Viscous Democracy for Social Networks

Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna

ON April 23, 2009, the social network site Facebook announced the preliminary results of a voting in which users were asked about a change in the terms of use of the network.[1] The vote came as the result of a petition by thousands of users, who criticized the social networking site for claiming too many rights over the user-generated content. In an attempt to legitimate this change, Facebook decided to let the users vote, and stated that if 30% of the roughly 200 millions "active" users had voted the decision would have been binding for them.[2] The outcome was that the new rules were preferred by 74.4% of the voters. And while only 600,000 users voted (1/100th of the prefixed *quorum*) the change in the terms of use went ahead nevertheless.[3] The global privacy watchdog, Privacy International, called the Facebook vote a "massive confidence trick".[4] The low voting turnout was largely foreseeable. Obviously, only a small fraction of Facebook users have the time, patience and dedication, and take the service seriously enough to actively participate in its governance.

Whether done to decide on a *motion* (e.g., to pick the best among a series of alternatives), or to elect some *representatives* (e.g., to constitute a senate), voting can be considered as a *collaborative decision making process*, whose final goal is to come to a determination reflecting as much as possible the opinion of citizens and thus leading to a decision that is good for the community as a whole.

In this perspective, the failure of the Facebook voting experiment is easily explained by the kind of *voting system* [2], [14] that was adopted. Attempting this kind of *direct democracy* voting in large online communities turns out not to be a good idea: when public decisions reach a certain level of complexity, it is unrealistic to assume that every participant might become engaged and informed enough to contribute to the decision making [8], [11].

This is confirmed by several authors [5], [16] which have observed that the degree of commitment of different participants in online communities and collaboration systems varies greatly. According to Nielsen's "90-9-1 Rule for Social Design" [16]:

> In most online communities, 90% of users are lurkers who never contribute, 9% of users contribute a little, and 1% of users account for almost all the action.

This may sound unfair, but it turns out to be central to the way open collaboration networks work. For instance, in the case of Wikipedia, Shirky [19] states that:

> Fewer than two percent of Wikipedia users ever contribute, yet that is enough to create profound value for millions of users ... among those contributors, no effort is made to even out their contribution. The spontaneous division of labor driving Wikipedia wouldn't be possible if there were concern for reducing inequality. On the contrary, most large social experiments are engines for harnessing inequality rather than limiting it.

Indeed, in a well-designed collaborative system even participants that show up only once can contribute positively to the achievement of the goals of the community. However, those "drive-by" participants do not (and should not) have unreasonable expectations about their weight in the decisions of the group. For instance in the case of the popular open source distribution Ubuntu, Leadbeater [15] points out that:

> Decision-making is very open: anyone can see what is decided and how; anyone can make suggestions about what should be done. But the way decisions are made is rarely democratic.

In a community in which there are a few core members with long-term commitments to the project, and many other members joining and leaving the project rapidly, egalitarian democracy is neither expected nor appropriate. Thus, the decision-making mechanism is often meritocratic. These considerations suggest that different forms of voting system should be considered for the very peculiar communities which are electronically-mediated social networks.

DIRECT, REPRESENTATIVE, AND LIQUID DEMOCRACY

Direct democracy is based on the idea of ensuring maximum equality and fairness by making all citizens vote directly for the different motions. As discussed before, direct democracy works better in practice for small cohesive groups, but when the decisions reach a high level of complexity, and the community reaches a large size, it becomes impractical for every citizen to be fully informed on every issue. Furthermore, direct democracy requires deliberation to work effectively, and deliberation is relatively more difficult to achieve through electronic communications than by direct face-to-face contact. Electronic communications reduce the set of modalities by which group members may communicate, affecting their performance, particularly when the group members are new to the specific technology being used [12].

Representative democracy, instead, involves a relatively small number of representatives who are elected by the citizens to take decisions on their behalf, about many different matters, during a relatively long period of time. Beyond the issue

---

[1] http://www.facebook.com/fbsitegovernance
[2] An account is defined to be active if it had some activity in the past thirty days.
[3] http://www.cio.com/article/490775/After_Vote_Facebook_Gets_New_Governing_Rules
[4] http://www.privacyinternational.org/article.shtml?cmd[347]=x-347-564312

of which representation structure is the most appropriate for a given context, representative democracy is only weakly democratic: although citizens may participate in elections, they cannot really choose their representatives in the strict sense, instead they have to choose among a restricted set of candidates whose views and values are often radically different from their own, and who tend to make fundamentally opposed choices when faced with controversial social decisions. As a result, voters' apathy is more common than political interest. Borrowing James Green-Armytage's words [11]:

> In traditional representation systems, voters' positions on hundreds of social issues must be reduced to choices between candidates or parties, resulting in massive information loss.

