A networked voting rule for democratic representation

Alexis R. Hernández¹,², Carlos Gracia-Lázaro²,
Edgardo Brigatti¹ and Yamir Moreno²,³,⁴

¹Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
²Institute for Biocomputation and Physics of Complex Systems (BIFI), and
³Department of Theoretical Physics, Faculty of Sciences, Universidad de Zaragoza, Zaragoza, Spain
⁴ISIFoundation, Turin, Italy

We introduce a general framework for exploring the problem of selecting a committee of representatives with the aim of studying a networked voting rule based on a decentralized large-scale platform, which can assure a strong accountability of the elected. The results of our simulations suggest that this algorithm-based approach is able to obtain a high representativeness for relatively small committees, performing even better than a classical voting rule based on a closed list of candidates. We show that a general relation between committee size and representatives exists in the form of an inverse square root law and that the normalized committee size approximately scales with the inverse of the community size, allowing the scalability to very large populations. These findings are not strongly influenced by the different networks used to describe the individuals’ interactions, except for the presence of few individuals with very high connectivity which can have a marginal negative effect in the committee selection process.

1. Introduction

The selection of an exemplar group of representatives to make decisions on behalf of a larger community is a widespread and critical problem for human societies [1,2]. Examples can be found in the election of a legislative assembly in indirect democracies, in the elections for the trade union, for supervisory or faculty board, for executive officers or non-governmental organization boards. The most widely used electoral systems can be classified into one of the following groups: first-past-the-post, two-round systems, proportional representation, ranked voting or in a mix of two or more of the previous groups [3–8]. In general, these systems seek to strike a balance between representativeness and effectiveness [9]. In the vast majority of electoral systems, including fully proportional representation...
systems [10,11], representatives gain a power of representation that is not completely proportional to the number of voters they represent, but rather the result of a given granularity. Interestingly, this problem is also relevant for artificial systems, such as software multiagent systems, i.e. in recommendation systems [12,13], mobile networks [14] and election-based mechanisms [15] for distributing data over an overlay P2P network [16].

In the last years, a general interest in the quantitative description, based on mathematical models, of social processes related to election procedures increased significantly. One of the reasons can be related to the successful introduction of agent-based modelling. In fact, agent-based models (ABM) provide a flexible and powerful theoretical framework for describing these phenomena [17–19]. One of the first opinions ABM was proposed by Clifford & Sudbury [20]. Although, at first, it was designed as a theoretical tool for studying the competition of species, it has been named voter model for the natural application to the dynamics of voting systems [21]. The voter model has been studied under different topologies [22–25] and conditions [26–28]. Nevertheless, although it can capture statistical features of real-world elections [29], its intrinsic limitations, such as the consideration of a unique cultural variable that can only take two values, restrict its applicability. Subsequently, new social models have been introduced to study the voting dynamics [30–32].

In this work, we analyse a quite different problem related to election procedures, which has not been addressed in previous works which use ABM. In fact, we study the general problem of selecting a group of candidates that best represents the voters. We consider systems where each voter is allowed to vote for only one candidate and the elected are the ones who obtain a better rank among their counterparts. In particular, we focus on the case of multi-winner elections, that is choosing a collective body of a given size (a committee). We model an idealized situation where voters are rational individuals, which means that they make a decision to maximize their representation, and they present a general knowledge of the candidates and direct access to them.

In classical elections, a fixed number of candidates participate and voters rank the candidates expressing their preferences. We introduce a new formal model, where the list of candidates is not fixed in advance, but they emerge as a self-organized process controlled by the voting rules. Moreover, voters express not preferences, but opinions, which determine their indications about whom they would like to see as their representative.

Our model introduces new mechanisms which give a fundamental importance to the accountability of the elected committee. The connection between representatives and constituents is fundamental and it is the basis for accountability, allowing us to check for incompetence and corruption. In classical voting systems, this link can be generally insured only for a small-sized community. In particular, for the national legislative assembly, it can be partially controlled by the small size of the electoral district. Our model introduces a radical difference for obtaining an efficacious accountability. In fact, it takes into account individuals’ first-hand trust relationship as a key ingredient to determine the elected representatives. Votes are assigned on the basis of a self-declared confidence circle, which is a network of trusted individuals which can be implemented on an online platform.

After having implemented this new voting rule, its effects are tested modelling the behaviour of the selected committee. The committee runs a series of ballots making choices about different issues. The quality of the elected committee is numerically assessed based on how much their final decisions are consistent with the personal opinions of the community. Note that many works study scenarios where representatives make decisions which are compared with some objective truth [1,33]. By contrast, as in [34], our approach is interested in discriminating the selected boards which best represent the community opinions.

