

# Should Social Network Structure Be Taken into Account in Elections?\*

Vincent Conitzer  
Duke University

April 2, 2011

## Abstract

If the social network structure among the voters in an election is known, how should this be taken into account by the voting rule? In this brief article, I argue, via the maximum likelihood approach to voting, that it is optimal to ignore the social network structure altogether—one person, one vote.

## 1 Introduction

The maximum likelihood framework provides a natural approach to determining an “optimal” voting rule. Under this approach, we suppose that there exists a “correct” outcome, which cannot be directly observed; however, voters’ votes are noisy estimates of this correct outcome. For example, in an election between two alternatives, we may suppose that one of the alternatives is better in some objective sense, and each voter is more likely to vote for the correct (better) alternative than for the incorrect alternative. To make this concrete, suppose that there is a fixed  $p > 1/2$  such that each voter, independently, votes for the correct alternative with probability  $p$ , and for the incorrect one with probability  $1 - p$ .

Under such a model, it makes sense to choose the *maximum likelihood estimate* of the correct outcome. For example, suppose that we have a specific profile in which there are  $n_a$  votes for alternative  $a$ , and  $n_b$  votes for alternative  $b$ . The likelihood of this particular profile happening given that  $a$  is the correct outcome is  $p^{n_a}(1 - p)^{n_b}$ , and given that  $b$  is the correct outcome it is  $(1 - p)^{n_a}p^{n_b}$ . If  $n_a > n_b$ , then the former is larger, so  $a$  is the maximum likelihood estimate; and vice versa. That is, the maximum likelihood estimate here is obtained by applying the majority rule: the alternative with more votes wins.

The study of this maximum likelihood approach to voting dates back to Condorcet [4], who proposed a noise model that generalizes the one above to more than

---

\*The technically rather simple observation presented in this short article was presented in full in a short rump session talk at the 2010 Dagstuhl seminar on the Computational Foundations of Social Choice. This work has been supported by an Alfred P. Sloan Fellowship and by NSF under award numbers IIS-0812113 and CAREER-0953756.

two alternatives. Two centuries later, Young [11, 12] showed that the maximum likelihood solution for this model coincides with the Kemeny rule [8]. A recent body of work has been devoted to the investigation of how alternative noise models lead to different rules [5, 3, 10, 2]. There is also a literature on *strategic* voting in such contexts—e.g., [1, 7]—but we will not consider such issues here. Another recent paper [6] explores the relationship between the maximum likelihood framework and the *distance rationalizability* framework, where we find the closest “consensus” election that has a clear winner, for some definition of consensus and some distance function, and choose its winner.

A common assumption in the existing work on the maximum likelihood approach is that votes are drawn independently (conditional on the correct outcome). It is easy to dispute this assumption. Voters do not make up their minds in a vacuum: they generally discuss the matter with others, and this is likely to affect their votes. Specifically, voters that talk to each other are more likely to reach the same conclusion, whether right or wrong.

This brief article considers elections in which there is some *social network structure* among the voters, representing which voters interact with which other voters (as well as how frequently or intensively they do so); and, moreover, this social network structure is known to the election designer. Hence, *the voting rule can take the social network into account*. This is not always as impractical as it may sound. For example, Facebook, Inc. decided to allow its users to vote on its future terms of use [13]. Naturally, Facebook knows the social network structure among its users very accurately. But, in order for the general idea to make sense, it is not necessary to have this level of detailed knowledge of the social network structure among the voters. For example, we may simply observe that the residents of the state of Hawaii are in some ways disconnected from the continental United States, so that there is perhaps less interaction between the two groups; and perhaps it makes sense to somehow take this into account in the design of a voting rule. Making such a statement can hardly be considered a taboo, since a voter’s state of residence is already taken into account via the electoral college. In fact, one may even go so far as to wonder whether a version of the electoral college is perhaps the maximum likelihood solution for a particular probabilistic model.