Voters' apathy together with the concentration of the power in the hands of a small political elite create fertile ground for corruption, entrenchment, conflicts of interest, etc., which may result in bad government [18].

Driven by the idea of meritocracy, sometimes online-communities have implemented decision systems which are midway in between the universal voting of direct democracy, and the model of representative democracy. In these cases, the right to vote is given to a subset of the community, not selected by the citizens through another voting, but chosen on the basis of commitment. For instance, while the on-line community of editors of Wikipedia discourages explicitly the use of voting for resolving editorial disputes[5], for changes affecting the whole Wikipedia, voting may take place – an example being the April 2009 vote[6] on the issue of adopting a Creative Commons license. In that case, not all the Wikipedia editors were allowed to vote, but only those that had contributed 25 edits or more by a certain date.

While this kind of solution has the merit of implementing a meritocratic decision system, it does not consider the peculiar nature of large online social networks, their structure, their properties[7] and their high conductance for viral phenomena. Instead of selecting who has the right to vote by a *rigid* threshold on commitment and activity (which do not necessarily imply that someone is trustable), it seems more appropriate to adopt a *fluid* system based on people's trust.

The idea of fluidity leads us to consider another form of voting system, namely *delegative democracy*[8] or *liquid democracy*[9], which is based on *transitive proxy voting* [8], [11], [23]. Under this system, a user can decide either to express directly her opinion on an issue, or to delegate her vote on that issue (or about multiple related issues) to a *proxy*, that is, another citizen that she trusts. If the proxy votes directly

---

[5] http://en.wikipedia.org/wiki/Wikipedia:VOTE

[6] http://meta.wikimedia.org/wiki/Licensing_update/Result

[7] Various properties have been discovered characterizing large social networks, such as the the fact that the distribution of nodes' degrees (i.e., the number of edges incident in a node) has a *heavy tail*, the diameter (i.e., the maximum possible distance between two nodes measured as length of the *shortest path* connecting them) is small [1], they exhibit a *small world* structure [21] and a peculiar community organization [9]. See also [22].

[8] The idea of delegate voting is not so young: an early proposal, dated 1884, was due to Lewis Carroll [6].

[9] http://democracialiquida.es/

on the issue, the weight of all delegated votes she received are added to her vote. Proxy delegation may be *transitive*: one's vote can be further delegated to her proxy's proxy.

Proxy-voting systems encourage people to co-operate to build direct, permanent political and social relationships with each other and with their individual supporters, forming a web of trust. People can achieve political influence proportional to their level of public support, that in social networks is usually related to their level of connectedness.

FROM LIQUID TO VISCOUS DEMOCRACY

Liquid democracy is based on local, personal acquaintance, trust and social relationships, and on the principle of transitive delegation: thus the collaborative decision making process in liquid democracy is a social cascade process, that seems perfectly suited for online social networks. Therefore we take it as a starting point for developing our proposal.

In the following we present a voting scheme that can be used for both deciding upon a concrete matter (e.g, deciding between two alternative motions, by taking the weighted count of the votes received by each motion) and to select a committee of representatives to deal with a particular group of issues (and during a particular period of time, relatively shorter than that of representative democracy). For the time being, the reader can assume we deal with the first case (deciding about an issue); we will later discuss the case of electing a committee.

Like in liquid democracy, the key aspect of the voting system we propose is that votes can be delegated transitively along the existing links of the social network. That is, any member of the network can choose a proxy among her contacts. Citizens can obviously also choose not to delegate their vote and express directly their opinion on the matter of the voting.

Besides the obvious organizational advantages, the constraint that votes can only be assigned to a direct connection has a twofold rationale. On one hand, voters can base their decision on a direct personal knowledge of the person they vote for, making direct propaganda essentially useless and thus decoupling popularity from credibility; on the other hand, the fact that mandates are attributed through a chain of direct connections should ensure a stronger sense of responsibility.

However, we cannot ignore the fact that personal ties in online networks are not as strong or direct as in real-world communities. People's social connections are a mixture of *strong ties* (family, close friends) and *weak ties* (distant friends, acquaintances) [10], and electronically-mediated networks allow people to maintain many more weak ties than before [7]. This means that the number of connections they have is larger than what one could consider an actual "friendship" network, and that members' trust in their connections is weaker on average. For this reason, it seems appropriate to introduce some *reluctance* in the delegation process that reduces the amount of transitivity, as we are now going to explain in detail.