2. The model

The system is composed of a population of $N_e$ electors and an internet-based platform. The platform allows the voters to self-declare who belongs to their confidence circle, which is a network of trusted individuals. The same platform is used by voters to manifest their opinions on $N_i$ issues. Issues are organized in questions which can be defined by a committee or by means of a self-organized process internal to the community. The answers of each individual $j$ are organized in a vector $v_j$. The vector is composed of $N_i$ cells and each cell can assume the value 1 if the answer is positive, $-1$ if it is negative or 0 if the question is left unanswered.

The following step allows us to find the better representative for each confidence circle. First, we consider an individual $j$ and we compute the vectors overlaps with all his neighbours $k$. This is obtained
Figure 1. Schematic of the vote process. Nodes stand for the individuals; the red ones belong to a cycle and will be confirmed as representatives if they collect more votes than the established threshold. The big numbers associated with the nodes represent the received cumulated votes. Arrows stand for the indication of each individual and the small numbers associated with them represent the number of transferred votes. Dotted arrows belong to a cycle, where there is no cumulative transfer of votes.

using the following expression:

\[
v^j \# v^k = \frac{\sum_{m=1}^{N_i} (v^j_m \cdot v^k_m) \delta(v^j_m, v^k_m)}{\sum_{m=1}^{N_i} (v^j_m \cdot v^k_m)^2},
\]

where the numerator counts the number of questions answered in the same way (only yes or not) and the denominator counts the number of questions answered simultaneously by both individuals; \(\delta\) stands for the Kronecker delta which is 1 if \(v^j_m = v^k_m\) and 0 otherwise. Each individual \(j\) will indicate as his representative the individual \(k'\) for which \(v^j \# v^{k'}\) is maximum. In the case where more than one individual generates the same maximum overlap value, the individual with a higher connectivity is chosen as the representative. For the exceptional case when also the connectivity is equal, the representative is randomly selected between the similar ones.

After the selection of the representative \(k'\) for every voter \(j\), as constrained by his confidence network, the final step consists in choosing the aggregate of representatives of the entire community. To this end, we construct a directed graph where a node represents each individual and a directed link connects the individual with his personal representative. In this graph, which, in general, can be composed by different disconnected clusters, cycles are present. They represent individuals that have been mutually indicated by themselves. As all the individuals outside the cycles are represented by the individuals belonging to them, individuals who belong to cycles are the proper potential representatives for the community (figure 1).

As a final step, among the individuals belonging to a cycle, only the ones with a number of votes larger than a threshold \(\theta\) are indicated as representatives. Votes are counted considering the cumulative flow defined by the directed graph. If the individual \(j\) is pointing to \(z\), \(z\) receives all the votes previously received by \(j\) plus one. This flow of votes is computed only following links outside the cycles. Inside the cycles, only the single vote of an individual is counted. In this way, the number of representatives is reduced and results to be a fraction of the total individuals which belong to a cycle.

3. Results

3.1. General analysis

The individuals’ opinions in relation to the selected issues are randomly generated with the following rule: given an issue \(i\), an individual does not present an opinion \((v_i = 0)\) with probability \(\frac{1}{2}\). The
probability to have an opinion \(v_i = +1(-1)\), is \(\frac{1}{3} + \epsilon_i \left(\frac{1}{3} - \epsilon_i\right)\), where \(\epsilon_i\) is a random variable following a normal distribution with mean value equal to zero and \(\sigma^2 = 0.05\). Such a simplified characterization of the single opinion presents some contact with real political opinions which, frequently, are polarized, presenting a natural bimodality of preferences in political and economical issues [35].

The confidence circle of each individual is modelled generating a network where nodes represent individuals and links the trust relationships present in the community. The confidence circle of an individual is obtained selecting a node and considering its first neighbours. Note that an important simplification of this approach is the fact that it generates individuals with symmetric trust relationships.

In the following analysis, three types of networks are considered. Homogeneous random networks, implementing the Erdös–Rényi model [36], where the degree distribution is peaked around a typical value \(\langle k \rangle\), heterogeneous networks, using the Barabasi–Albert model [37], with a power-law degree distribution \(P(k) \propto k^{-3}\), and networks with the small-world property using the Watts and Strogatz model [38]. Our aim is not to model specific aspects of a real social network, but to use simple examples just to discuss the possible influence of some relevant network properties (such as the heterogeneity in the degree distribution, the average degree and the small-world property), on the behaviour of our model.

