i. Recall that n_i follows a Poisson distribution with mean $s_i n_e$. Then, Myerson shows that, as n_e approaches ∞, the probability of casting a pivotal vote[2] in support of $c \in \{1,2\}$, in a 2-candidate election, converges to

$$Pr(n_1 = n_2|\mathbf{s}) \approx \frac{e^{n_e(2\sqrt{s_1 s_2} - s_1 - s_2)}}{4\sqrt{\pi n \sqrt{s_1 s_2}}} \frac{\sqrt{s_1} + \sqrt{s_2}}{\sqrt{s_c}}$$

This requires that $s_1 + s_2 \leq 1$, allowing for some voters to abstain. To extend this model to a multi-candidate election, we treat ballots supporting other candidates as abstentions. We also require that this be a winning tie: i.e. $n_c > n_i, \forall i \neq 1,2$. Since n_i are drawn independently from Poisson distributions with mean $s_i n_e$, the probability that $n_c - n_i > 0$ follows a Skellam distribution, which is approximated as

$$Pr(n_c > n_i) \approx \frac{(1 + (b_c + b_i)^2) e^{-(\sqrt{b_c} - \sqrt{b_i})^2}}{2(b_c + b_i)^2} - \frac{e^{-(b_c + b_i)}}{4\sqrt{b_c b_i}} - \frac{e^{-(b_c + b_i)}}{8 b_c b_i}$$

where $b_c = s_c n_e$ and $b_i = s_i n_e$, and $s_c > s_i$ [13]. If we make the simplifying assumption that the events $n_1 = n_2$ and $n_1 > n_i \forall i \neq 1,2$ are independent[3], then the probability that candidate x and y are in a pivot condition is the intersection of the events where x and y are tied, and x has more votes than every other candidate $i \neq x, y$. So, we approximate $T(y,x)$ as

$$T(y,x) = Pr(n_x = n_y|\mathbf{s}) \prod_{i \in C \setminus \{x,y\}} Pr(n_x > n_i)$$

As a result, the Poisson Model requires $\binom{m}{2}$ probability calculations, each of which takes $O(m)$ computations, giving us a rough runtime of $O(m^3)$.

3 COMPARISON VIA SIMULATIONS

To benchmark these heuristics against the Full Voter model, we construct a framework where a voter v is queried for a strategic response based on a particular set of observations. For each trial, m candidates are generated with positions drawn uniformly at random from $[0, 100]$. The voter v and her $d = 25$ out-neighbors also draw their preferences from the same distribution. The value of d captures the amount of information available to the voter, and also the "resolution" of Figures 5–7. We chose a moderate value of $d = 25$, though we do not expect changing the value of d to significantly impact the results in this section. Each of the out-neighbors are assumed to vote truthfully, and v constructs a strategic ballot based on her Voter Response Function.

Figures 1 through 3 show the time (in seconds) required to construct one voter response under different voter models, with varying number of candidates m and size of electorate n. Each datum is averaged over 1000 trials. The top bar shows the time required for the Full voter. We see that the sampling heuristics, particularly TieR and TieH perform well, up to 2 orders of magnitude faster than Full; their runtimes are plotted in the Insets of Figures 2 and 3 using a different scale. Additionally, the runtime of TieH is unaffected by n. The Poisson Voter is not included in these benchmarks because its runtime is negligible.

We also compare the voter response of the heuristics to the response of the Full Voter. Figure 4 shows the rate of disagreement between the heuristics and the Full Voter. TieH ($L = 2000$) has a disagreement rate of 0.124, which means it computes a strategic ballot different from Full 12.4% of the time. Most heuristics are comparable in their accuracy, except TieR which performs noticeably worse. The accuracy of TieR may be increased by increasing the sample size l, but the gain is modest compared to the increase in runtime. It is interesting to note that even though the Poisson Voter is based on a fundamentally different model of voting, its

[1] The lower bound is obtained by reserving two ballots for the tied winners and then splitting the remaining ballots evenly between m candidates.