The ballot and the tally

To describe a voting system we need to specify how the voters express their preference (sometimes called the *ballot*),
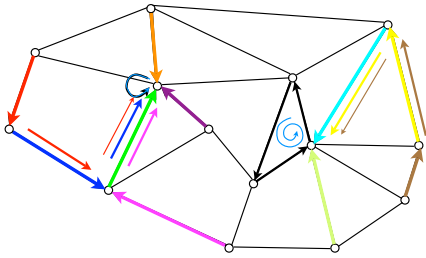
Fig. 1. Example delegation graph. Thin gray lines indicate the underlying social network. Colored arrows represent transitive delegation of their originating nodes, with thickness indicating the amount of weight actually transferred

and the algorithm that determines the final outcome (the *tally*). Note that there are also technical issues with how the voting is concretely carried on, such as establishing the identity of participants[10] – this is an important topic on its own, that can be solved through various technical means, but that we are not going to address here.

The ballot. Various possible choices exist for defining the ballot. *"One-vote"* voting systems are those in which a voter picks exactly one candidate; in our case, one contact. In a *"ranked"* voting system, each voter would rank her contacts in order of preference, and in a *"rated"* voting system, voters would give a score to each contact.

In the following we consider the simpler *one-vote* kind of ballot, where participants choose to delegate their decision to exactly one of their contacts, or to vote for themselves, which corresponds to not delegating the vote further.

This ballot can be interpreted as the creation of a *delegation graph*. A delegation graph is a directed graph built over the undirected underlying friendship social graph: it can contain cycles and it can have self-loops representing the choice of some electors not to delegate their vote and instead to express directly their opinion on the matter of the voting. Figure 1 illustrates an example of delegation graph induced over a social network.

The tally. Our system considers that each person in the network receives a certain amount of score (weight); the score will then be used to decide among the possible alternative motions (or to elect a committee), but the actual way scores are used is not part of the voting system we discuss.[11]

There are many ways of computing this score from a delegation graph, a trivial one being the sum of all votes received. Here we propose a more complex tally, namely *transitive proxy voting with exponential damping*. This is similar to standard proxy voting of liquid democracy, but with a *damping factor* that introduces some reluctance in the way delegated votes are transferred. This reluctance, controlled by

a parameter α, corresponds intuitively to the idea that, in an electronically mediated social network, typically you cannot fully trust your connections, and you want to refrain from giving them all of your delegation. But more importantly, you do not know how far your liquid vote can go on the network hop-by-hop: even if you fully trust your proxy, can you transitively fully trust your proxy's proxies?

Reluctance makes the vote less liquid, reducing its strength with each delegation step, and thus limiting the distance it can travel. Reluctance makes the vote *viscous*. We might call this form of proxy voting a *viscous democracy* because of the way trust (and consequently a vote's weight) decays with the distance.

VISCOUS DEMOCRACY AND SPECTRAL RANKING

We are now ready to introduce our proposed voting system. The computation we suggest is known to sociologists as *Katz's index* [13]: every vote transfers by transitivity to distances larger than one, but with an attenuation factor. The delegation graph has out-degree one (because a one-vote ballot is used), which makes our case much easier to analyze. In fact, the score of node i is simply proportional to:

$$\sum_{p \in \mathrm{Path}(-,i)} \alpha^{|p|} ,$$

where $\mathrm{Path}(-,i)$ is the set of all delegation paths ending at node i and $|p|$ is the length of one such path.[12] We remark that computing Katz's index on the delegation graph is completely different from computing Katz's index on the social network (which would be its standard application).

Techniques similar to Katz's index, the so-called *spectral ranking* methods, have been used by psychologists, sociologists, management theorists, etc., at least for the past 50 years, to estimate authority, power, influence and centrality. The currently most popular incarnation of this idea is probably Google's PageRank [17] (see the sidebar "**A brief history of Spectral Ranking**" for details). Our proposal adds a new flavor to an old ingredient, showing how decades-old techniques can be applied to voting in online social networks.[13]

THE PARAMETER α

There is a single parameter $\alpha \in]0, 1[$ controlling the voting process, which can be understood as the *delegation factor*, the amount of its own power that a person can delegate to another person; in other words, $1 - \alpha$ is the amount of viscosity in the system.

If the delegation factor is small (close to 0), mandates become undelegable. This means that if a person receives enough delegations (votes) from other nodes, she alone can not make a third party "more powerful" than herself even if she votes for him/her. In the limit, when the delegation factor

---

[10]A quite popular anecdote is that about "Hank the Angry Drunken Dwarf" winning the 1998 People magazine online poll to determine the most beautiful people in the world: `http://en.wikipedia.org/wiki/Hank_the_Angry_Drunken_Dwarf`

[11]For example, when deciding about a motion one may take only the voters that decided not to delegate their vote, and make them choose about the possible alternatives, weighting their choices proportionally to the scores obtained. If everybody delegated their votes, one can take people belonging to *cycles* instead, as they retain part of their voting power, albeit in a weaker sense. An application of our voting system to the case of elections will be discussed later in the paper.

[12]Given the possible presence of cycles in the delegation graph, some of these paths may be arbitrarily long; however for $|\alpha| < 1$ the above sum converges.