So, given that we know the social network structure (or some estimate thereof), how should this affect the voting rule? We can state various suggestive intuitive arguments, such as the following.

1. *Well-connected voters benefit from the insight of others so they are more likely to get the answer right. They should be weighed more heavily.*
2. *Well-connected voters do not give the issue much independent thought; the reasons for their votes are already reflected in their neighbors’ votes. They should be weighed less heavily.*
3. *It is oversimplifying matters to discuss the matter merely in terms of voters’ weights. We can do better by considering richer options, such as an electoral college.*

In this brief article, I will argue that all of these three suggestions are wrong—or, arguably, 1 and 2 are both correct, but cancel each other out. That is, optimal rules ignore the social network structure altogether, affirming the familiar principle of *one person, one vote*. The mathematical result is quite trivial once the model has been set up.

## 2 Model

In this section, I give a concrete probabilistic model that takes the social network structure into account; I will discuss the benefits and drawbacks of this model in the final section. Let  $v$  denote a vertex in the social network graph, corresponding to a voter. Let  $N(v)$  denote  $v$ 's neighbors. Let  $V_v$  denote voter  $v$ 's vote, and  $V_{N(v)}$  the subprofile consisting of the votes of all the neighbors of  $v$ . Let  $c$  denote the “correct” outcome. For ease of reading, the reader is encouraged to keep the simple example of an election between two alternatives in mind, so that there are two outcomes and two possible votes. However, this is not at all necessary for the main result: there can be arbitrarily many alternatives, and in fact we will not require any restriction on the outcome space, on the space of possible votes, or on the relationship between the two. For example, an outcome could consist of an individual alternative, a subset of the alternatives, or a ranking of the alternatives; and so could a vote, independently of what an outcome consists of.

For the maximum likelihood approach, we are interested in the probability of the full profile of votes, given  $c$ . I will assume that this probability can be factored as  $\prod_v f_v(V_v, V_{N(v)}|c)$ , where  $f_v$  is a function associated with vertex  $v$ . (The dependence of  $f$  on  $V_{N(v)}$  captures that voter  $v$ 's vote can be affected by her neighbors' opinions (votes)—as opposed to more standard models that take the form  $\prod_v f_v(V_v|c)$ .) This type of factorization assumption is commonly made in undirected graphical models of probability distributions (Markov random fields).<sup>1</sup> It is worth noting that, in the special case where the social network graph is a clique, this assumption is not restrictive at all. I now make the following assumption which does introduce a restriction that is crucial for the result.

**Assumption 1** *For every  $v$ , there exist functions  $g_v$  and  $h_v$  so that  $f_v$  can be factored as  $f_v(V_v, V_{N(v)}|c) = g_v(V_v|c) \cdot h_v(V_v, V_{N(v)})$ .*

Here,  $g_v$  is intended to capture the effect that a voter is more likely to vote for the correct outcome, and  $h_v$  to capture the effect that voters are likely to vote similarly to their neighbors. The key assumption is that  $g_v$  does not depend on  $V_{N(v)}$ , and  $h_v$  does not depend on  $c$ . That is, the tendency to vote for the correct outcome is in a sense independent from the tendency to agree with one's neighbors.

For example, in the two-alternative case, Assumption 1 precludes models where, given that all of a vertex's neighbors vote for the *correct* outcome, this vertex also does

<sup>1</sup>In this literature, factors can only be associated with *cliques* of the graph. In contrast, here, a factor can be associated with the neighborhood of any vertex. Since any clique is contained in the neighborhood of any one of its vertices, the model studied here is in a sense less restrictive—though this is in any case not the main assumption of this article, which follows next.

so with probability 1; but given that all of the neighbors vote for the *incorrect* outcome, this vertex still has a positive probability of choosing the correct outcome. If this were the noise model, then if we found a vertex voting differently from all its neighbors, we should conclude that that vertex has identified the correct outcome, no matter what everyone else votes. So, in that case, the social network structure would matter.