[2] This accounts for both a direct tie, and where c is one vote away from a tie.

[3] The two events $n_1 = n_2$ and $n_1 > n_i$ are not independent in general. For example, suppose that $s_1 = s_2 \gg s_3$. If we know that $n_1 < n_3$, then we know that n_1 is likely small, which makes the event $n_1 = n_2$ much less likely.

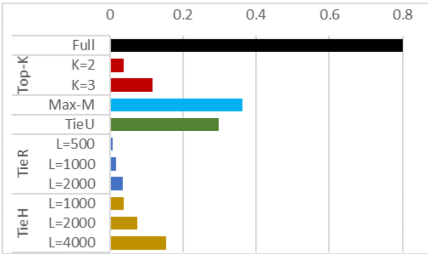

Figure 1: Runtime in seconds to construct one voter response ($m = 5$, $n = 100$).

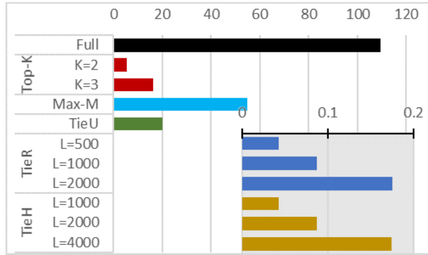

Figure 2: Runtime in seconds to construct one voter response ($m = 5$, $n = 500$). Inset show runtimes of heuristics using a different scale (also in seconds).

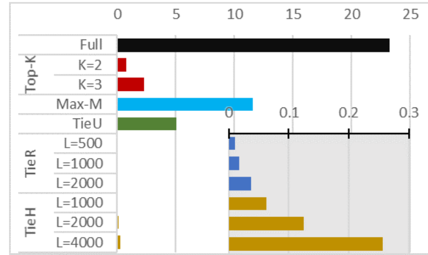

Figure 3: Runtime in seconds to construct one voter response ($m = 6$, $n = 100$). Inset show runtimes of heuristics using a different scale (also in seconds).

accuracy is comparable to the other heuristics according to this benchmark. The results of Figure 4 are representative of other settings of n and m.

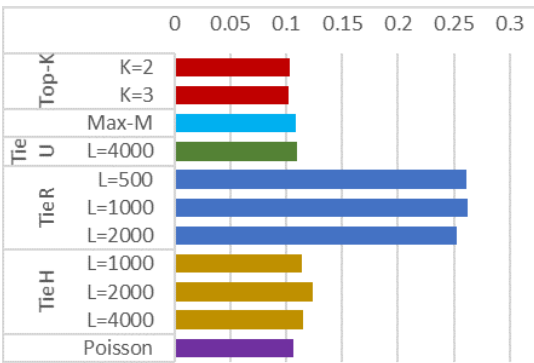

Figure 4: Rate of Disagreements from the `Full` Voter Model ($m = 5$, $n = 500$).

`Poisson` and `TieH` Voters appear to be our best heuristics for approximating the behavior of `Full` Voter. Next, we examine exactly *when* they disagree with `Full` Voter. For this benchmark, we consider the space of all possible observations that the voter may encounter, and determine which observations induce differences in voter response. For simplicity, we consider the case where $m = 4$ and $n = 500$. We assume the voter observes the ballots of $d = 25$ other people in her social network. We fix the voter's preference to be 0, the candidates positions to be $(10, 15, 20, 25)$; that is, v likes candidate c_1 the most, and c_4 the least. This results in a plot on 3-axes: (b_1, b_2, b_3) representing the number of observed ballots supporting candidates 1, 2, and 3 (ballots supporting 4 may be inferred). We project this to the 2 dimensional heatmap shown in Figure 5. Along the x-axis, we plot b_1, the number of ballots the voter observes supporting her favorite candidate; along the y-axis, b_2, the support for her second choice. Each cell represents multiple points in the observation space. For instance, the cell $(4, 5)$ corresponds to all observations $(4, 5, b_3)$, where $b_3 \in \{0, 1, \ldots 16\}$. Cells in solid green represent conditions where `Full` Voter always casts a sincere ballot for candidate 1; Cells in solid white represent conditions where `Full` Voter never votes sincerely. Because the heatmap is a projection from a higher dimensional plot, each cell may represent more than one possible observation. When `Full` votes sincerely in only some of those observations, the cell is shaded in lighter green. The triangle on the bottom right, show in gray, are inadmissible conditions where the total number of ballots supporting candidate 1 and 2 exceed 25.