[13]Also Yamakawa *et al*. [23] use (a weighted form of) Katz's index, but on a matrix representing voters and motions. Each voter can give a positively weighted vote to one or more voters and motions, and the sum of the outgoing weights is forced to be one. There is however no underlying social network in their setting.

tends to zero, only direct votes count, and the resulting process can be described as a simple majority voting.

If the delegation factor is large (close to 1), most nodes that delegate their mandate to someone else will not have large scores. In the limit, when the delegation factor tends to one, the system becomes liquid democracy and the winners are chosen simply by the size of the sub-tree to which they belong (i.e., the number of people that voted for them, directly or indirectly).

In Figure 2 we depict an example showing how the delegation factor can determine the degree of viscosity, hence the scores: when $\alpha = 0.2$ (left), the system is more viscous and node 5 (which has many more *direct* supporters than the others) has the largest score; when $\alpha = 0.9$ (right), the higher degree of transitivity makes 7 the node with the largest score. Also observe that in this case node 8 is slightly stronger than node 9 (contrarily to what happens for smaller values of α), because although 8 has fewer direct supporters she receives part of the influence of node 7.

Finally, there are at least two more properties of this system that are worth mentioning (a more complete list of features, as well as a detailed comparison with PageRank, can be found
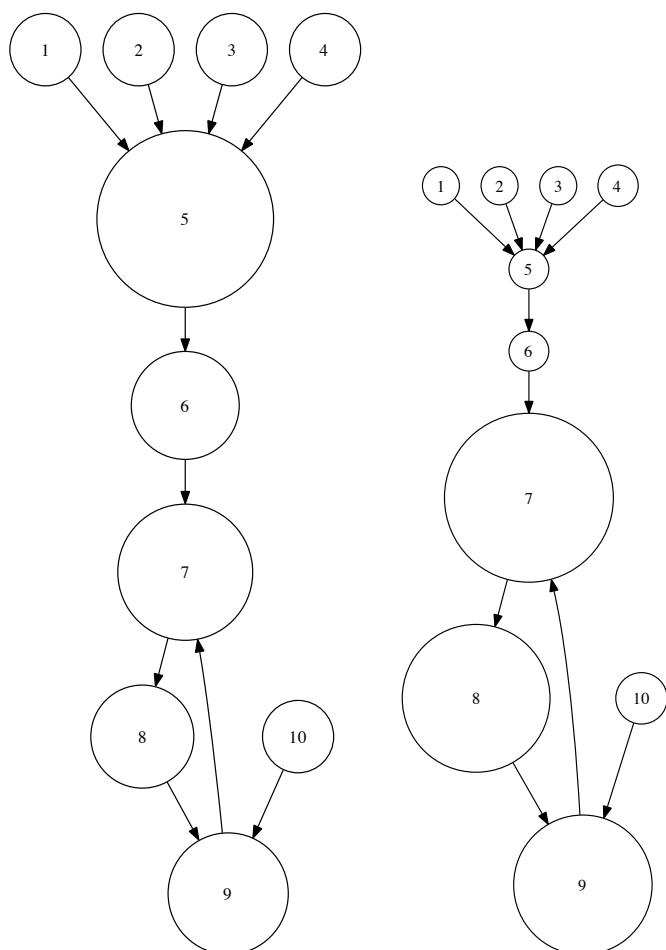

Fig. 2. Scores computed on the same delegation graph with a low delegation factor ($\alpha = 0.2$) and a high delegation factor ($\alpha = 0.9$). The size of the nodes is proportional to their score.

in [3]): (i) given the right delegation graph, everyone could be a winner (provided that the social network is connected), so there are no a-priori losers in the social network—while of course some nodes have a better chance than others—and (ii) there is continuity in the decision with respect to α: if one node can obtain a larger or smaller score than another with appropriate choices of α, there is always another choice that makes the two nodes have the same score.

IMPORTANCE OF THE SOCIAL NETWORK AND ABSTENTIONISM

There are two interesting questions raised by this system. The first is to what extent the structure of the social network determines the outcome of the election. The second is how to deal with missing votes. It turns out that the answers to these questions are closely related: indeed, when an individual abstains from voting we can use the local structure of the network around that individual to have hints about how her vote would have influenced the final outcome.

The results of an election in *viscous democracy* depend on the delegation factor α and on the delegation graph, which is in turn constrained by the underlying social network, because each voter can only vote for one of her neighbors. Above, we have explained that anybody could in theory win the election, given the right delegation graph. This result is optimistic, though, in the sense that in practice the likelihood that a node on the fringe of the network will win an election may be small.