The exploration of the system behaviour can be obtained considering two fundamental observables. The first one is the normalized committee size which is measured as the ratio between the number of elected individuals \((E)\) and the total number of individuals of the community: \(F_{\text{rep}} = E/N_e\). To have manageable committees, a small value of \(F_{\text{rep}} = E/N_e\) is preferable. The second one is the representativeness. This is defined as the fraction of decisions expressed by the elected committee \((e_j)\) which matches with the community decisions \((c_j)\) over all the considered \(N_i\) issues: \(R = \sum_{j=1}^{N_i} \delta(e_j - c_j)/N_i\). Then, for \(R = 1\) we have a perfect committee, which make all the decisions in line with the popular will and for \(R = \frac{1}{2}\) (for binary decisions) we have a non-representative committee, whose decisions are completely uncorrelated to the popular will.

The decision expressed by the elected committee is obtained through a majority vote where each representative’s vote is weighted by the numbers of popular votes he received in the election procedure. The community decision is obtained by a direct process (plebiscite), where every individual votes in accordance with the opinion expressed in his vector \(v^j\). Note that if the individual has no opinion on a particular issue, he abstains from voting.

In figure 2, we show the representativeness \(R\) and the normalized committee size \(F_{\text{rep}}\) as a function of the threshold \(\Theta = \theta/N_e\) for different values of the number of issues \(N_i\), as obtained in a typical system with confidence circles defined from an Erdös–Rényi network. As expected, the representativeness and the normalized committee size decrease when increasing the threshold value. In particular, the decrease of the normalized committee size is very fast.

For fixed threshold values, decreasing the number of issues, quite intuitively, increases the representativeness. Similarly, it increases the normalized committee size. This last effect is not obvious.
Figure 3. (a) Representativeness versus normalized committee size, (b) logarithmic plot of $1 - R$ versus normalized committee size for $N_i = 40$. The continuous line has slope $-\frac{1}{2}$. For both figures, $N_e = 10\,000$ and confidence circles are defined from an Erdős–Rényi network with $\langle k \rangle = 40$. Results are averaged over 100 different realizations.

and it is produced by the fact that for a small number of issues the distribution of votes has a less pronounced peak and the threshold is not very efficient in selecting between the indicated individuals. Anyway, this tendency rapidly saturates and, for $N_i > 40$, the curves do not show any relevant dependence on this parameter.

The ideal committee corresponds to a small group of representatives which expresses a high level of representativeness. This is obtained by selecting an intermediate value for $\Theta$, which can be identified seeking for a representativeness close to 0.9, and looking at the corresponding committee size. For this reason, in the following we will plot the representativeness versus the normalized committee size, which allows a clear visualization of this fundamental relationship. In figure 3, we can observe that for fixed values of $R$, the normalized committee size increases when the number of issues increases. The relation between $R$ and $F_{\text{rep}}$, as reported in figure 3, can be characterized by a simple relation: $1 - R \propto 1/\sqrt{F_{\text{rep}}}$ of the obtained representativeness in relation to the ideal one scales as the square of the inverse of the normalized committee size. This means that, for example, for improving $R - 1$ by a factor of 2, the number of elected must quadruple. This relation can be justified considering that $1 - R$ is proportional to the error in the estimation of the mean opinion $v_i$ (the result of the plebiscite) using a sample of size $F_{\text{rep}}$. In the case where the estimation of the mean opinion uses an independent and identically distributed sample of size $n$, it is well known that the standard error of the sample mean scales as $n^{-1/2}$. Interestingly, the same scaling law is preserved using our voting rule, which effectively can be seen as a particular data sample strategy.

A similar analysis was conducted looking at the dependence on the system size $N_e$ (figure 4). Surprisingly, fixing $R$, the committee size decreases with the system size. For example, for the parameters used in figure 4, a representativeness of 0.9 is obtained with a committee of 29 members for a community of 1000 individuals, and with just 15 representatives for $N_e = 30\,000$. In particular, fixing $R = 0.9$, $F_{\text{rep}}$ decreases approximately with the inverse of $N_e$ using a Barabasi–Albert network, and even faster for an Erdős–Rényi network.

In figure 5, we can see that the representativeness is not strongly dependent on the connectivity of the network, and for $\langle k \rangle > 40$ the curves present very similar behaviours. The heterogeneity in the degree distribution of the network seems to have a relative small impact on the results too, as it can be appreciated by comparing the results of the Erdős–Rényi network with the Barabasi–Albert one. Finally, the small world property of the Watts–Strogatz model does not influence our results.