Naturally, as b_1 increases, `Full` will have a tendency to vote for 1 as well. The large triangular cutout on the left shows when candidate 2 has enough support that v will change to a strategic vote for 2. The upper left region shows situations where neither 1 nor 2 have much support, and the election is a race between 3 and 4. This plot shows that the `Full` Voter tends to vote sincerely when her favorite candidate has even moderate amount of support (to bolster her chance for victory), or when the race is between her top and second choices. She only votes strategically when she believes the likely winners do not include candidate 1.

Figure 6 uses the same axes, and highlights the conditions where `TieH` ($L = 1000$) *disagrees* with `Full`. Cells in red show *increased* preference for sincere voting; cells in green show *decreased* preference for sincere voting. Cells in white show the two models in general agreement (i.e. the observations represented by the cell result in the same number of wins for each candidate, though not necessarily for the same observations). Due to the stochastic nature of `TieH`, the instances of disagreement are spread out in the observation space. However, there is a trend for them to concentrate near the borders where the `Full` Voter transitions between different ballots. This is more clearly seen in Figure 7, which maps the same differences between `Poisson` and `Full`. Here, we see that `Poisson` systematically casts more sincere votes, and the disagreements concentrate on the transition between sincere voting and strategically voting for the second choice.

4 ADDRESSING THE DESIDERATA

We return to the desiderata proposed earlier in this paper and consider how well our proposed heuristics implement them. Figure 8 summarizes the degree to which each proposed heuristic fulfills the desiderata, including a column for the `Full` Voter Model, and a row for a heuristic's Fidelity when compared to `Full`. By their design, the heuristics adhere to the Knowledge requirement, that each voter acts only upon observations gleaned from her social network. Similarly, all heuristics fulfill the Anonymity and Optimistic

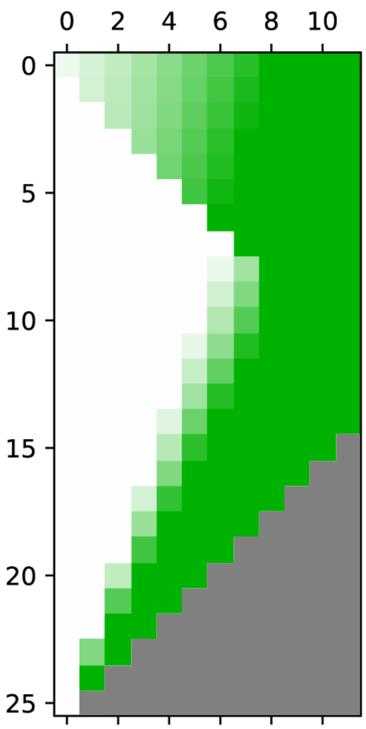

Figure 5: Conditions where `Full` votes sincerely in green.

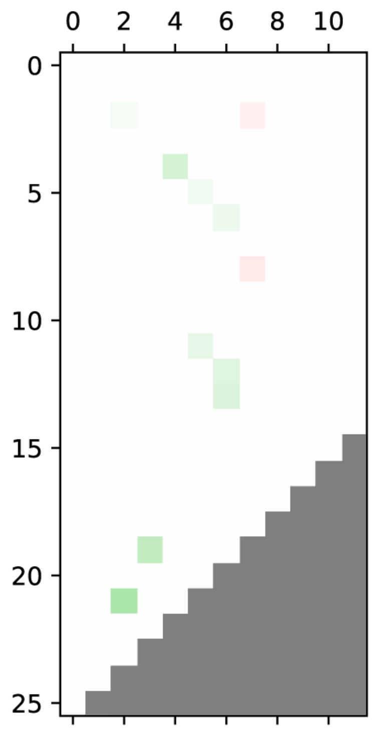

Figure 6: Conditions where `TieH` disagrees with `Full` in red or green.