The problem of missing votes can be dealt with naturally within our system considering all the possible what-if scenarios; if a member with k contacts does not vote, her vote may have been cast in k different ways, and each would have produced a certain score of the nodes: in absence of more information, we can consider these scores equally probable, and take their average as the results of the vote. In other words, we treat non-voters as if they are equally likely to trust any of their contacts, hence making the final outcome dependent on the votes that have been cast and on the local neighborhood structure around non-voters for the votes that have not been cast.

One way to try to understand what is going on here is by running an election where nobody expresses a preference (i.e., an election with 100% of abstention). This amounts to computing the expected score of each node considering only the social network.

Although this is a measure of centrality of the nodes (we called it *voting centrality* in [3]), it does not seem to be much correlated to other simple centrality measures (e.g., the number of contacts). This score is also not the standard PageRank computed on the social network graph: this is because there is a subtle, but important, difference between a node being equally likely to trust any of its contacts and a node spreading its trust equally among its contacts, like it would in the standard PageRank computation. A deep understanding of what this measure actually means is still missing, but the bottom line is that the social network *influences* but does not *determine* the outcome of all possible elections using viscous democracy—for example, as we said above, everybody can become a winner provided that the network is connected.

A brief history of Spectral Ranking: from the forties to PageRank

The renowned PageRank index, which is at the basis of the initial success of Google's ranking mechanism, has been re-discovered over and over for 50 years.

Indeed, spectral techniques for computing "best" entities when some relationship between them are known date at least to the late forties, when John R. Seeley (*Canadian Journal of Psychology*, 1949) publishes his study on ranking children given a matrix that contains information about whether each child likes another child. Seeley argues that the rank of a child should be given by the sum of the ranks of the children who like him. Seeley imposes this requirement through a linear system: in modern terms, he computes the left dominant eigenvector of the normalized matrix to establish which children are more popular.

Just a few years later, in his unpublished thesis (*The algebraic foundations of ranking theory*, 1952) T.H. Wei discusses how to rank football teams given a matrix representing how much each team is better than another (e.g., 1 for a win, 1/2 for a tie, 0 for a loss). He argues that given a starting equal score to every team, we can get a more precise score by adding for each team the scores of the team it defeated, plus half the score of the team with which it tied. Wei shows that iterating this process he obtains the *right* dominant eigenvector, and he uses it to identify the best teams.

At this point *spectral ranking*—using eigenvectors to compute ranks—was an established idea.

In the fifties Leo Katz introduces its index (*Psychometrika*, 1953). Katz starts from the consideration that given a zero/one matrix expressing whether each person of a group "likes" or "votes for" another (implying endorsement) we should estimate the importance of a person not only from the number of its voters, but also from the number of its voters' voters, and so on *ad infinitum*. To obtain a finite value, Katz suggests an *attenuation factor* α that is used to reduce the weight of longer and longer voting paths. It can be shown [4] that when the attenuation factor of Katz's index approaches the reciprocal of the dominant eigenvalue of the matrix we get back (as a limit) the standard spectral ranking of Seeley and Wei, thus suggesting the name *damped spectral ranking* for the former.

PageRank combines the dominant eigenvector ideas of Seeley and Wei with Katz's approach (damping). Again, the motivation is different, and related to the *random surfer model*—PageRank models the behavior of a surfer that moves randomly through links and, with probability α, jumps to a random node. Seeley's ranking is PageRank without jumps to random nodes, whereas Katz's index is PageRank without the normalization (divide by the number of linked nodes) usually applied to each matrix row.

For more mathematical details and further historical remarks, the reader can consult [20].

ELECTING A COMMITTEE: PROPORTIONALITY AND NON-MONOTONICITY

In this section, we discuss briefly the application of viscous democracy to elections: i.e., to select a committee of representatives to deal with a particular group of issues. In this case, self-loops in the delegation graph indicate citizens that accept to be possibly elected in the committee: in other words, nodes with a self-loop indicate their willingness to be considered as candidates.

When a committee having s seats must be selected, we can simply choose the s top scoring candidates. However, there is an opportunity of selecting a committee that represents the diversity of users, by ensuring *proportionality*. The criterion of proportionality states that, ideally, each political alliance should have a share of the seats proportional to its share of the votes [14].

The concept of *party* or *alliance* can be mapped easily into our voting system. In absence of specific alliances declared beforehand, a voting system for social networks may interpret the connected components of the delegation graph as alliances, as they represent communities of like-minded people who delegate to other members of their community but not to aliens. This way of selecting communities is much more fine-grained than simply choosing the connected components of the underlying social network: indeed, casting a vote implies an expression of will that singles out a special relation among a set that in practice contains many weak ties [7].