These results are consistent with the inspection of the relation between the individual’s connectivity and the number of votes they received. As can be appreciated in figure 6, even if it is necessary a reasonable connectivity (higher than the mean value) to obtain votes and a higher connectivity increases the probability to obtain more votes; for Erdős–Rényi networks this effect is weak and not impactful. For the case of the Barabasi–Albert network, this effect is more relevant and probably it is responsible for the influence that this topology has on the behaviour of the representativeness and normalized committee size (figure 5), determining a slightly weaker performance. In fact, it moderately decreases
3.2. Comparison with other voting rules

In the following, we explore if our model, at least theoretically, selects committees who present a representativeness comparable with other traditional methods of body selection. Obviously, representativeness is compared among committees of the same size.

A first possible comparison is with a model of a traditional majority voting for the selection of representatives in a closed list of previously selected candidates. This is probably the most common practice in selecting committees. An example can be found in the election of legislative assembly with the system of multi-member districts. In our simulations, a list of $N_c$ candidates is randomly selected among the community and each individual votes for the candidate who presents the higher overlap with its opinion vector. Decisions are taken with the same weighted voting rule. This modelling approach mimics a voter who presents a perfect knowledge of the candidates, and it assumes that he makes a rational
Figure 6. (a,c) Mean connectivity \( \langle k_V \rangle \) of an individual which received a number of votes \( V \). (b,d) Mean number of votes \( \langle V_k \rangle \) collected by individuals with a given connectivity \( k \). Data are obtained using 1000 different realizations in an Erdős–Rényi network (a,b) and a Barabasi–Albert one (c,d), with \( \langle k \rangle = 40 \), \( N_i = 40 \) and \( N_e = 10000 \).

decision to maximize his representation. Also for this voting rule, representativeness is computed by comparing the decisions taken by the committee, obtained with a weighted majority voting process, with the results of the direct popular vote. As can be appreciated in figure 7, our model is by far more efficient, reducing the size of the committees in more than a half.

Finally, we compare our method to an idealized perfect voting rule (PVR). This rule represents a situation of rational individuals that have a perfect knowledge of all the other individuals, which means that they perfectly know the opinion of all the other individuals. Moreover, they are globally networked, which means that they have a direct access to all other individuals, allowing their acts checked. In this situation, a voter indicates an individual which presents the higher overlap with his opinion vector. The selected committee is composed of first \( F_{\text{rep}} \cdot N_e \) individuals which poll more. Also in this case, the committee decisions are taken by means of a weighted majority vote. This voting rule, although unrealistic, is still useful, at least, in two respects. First, very small communities can exhibit similar characteristics. Second, the model is a useful yardstick for evaluating the levels of representativeness of other more realistic models.

Note that these voting rules can be considered, from a more abstract point of view, as an optimization problem: to find the vectors with greater overlap in an ensemble of given vectors, with no other constraints. By contrast, our original voting rule corresponds to the same optimization problem constrained by the fact that the overlap is inspected only locally, on a small subset of the ensemble of given vectors, because of the presence of the confidence circles.

In figure 7, we compare the representativeness of the PVR with our networked rule for different committee sizes. It is quite impressive that the representativeness of our voting rule is comparable with the PVR. Actually, when the number of opinions is relative small, the voting rule here presented performs even slightly better than the perfect rule. This effect disappears when the number of opinions is increased, and in this situation our networked rule is practically indistinguishable from the PVR.
Figure 7. Representativity versus the fraction of representatives for the proposed networked voting rule (NVR), the traditional majority voting (TMV) and a perfect voting rule (PVR). Parameters: $N_e = 10000$, $\langle k \rangle = 40$, $N_i = 40$ (a) and $N_i = 160$ (b). For the TMV, we consider $N_i = 100$. See the main text for detailed explanations of the voting rules.

4. Discussion

We introduced a general framework for studying the problem of selecting a committee of representatives to make decisions on behalf of a larger community. In our approach, we do not probabilistically study some properties of a particular voting rule, but we realistically implement it taking into account the effects of the individuals’ opinions and of their social relations. Moreover, we model the behaviour of the selected committee which is called to make choices about different issues. In this way, we are able, not simply to compare their choices with some abstract truth, but to clearly quantify the relation between representativeness and committee size.

Based on this scheme, it is possible to study the properties of a new networked voting rule, introduced with the aim of obtaining high representativeness together with strong accountability of the elected. This rule also presents the interesting feature that the candidates are not fixed in advance, but they emerge as a self-organized process.