### 3 Detailed example

For concreteness, let us consider a simple example with two vertices (voters) connected by an edge, voting over two alternatives (so a vote corresponds simply to one of the alternatives). Let  $-c$  denote the alternative other than  $c$  and let  $-v$  denote the voter other than  $v$ . For each of the two vertices  $v$ , we let  $g_v(V_v = c|c) = .7, g_v(V_v = -c|c) = .3$ , indicating that votes for the correct outcome are more likely. We also let  $h_v(V_v = c, V_{-v} = c) = 1.142, h_v(V_v = c, V_{-v} = -c) = .762$ , indicating that voters are more likely to agree with each other. We can now calculate that  $P(V_v = c|c) = P(V_v = c, V_{-v} = c|c) + P(V_v = c, V_{-v} = -c|c) = .7 \cdot 1.142 \cdot .7 \cdot 1.142 + .7 \cdot .762 \cdot .3 \cdot .762 = .761$ .

In contrast, a model in which the two voters do not interact can be obtained by setting  $h = 1$  everywhere; in this model,  $P(V_v = c|c) = .7$ . Thus, there is some truth to the first argument in the introduction: well-connected voters benefit from the insight of others so they are more likely to get the answer right. However, it does not yet follow from this that they should be weighed more heavily, and as we will see this is in fact not the case.

### 4 Main result

We now arrive at the main result, which states that (under Assumption 1) the social network structure should not be taken into account by the voting rule.

**Proposition 1** *Under Assumption 1, the maximum likelihood estimate does not depend on the social network structure—more precisely, it does not depend on the functions  $h_v$  (assuming they are positive everywhere). Specifically, under these conditions, the maximum likelihood estimate of the correct outcome is  $\arg \max_c \prod_v g_v(V_v|c)$ .*

As noted before, the proof is quite trivial:

**Proof:** By Assumption 1, we can write the likelihood as

$$\prod_v f_v(V_v, V_{N(v)}|c) = \prod_v g_v(V_v|c) \cdot h_v(V_v, V_{N(v)}) = \left( \prod_v g_v(V_v|c) \right) \cdot \left( \prod_v h_v(V_v, V_{N(v)}) \right)$$

Therefore, the maximum likelihood estimate of the correct outcome is

$$\arg \max_c \prod_v f_v(V_v, V_{N(v)}|c) = \arg \max_c \left( \prod_v g_v(V_v|c) \right) \cdot \left( \prod_v h_v(V_v, V_{N(v)}) \right) = \arg \max_c \prod_v g_v(V_v|c)$$

which does not depend on the  $h$  functions. ■

Because of Proposition 1, results derived in the maximum likelihood framework with the standard independence assumption still hold in our setting. For example:

**Corollary 1** *In a two-alternative election where Assumption 1 holds, if  $g_v$  is the same for all voters  $v$  (so  $g_v = g$ ), and  $g(c|c) > g(-c|c)$ , then the majority rule gives the maximum likelihood estimate of the correct winner.*

**Proof:** By Proposition 1, the maximum likelihood estimate of the correct winner is  $\arg \max_c \prod_v g(V_v|c)$  (regardless of the  $h_v$  functions). The remainder of the proof is as described in the introduction, letting  $p = g(c|c)$ . ■

This also provides justification for the strong-seeming assumption of independence across voters in the maximum likelihood voting literature, showing that modeling interaction among the voters does not necessarily change the solution.

## 5 Discussion

As noted earlier, while the result in this brief article is presumably easiest to illustrate in the context of an election between two alternatives, everything in the article applies just as well to settings with more alternatives. Also, there is no requirement that votes have to take any particular form—a vote can be a ranking of the alternatives, a subset of approved alternatives, or anything else. Similarly, the correct “outcome” can be an alternative, a ranking of the alternatives, or anything else. Finally, no particular structure on the social network graph is required for the result—for example, all voters may be directly connected to each other. The crucial assumption is Assumption 1, which separates the correct-outcome effect from the interaction effect. An argument for voting rules that take social network structure into account could be based on models that violate Assumption 1.