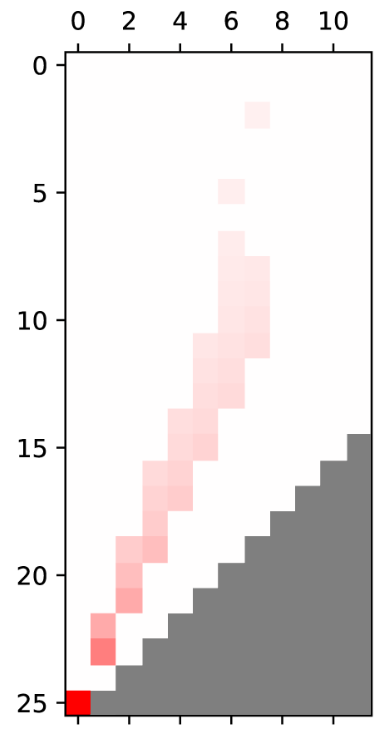

Figure 7: Conditions where `Poisson` disagrees with `Full` in red or green.

	`Full`	`Top-K`	`Max-M`	`TieU`	`TieH`	`TieR`	`Poisson`
Knowledge	✓	✓	✓	✓	✓	✓	✓
Rationality	✓	~	~	~	~	~	~
Anonymity	✓	✓	✓	✓	✓	✓	✓
Equilibrium	*	*	*	*	*	*	*
Tractability	✗	✗	✗	*	✓	✓	✓
Optimistic	✓	✓	✓	✓	✓	✓	✓
Fidelity	–	✓	✓	✓	✓	✗	✓

Figure 8: Adherence to the Desiderata by the Heuristics. ~ indicates approximately rational. The asterisks in the Equilibrium row indicate each algorithm will either converge to an equilibrium or display a cyclic behavior. The symbols in the Tractability row provide rough indicators on the speed of the algorithms, ranging from slow (✗), to moderate (*), to fast (✓).

criteria. Our models also fulfill the Equilibrium desideratum in the sense that it terminates when the population converges, or exhibits cyclic behavior to show an equilibrium cannot be reached.

As explained earlier, the `Full` Voter Model is based on rational actions of the voter, and so fulfills the Rationality desideratum. The other models approximate this rational behavior to varying degrees of fidelity. With the exception of the `TieR` models, all our proposed heuristics replicate the decisions of the fully rational voter most of the time (about 88% to 90% for $m = 5$ and $n = 500$).

The main difference between our models is in the Tractability desideratum. As noted before, the `Full` Voter Model scales very poorly with both the number of candidates m and the number of other voters n. The heuristics `Top-K`, `Max-M` and `TieU` simplify this computation in different ways, but still require this enumeration. As a result, they scale poorly to larger elections. The `TieR` heuristic bypasses the need for enumerating pivot space by using Monte Carlo sampling on the entire outcome space; this is fast, but unfortunately loses accuracy quickly in larger elections due to the sparsity of pivot outcomes. On the other hand, the `TieH` heuristic generates more consistent runtimes by allow the heuristic to *not* sample uniformly from the pivot space. The cost in accuracy for making this assumption is small, while the performance gains are enormous. Surprisingly, the `Poisson` heuristic also performs very well in our experiments. While it makes fundamentally different assumptions about the voters, it produces results that are comparable in accuracy to `TieH` in a negligible runtime. Thus, `Poisson` and `TieH` are the heuristic models that fulfill the best combination of desiderata for use in larger simulations.

5 CASE STUDY: THE MICROMEGA RULE

In this section, we illustrate an application of our heuristics by studying the Micromega Rule. In particular, we focus on the effects of different population sizes, an aspect that relies on our heuristics' ability to scale to larger graphs.