The results of our simulations suggest that this rule is able to obtain a high representativeness for relatively small committees. In fact, these outputs are comparable with an ideal PVR, and they perform clearly better than a classical voting rule based on a closed list of candidates. Moreover, for fixed representativeness, the normalized committee size approximately scales with the inverse of the community size, allowing the scalability of this approach to very large populations. Finally, we were able to characterize the relation between the committee size and the representatives by means of a general inverse square root law.

These findings are robust and they are not strongly influenced by general properties of the network used to describe the individuals’ interactions. It seems that only the heterogeneity can have some role in modifying the relation between representativeness and committee size. In fact, the presence of a few individuals with very high connectivity can negatively influence the selection process of the committee. This fact is consistent with the risks related to a dominant position of single individuals, which may lead to a general sub-representation of the general opinion of the community.

The introduction of this networked voting rule, based on a decentralized large scale platform, can be considered as an interesting tool for implementing an hyper-representative mechanism of committee selection based on a distributed social mechanism where the use of a block chain encryption mechanism could guarantee the security of the voting process. This algorithm-based approach facilitates the participation of the entire population both as electors and as representatives, dismissing the importance of traditional authorities. Our analysis highlights the feasibility and potential of this election rule from the point of view of an organizational theory; specific and thorny political theory considerations and risks are not addressed.

Data accessibility. The code used for this research is available at https://www.researchgate.net/publication/321804236_AxelDemX3.

Authors’ contributions. A.R.H. is responsible for the conceptual idea behind the paper; all the authors participated in the design of the model; A.R.H. and C.G.-L. have written the code and ran the simulations; all authors participated in the
analysis and discussions of the results; E.B. was the main person responsible in the writing process, but all authors commented on the manuscript and gave their final approval for publication.

Competing interests. We have no competing interests.

Funding. C.G.-L. and Y.M. acknowledge support from the Government of Aragón, Spain through a grant to the group FENOL, by MINECO and FEDER funds (grant no. FIS2014-55867-P) and by the European Commission FET-Open Project Ibsen (grant no. 662725).

Acknowledgements. A.R.H. thanks COSNET Laboratory at the Institute BIFI for their partial support and hospitality during the realization of most of this work.

References

1. Condorcet MJ. 1785 Essai sur l’application de l’analyse à la probabilité des décisions rendues à la pluralité des voix. Paris, France: de l’Imprimerie Royale. (https://doi.org/10.3931/e-rara-3791)

2. Kelly J. 1991 Social choice bibliography. Soc. Choice Welfare 8, 97–169. (http://www.jstor.org/stable/41105976)

3. Chamberlin J, Courant P. 1983 Representative deliberations and representative decisions: proportional representation and the Borda rule. Am. Polit. Sci. Rev. 77, 718–733. (doi:10.2307/1957270)

4. Gallagher M. 1991 Proportionality, disproportionality and electoral systems. Elect. Stud. 10, 31–51. (doi:10.1016/S0261-3794(91)80004-C)

5. Gallagher M, Mitchell P (eds). 2005 Project Ibsen (grant no. 662725).

6. Naamani-Dery L, Kalech M, Rokach L, Shapira B. 2017 Tokeeportochangefirstpastthepost

7. Blais A. 2008

8. Bormann NC, Golder M. 2013 Democraticelectoral choiceofelectoralssystemsinadvanced democracies. Am. Polit. Sci. Rev. 93, 609–624. (doi:10.2307/25853577)

9. Bormann NC, Faliszewski P, Slinko A. 2015 Achieving proportionalrepresentationandtheBorda rule. Am. Polit. Sci. Rev., 99(4), 609–624. (doi:10.1016/j.socho.2013.01.005)

10. Clifford P, Sudbury S. 1973 Amodelforspatial consensus topersonalizeddecisionmaking. In Proc. of IJCAI-2011 16–22 July 2011, Barcelona, Spain, pp. 1141–1147. Buenos Aires, Argentina: AAAI Press. See http://dl.acm.org/citation.cfm?id=2832249. 2832407.

11. Dixit AK, Weibull JW. 2007 Political polarization. Proc. Natl. Acad. Sci. USA 104, 7331–7336. (doi:10.1073/pnas.0702770104)

12. Erdoes P, Rényi A. 1960 On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. 5, 17–60.

13. Barabási AL, Albert R. 1999 Evolution of scaling in random networks. Science 286, 509–512. (doi:10.1126/science.286.5439.509)

14. Watts DJ, Strogatz SH. 1998 Collective dynamics of ‘small-world’ networks. Nature 393, 440–442. (doi:10.1038/30918)