In this article, I have not proposed any concrete mechanism<sup>2</sup> by which voters influence their neighbors’ opinions (for example, voters could sequentially send each other particular kinds of messages before voting); all that is needed is Assumption 1 on the distribution. On the one hand, it can be argued that it is an advantage of the model that it does not rely on such a concrete mechanism, because any analytically tractable mechanism will presumably be at best a very rough approximation of the truth. On the other hand, it still seems desirable to propose and analyze such concrete mechanisms, to see whether they violate the assumptions made here, and, if so, what their effect is on the maximum likelihood approach. In this sense, this brief article is potentially but a first step in a richer body of research.

It is instructive to compare the result in this article to a result by Nitzan and Paroush [9]. They consider a model with two alternatives where some voters are more likely to identify the correct alternative than others, and find that the optimal rule is a *weighted* majority rule, where more skillful voters receive a larger weight. In contrast, in the framework of this article, better-connected voters are more likely to vote correctly (as we saw

<sup>2</sup>“Mechanism” here is in the common sense of the word, not a reference to mechanism design.

in Section 3), but are not given a correspondingly larger weight (Corollary 1). There is no contradiction here: in the Nitzan-Paroush model, voters vote independently—there are no edges in the graph. Hence, more skillful voters really do contribute more valuable information, and should be weighed more heavily. In contrast, in the framework of this article, better-connected voters are not really contributing more valuable information, because some of their information is already reflected in their neighbors' votes. In fact, the Nitzan-Paroush framework can easily be accommodated in the framework of this article, by reflecting skill in the  $g$  functions. If we do so, then there are two reasons that a voter may be more likely to vote correctly than others: one is being more skillful, and the other is being better connected. By Proposition 1, the former results in a larger weight (exactly as in the result by Nitzan and Paroush), but the latter does not—a generalization of Corollary 1.

## References

- [1] David Austen-Smith and Jeffrey S. Banks. Information aggregation, rationality, and the Condorcet Jury Theorem. *American Political Science Review*, 90:34–45, 1996.
- [2] Vincent Conitzer, Matthew Rognlie, and Lirong Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, Pasadena, CA, USA, 2009.
- [3] Vincent Conitzer and Tuomas Sandholm. Common voting rules as maximum likelihood estimators. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 145–152, Edinburgh, UK, 2005.
- [4] Marie Jean Antoine Nicolas de Caritat (Marquis de Condorcet). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. 1785. Paris: L'Imprimerie Royale.
- [5] Mohamed Drissi-Bakhkhat and Michel Truchon. Maximum likelihood approach to vote aggregation with variable probabilities. *Social Choice and Welfare*, 23:161–185, 2004.
- [6] Edith Elkind, Piotr Faliszewski, and Arkadii Slinko. Good rationalizations of voting rules. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 774–779, Atlanta, GA, USA, 2010.
- [7] Timothy Feddersen and Wolfgang Pesendorfer. The swing voter's curse. *American Economic Review*, 86:408–424, 1996.
- [8] John Kemeny. Mathematics without numbers. *Daedalus*, 88:575–591, 1959.
- [9] Shmuel Nitzan and Jacob Paroush. Optimal decision rules in uncertain dichotomous choice situations. *International Economic Review*, 23(2):289–297, 1982.

- [10] Michel Truchon. Borda and the maximum likelihood approach to vote aggregation. *Mathematical Social Sciences*, 55(1):96–102, 2008.
- [11] H. Peyton Young. Condorcet’s theory of voting. *American Political Science Review*, 82:1231–1244, 1988.
- [12] H. Peyton Young. Optimal voting rules. *Journal of Economic Perspectives*, 9(1):51–64, 1995.
- [13] Mark Zuckerberg. Voting begins on governing the Facebook site, 2009. <http://blog.facebook.com/blog.php?post=76815337130>.