In political science, the Micromega rule frequently deals with districting and systems of proportional representation. While it is intuitive that properties of the voting rules determine the qualities of successful parties, Josep Colomer [7] posits that this influence runs both ways: that existing political parties will favor electoral rules that improve their future electoral chances. In particular, he formulated the Micromega rule, which predicts that a government comprised of few, large parties will favor smaller assemblies and larger districts, while those formed from smaller parties will favor assemblies with many seats and smaller individual districts. These topics have been explored in computational social choice by Bachrach, Lev, Lewenberg, and Zick [2]. They compare the outcomes from a district based election to that of a popular vote. They devise the Misrepresentation Ratio to measure this deviation, and show that misrepresentation occurs in voter populations of all sizes.

We explore a variation of the Micromega rule. We posit that large parties are more effectively able to consolidate their voter base in *large electoral districts* (i.e. those having a larger population of voters), while less populous districts will see more continued support for less popular candidates. It is this continued support that allows the party to remain viable in future elections. The use of voter heuristics is essential to this exploration of the Micromega rule because our hypothesis depends on population size. Without using heuristic voter responses, it would not be feasible to simulate communities of any significant size, and we would not be able to sample election results from a large enough range of community sizes of test our hypothesis.

We present our electoral districts as simulated social networks. We use directed and undirected versions of the Erdős-Rényi (ER) random graph model [9], and the Barabási-Albert (BA) preferential attachment model [1]. In the (directed) ER graph with density parameter pr, every (directed) edge exists with probability pr. To create a (directed) BA graph with attachment parameter pr, we recursively add new nodes to an existing graph, attaching it to pr existing vertices via a (directed) edge; the existing vertices are picked with probability proportion to their (in-)degree. We consider graphs of sizes $n \in \{200, 400, 600, 800\}$. We focus on the scenario with $m = 5$ candidates. We fix the average in-degree of each node to be approximately $d = 30$.[4] A value of $d = 30$ was selected as a reasonable number of informational sources voters may consider in such a scenario. Increasing d allows voters to sample more information from the network, and therefore give them a more precise estimate on the outcome of the election. This will likely increase the amount of strategic play in all conditions, since voters will be less optimistic that a favorite candidate will recover if they are behind in the polls. However, we do not believe changes to d will significantly affect the qualitative patterns that emerge.

Based on our results from the previous section, we will use both the `Poisson` and `TieH` models for this simulation. The `Poisson` model is parameter free, though it exhibits systematic bias as compared to the `Full` voter; for the `TieH` model, we set the sample size parameter at $L = 2000$. To measure the degree of support for less popular parties, we take the **SF Ratio**, the ratio of support between the second- and first-runner up candidates [8]. Each data point is the average of 200 replications.

[4]Importantly, the parameter pr is *doubled* when constructing directed BA graphs.

5.1 Results

Figure 9 plots the average SF-Ratio for each condition under the `TieH` model. Here, we observe clear downward trends in the directed ER and BA graphs — the amount of support for the third place candidate diminishes exponentially with increasing voter population, reflecting an exponentially increasing ability for voters to vote strategically. Most interesting is the difference in SF Ratios of the ER and dER graphs. Structurally, there is only one difference between these two models. In the ER graph, influence is reciprocal – if v observes another voter, then that other voter also observes v. The same is not generally true in a dER graph, which allows their voters an increased ability to communicate and propagate information through strategic play.

The most significant downward trend is observed in the dBA graph. The dBA model generates directed acyclic graphs that have a strongly hierarchical structure. Information in these graphs flow from the older and higher degree nodes, toward the younger, lower degree nodes. This hierarchical structure prevents effective communication, which manifests in the high SF Ratios, particularly in smaller graphs.

Figure 10 plots the average SF-Ratios under the `Poisson` model. We observe the same strong trend in the directed BA graphs. However, there is no discernible pattern in any of the other graph types as the population size changes (data not shown). It is possible that this difference is due to the systematic bias in how the `Poisson` model attempts to estimate the `Full` voter's behavior, or that the trend observed in the directed ER graphs is relatively fragile compared to that of the directed BA graph. This can be interpreted as an alternate cause of the Micromega Rule — that large electoral districts facilitate the consolidation of support for larger parties in certain types of social networks, where as districts with smaller voting populations show more support for "fringe" candidates.