This allows for proportionality understood as picking, for each connected component of the delegation graph, the top-k scoring nodes, in which k is proportional to the size of the connected component. For example, suppose that we have to assign s seats and that we have c communities with $n_1, \ldots, n_c$ members, respectively: we can assign to the i-th community $k_i = n_i \cdot c/(n_1 + \ldots + n_c)$ seats, choosing the k_i top scoring nodes within that community. Like in all voting systems based on proportionality, however, some attention must be paid to how fractional seats should be assigned (in the above formula, in fact, k_i may not be an integer). There are many known (and incompatible) solutions to this problem, the most widely used (adopted, e.g., for electing the European Parliament) being the so-called *D'Hondt rule* [14]: this rule assigns the available seats one at a time, giving it to the community with largest ratio $n_i/(s_i + 1)$ where s_i is the number of seats assigned to the i-th community so far.

If our system is used for proportional voting, we have (like in many other multiple-winner voting procedures) no guarantee of *monotonicity*: this means, counter-intuitively, that not voting for a node y may increase its chances of being elected. Consider, for example, the situation depicted in Figure 3, and assume that we are going to assign 3 seats using the D'Hondt rule.

If C votes for herself, we have 3 communities of sizes $\{12, 8, 5\}$ and C is not elected, because the leftmost community will receive two seats, and the one in the middle will

get the remaining one[14]. Instead, if C votes for B, we have 2 communities of sizes $\{12, 13\}$ and C will be elected as the merged community will receive two seats[15] assigned to its top-two scoring members, i.e., B and C. Thus, in this case it is better for C to vote for B instead of for herself, violating the monotonicity property.
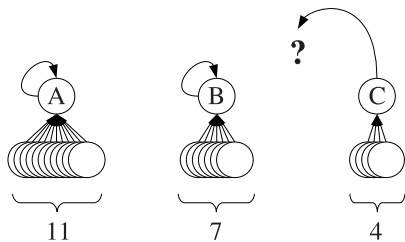

Fig. 3. If C wants to be elected as one of the 3 representatives of this network, under D'Hondt proportionality rules, should she vote for herself or for B?

Non-monotonicity makes the system vulnerable to tactical voting, albeit in practice no user has enough information about the structure of the network and about the other users' decisions to really implement it. In particular, if all votes are cast simultaneously and kept secret until the poll, tactical voting is made even more difficult although in principle not impossible. It is worth observing that anyway the whole voting system is essentially as reliable as the underlying social network: forms of collusion (e.g., accepting friendship from strangers if they promise to vote for me) are of course possible, as long as the social network contains no way to refrain one to accept strangers as contacts.

SIMULATED VOTING IN DBLP & Y! MEME

The effectiveness of a voting system in practice depends on qualitative factors, such as whether the voters are able to understand the voting system and accept it, whether the decisions reached by the community are in some sense correct and whether the members of the community agree to go along with such decisions. Although running a real voting experiment in a sufficiently large social network would be extremely difficult, we would like to present some simulations suggesting that our voting system produces results that are consistent with common sense and reflect the structural properties of the social networks on which they are run.

Simulated voting in a CS community

As a first example, we simulated an election in the computer science community, using the DBLP co-authorship network.[16] In this network each node represents a computer scientist and we interpret the co-authorship relation as a social tie,

---

[14]D'Hondt multipliers are (12,8,5). The first community wins one seat as it has the largest figure. Now multipliers are (6,8,5). The second community wins one seat. Now multipliers are (6,4,5). The first community wins one more seat. Now multipliers are (3,4,5), it would be the turn of the third community but there are no more seats to fill.

[15]D'Hondt multipliers are (12,13). The second community wins one seat. Now multipliers are (12,6.5). The first community wins one seat. Now multipliers are (6,6.5). The second community gets the last seat.

[16]http://www.informatik.uni-trier.de/~ley/db/

TABLE I
COMPUTER SCIENTISTS THAT OBTAINED THE TOP 10 SCORES IN TWO COMMUNITIES, BROADLY IDENTIFIED WITH "THEORY AND ALGORITHMS" (LEFT) AND "DATABASES AND DATA MANAGEMENT" (RIGHT). THE SCORE IS COMPUTED USING THE VOTING ALGORITHM DESCRIBED IN THE TEXT, WITH $\alpha = .85$; IN PARENTHESIS, THE NUMBER OF PUBLICATIONS IS SHOWN FOR COMPARISON.

Micha Sharir	.01594 (411)	Hector Garcia-Molina	.00683 (374)
Noga Alon	.00178 (402)	Jeffrey D. Ullman	.00097 (244)
Erik D. Demaine	.00050 (305)	Michael Stonebraker	.00034 (230)
Avi Wigderson	.00042 (243)	Randy H. Katz	.00028 (179)
Oded Goldreich	.00042 (269)	David Maier	.00027 (207)
Leonidas J. Guibas	.00038 (299)	David A. Patterson	.00021 (144)
David Eppstein	.00037 (304)	David J. DeWitt	.00020 (179)
Michael T. Goodrich	.00037 (228)	Rajeev Motwani	.00016 (182)
David Peleg	.00031 (279)	Raghu Ramakrishnan	.00015 (208)
Mikhail J. Atallah	.00031 (190)	David E. Culler	.00014 (140)

indicating that two scientists are connected if they have worked together on an article. We simulated a voting to elect the most representative author in each area, using the following criteria: every author x considers her co-authors in decreasing order of number of papers co-authored, and votes for the first one that has been more productive than x (i.e., that has written more papers than x); if no such co-author exists, x votes for herself.