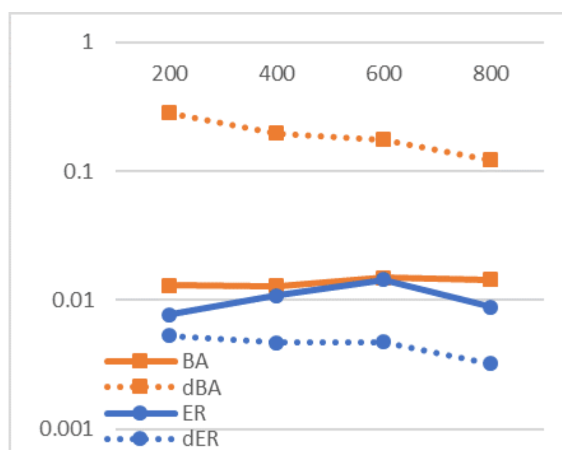

Figure 9: Average SF Ratios of graphs of different sizes under the `TieH` model. Note y-axis is in log scale.

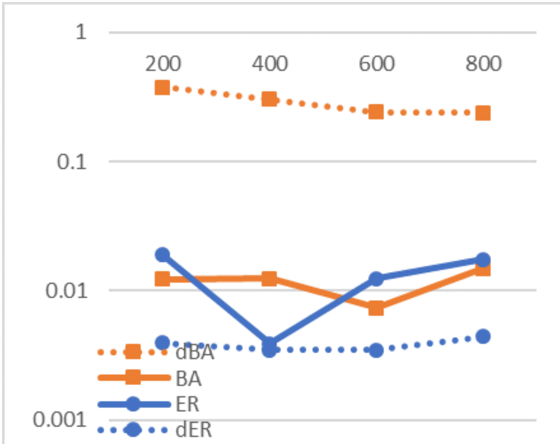

Figure 10: Average SF Ratios of graphs of different sizes under the `Poisson` models. Note y-axis is in log scale.

It is worthy to note that the undirected BA graphs show no significant trend in either direction. Amongst the models used, the BA preferential attachment model may be the most representative of real world social networks. So while we demonstrate an alternate cause for the Micromega rule in certain types of social networks, the result does not necessarily generalize to real world networks.

6 CONCLUSION

In this paper, we proposed a number of heuristic voter models for strategic voting in social networks. While the `Full` Voter model in [25] works well in small graphs, the exact computation of expected utilities proves infeasible for larger graphs. Voters in our models are boundedly rational and our heuristics lighten their cognitive burden in ways that would be natural for a human voter. Our heuristics perform up to 2 orders of magnitude faster, and retain a high level of fidelity when compared to the `Full` Voter model. As a result, this allows us to highlight the differences in strategic behavior when voters are part of populations of different sizes.

To illustrate this, we use our heuristic voter model to investigate the Micromega rule. We show that in certain networks, particularly the directed Erdős-Rényi and directed Barabási-Albert models, smaller populations offer more support for fringe candidates than larger electorates. The orientation of directed edges in the dBA graphs lends it a strict hierarchical structure, which reinforces the Micromega rule dramatically in our simulations. Other preferential attachment models such as Bollobás's scale-free graphs offer parameters which may be tuned to allow for different degrees of hierarchy [5]. It would be interesting to explore the impact of hierarchical structure on strategic voting in future work.

Our heuristic models also pave the way to simulating strategic voting behavior in truly large scale networks. This opens up the possibility of simulating on real world datasets, where nodes number in the millions. We may also consider comparing the results of our models to human voting behavior in controlled settings. Moreover, our model could be extended to include other scoring rules, such as Borda and k-Approval, and other social choice functions in general.

REFERENCES

[1] Réka Albert and Albert-László Barabási. 2002. Statistical mechanics of complex networks. *Reviews of modern physics* 74, 1 (2002), 47.

[2] Yoram Bachrach, Omer Lev, Yoad Lewenberg, and Yair Zick. 2016. Misrepresentation in district voting. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*.

[3] Duncan Black. 1948. On the rationale of group decision-making. *The Journal of Political Economy* (1948), 23–34.

[4] Paolo Boldi, Francesco Bonchi, Carlos Castillo, and Sebastiano Vigna. 2011. Viscous Democracy for Social Networks. *Commun. ACM* 54, 6 (2011), 129–137.

[5] Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan. 2003. Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 132–139.

[6] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489, 7415 (2012), 295–298.

[7] Josep M Colomer. 2005. It's parties that choose electoral systems (or, Duverger's laws upside down). *Political Studies* 53, 1 (2005), 1–21.

[8] Gary W Cox. 1997. *Making votes count: Strategic coordination in the world's electoral systems*. Vol. 7. Cambridge Univ Press.

[9] Paul Erdős and Alfréd Rényi. 1961. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica* 12, 1 (1961), 261–267.

[10] William Feller. 1968. *An Introduction to Probability Theory and Its Applications*. Vol. 1. John Wiley & Sons New York.

[11] Umberto Grandi. 2017. Social choice and social networks. *Trends in Computational Social Choice. AI Access* (2017), 169–184.

[12] Melvin J Hinich and Michael C Munger. 1997. *Analytical politics*. Cambridge University Press.

[13] Govinda M Kamath, Eren Şaşoğlu, and David Tse. 2015. Optimal haplotype assembly from high-throughput mate-pair reads. In *Information Theory (ISIT), 2015 IEEE International Symposium on*. IEEE, 914–918.

[14] Adam DI Kramer. 2012. The spread of emotion via Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 767–770.

[15] Michael F Meffert and Thomas Gschwend. 2009. Strategic voting under proportional representation and coalition governments: a simulation and laboratory experiment. *Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung* 7 (2009).

[16] Reshef Meir, Omer Lev, and Jeffrey S Rosenschein. 2014. A local-dominance theory of voting equilibria. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*. ACM, 313–330.

[17] Reshef Meir, Maria Polukarov, Jeffrey S Rosenschein, and Nicholas R Jennings. 2010. Convergence to equilibria in plurality voting. In *Proc. of 24th Conference on Artificial Intelligence (AAAI-10)*. 823–828.

[18] Roger B Myerson. 2000. Large poisson games. *Journal of Economic Theory* 94, 1 (2000), 7–45.

[19] Roger B Myerson and Robert J Weber. 1993. A Theory of Voting Equilibria. *American Political Science Review* 87, 01 (1993), 102–114.

[20] Svetlana Obraztsova, Omer Lev, Maria Polukarov, Zinovi Rabinovich, and Jeffrey S Rosenschein. 2015. Farsighted voting dynamics. In *AGT@ IJCAI workshop, Buenos Aires, Argentina*.

[21] Zinovi Rabinovich, Svetlana Obraztsova, Omer Lev, Evangelos Markakis, and Jeffrey S Rosenschein. 2015. Analysis of Equilibria in Iterative Voting Schemes.. In *AAAI*, Vol. 15. Citeseer, 1007–1013.

[22] Annemieke Reijngoud and Ulle Endriss. 2012. Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 635–644.

[23] Amirali Salehi-Abari and Craig Boutilier. 2014. Empathetic Social Choice on Social Networks. In *Proceedings of The 13th International Conference on Autonomous Agents and Multiagent Systems*. 693–700.

[24] Amirali Salehi-Abari and Craig Boutilier. 2015. Preference-oriented Social Networks: Group Recommendation and Inference. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys-15)*.

[25] Alan Tsang and Kate Larson. 2016. The echo chamber: Strategic voting and homophily in social networks. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 368–375.

ACKNOWLEDGMENTS

This research was made possible by the funding and support from the Natural Sciences and Engineering Research Council of Canada (NSERC).