The protocol we are using for this simulated voting contains admittedly more information than the underlying social network of co-authorship: this fact would anyway be true also of a real-world voting experiment (a voter would choose a proxy among her contacts on the basis of information that is of course not available or deducible from the network alone), but the reader should be warned that the results of the voting depend largely on the tallying rules adopted.

At the end of this simulation, we considered the top-scoring authors within each connected component of the delegation graph: such components are thought to approximate the different research communities, as explained above. In Table I you can find the top ten results of two communities, broadly identified as "Theory and algorithms" and "Databases and data management"[17]. Of course, because of the way the simulation is run, the most prolific authors tend to be favored, but the relation between number of publications and score is not at all trivial, as it can be seen even from the few examples shown in the table[18]; it is anyway clear that the results are in line with what the common sense would suggest.

Simulated voting in a Social Network

Yahoo! Meme[19] is a microblogging platform that allows for viral spread of information *memes*. Most of the memes currently being posted by users are interesting or funny photos or cartoons. The social network here is directed, and a link represents a *follower-followed* relation; of course, all our

---

[17]The data and code used for this experiment are available at http://law.dsi.unimi.it/dblp/.

[18]Kendall's τ correlation coefficient between number of publications and voting scores is 0.27 (0 means total uncorrelation and ± 1 means perfect positive/negative correlation).

[19]http://meme.yahoo.com/

machinery naturally extends to this case—followers can only vote for one of the users they follow.

We performed a simulated election on an early snapshot of Yahoo! Meme containing tens of thousands of users. First, the *influence* of a user u is computed considering all the memes that u have posted and summing the number of copies of those memes that have been re-posted (or "re-tweeted" in the microblogging jargon) by followers u and recursively, by followers-of-followers, and so on. Next, we assume that users vote for the person they follow most often, this is, for the person from whom they have re-posted the most memes, breaking ties arbitrarily. Finally, people vote for themselves if they are not following anyone more *influential* than themselves.

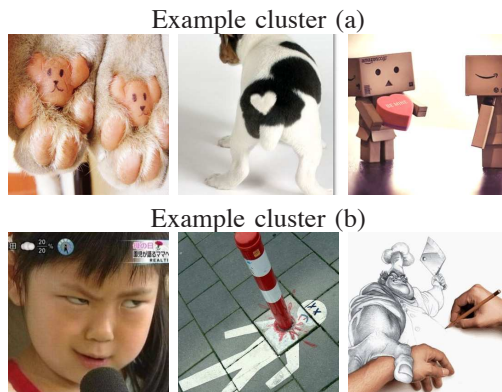Example cluster (a)

Example cluster (b)

Fig. 4. Representative images from two example clusters after a simulated voting in Yahoo! Meme.

The result of this election is a series of influential users "elected" in different communities. As with DBLP simulated election, even in this case we can observe that voting surfaces homogeneous communities.

First, we notice that communities tend to be homogeneous in terms of the countries users are located in. Second, inspecting visually the memes that had the larger number of re-posts in these communities, we notice they correspond to clearly coherent topics. For instance, in Figure 4 we display thumbnails of memes posted in two communities that tend to often post "cute" and "funny" images respectively.

Conclusions

Any voting system has trade-offs and viscous democracy is not exempt from those; we have discussed some of them above (for instance, its vulnerability to tactical voting and its dependence on the choice of the parameter α), as well as some possible extensions (e.g. give different weights to each delegation arc, allow users to vote for motions and to delegate fractionally [23]). But all these variants risk to make the system less simple and thus more difficult to understand and adopt.

People interact in social networks in many ways. Sorted by increasing complexity, they share information, cooperate, produce collaboratively, and take collective action [19]. We do not expect many online communities to take binding decisions collectively today, but they will do increasingly in the future. Novel interactive environments call for novel collective decision systems that will be developed over the next years.

References

[1] ALBERT, R., JEONG, H., AND BARABASI, A. L. Diameter of the world wide web. *Nature 401* (1999), 130–131.

[2] ARROW, K. J. *Social Choice and Individual Values*. John Wiley & Sons, 1951, 2nd ed., 1963.

[3] BOLDI, P., BONCHI, F., CASTILLO, C., AND VIGNA, S. Voting in social networks. In *Conference on Information and Knowledge Management (CIKM)* (2009), pp. 777–786.

[4] BOLDI, P., SANTINI, M., AND VIGNA, S. PageRank: Functional dependencies. *ACM Trans. Inf. Sys. 27*, 4 (2009).

[5] BUTLER, B., SPROULL, L., KIESLER, S., AND KRAUT, R. Community effort in online groups: Who does the work and why? In *Leadership at a distance*, S. Weisband, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 2007, pp. 171–194.

[6] CARROLL, L. *The Principles of Parliamentary Representation*. Harrison and Sons, 1884.

[7] DONATH, J., AND BOYD, D. Public displays of connection. *BT Technology Journal 22*, 4 (October 2004), 71–82.

[8] FORD, B. Delegative democracy. http://www.bford.info/deleg/deleg.pdf, May 2002.

[9] GIRVAN, M., AND NEWMAN, M. E. J. Community structure in social and biological network. *PNAS 99* (2002), 7821–7826.

[10] GRANOVETTER, M. S. The strength of weak ties. *The American Journal of Sociology 78*, 6 (1973), 1360–1380.

[11] GREEN-ARMYTAGE, J. An introduction to proxy-based democratic systems. http://fc.antioch.edu/~james_green-armytage/vm/proxy2007.htm, Written May 2007, posted online January 2009.

[12] HOLLINGSHEAD, A. B., MCGRATH, J. E., AND O'CONNOR, K. M. Group task performance and communication technology: a longitudinal study of computer-mediated versus face-to-face work groups. *Small group research 24*, 3 (1993), 307–333.

[13] KATZ, L. A new status index derived from sociometric analysis. *Psychometrika 18*, 1 (1953), 39–43.

[14] KATZ, R. S. *Democracy and Elections*. Oxford University Press, 1997.

[15] LEADBEATER, C. *We-Think: Mass innovation, not mass production*. Profile Books, July 2009.

[16] NIELSEN, J. Participation inequality: Encouraging more users to contribute. Jakob Nielsen's Alertbox, October 9, 2006. http://www.useit.com/alertbox/participation_inequality.html.

[17] PAGE, L., BRIN, S., MOTWANI, R., AND WINOGRAD, T. The PageRank citation ranking: Bringing order to the web. Tech. Rep. 66, Stanford University, 1999.

[18] PANI, M. Hold your nose and vote: Why do some democracies tolerate corruption? Working Paper WP/09/83, International Monetary Fund, April 2009.

[19] SHIRKY, C. *Here Comes Everybody: The Power of Organizing Without Organizations*. Penguin Press, February 2008.

[20] VIGNA, S. Spectral ranking, 2009. arXiv:0912.0238v2.

[21] WATTS, D. J., AND STROGATZ, S. H. Collective dynamics of 'small world' netwroks. *Nature 393* (1998), 440–442.

[22] WILLINGER, W., ALDERSON, D., AND DOYLE, J. C. Mathematics and the Internet: a source of enormous confusion and great potential. *Notices of the American Mathematical Society 56*, 5 (2009), 586–599.

[23] YAMAKAWA, H., YOSHIDA, M., AND TSUCHIYA, M. Toward delegated democracy: Vote by yourself, or trust your network. *International Journal of Human and Social Sciences 1*, 2 (2007).

Paolo Boldi obtained his PhD in Computer Science at the University of Milano, where he is currently Associate Professor. His research interests touched many different topics in theoretical and applied computer science, such as: domain theory, non-classical computability theory, distributed computability, anonymous networks, sense of direction, self-stabilizing systems. Recently, his works focused on problems related to the World-Wide Web, a field where his research has also produced software tools used by many people working in the same area.
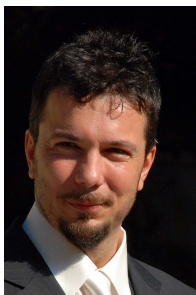
Francesco Bonchi received his PhD in Computer Science from University of Pisa, in 2003. He is currently a researcher at Yahoo! Research where his recent research interests include mining query-logs, social networks, and social media. He is co-editor of the book "Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques" Chapman & Hall/CRC Press. He has been organizing the Privacy, Security and Trust workshop at SIGKDD in 2007 and 2008. He is co-chair of ECML-PKDD 2010.

Carlos Castillo received his PhD in Computer Science from University of Chile, in 2004. He is currently a researcher at Yahoo! Research Barcelona. Before that, he was a postdoctoral fellow at Sapienza Universit di Roma and Universitat Pompeu Fabra in Barcelona. He has researched on Web crawling, Web characterization, and Web ranking; and is currently active in the areas of Web Usage and Link Mining. In the area of Adversarial Web IR, he has co-organized several challenges and workshops such as GraphLab and AIRWeb.

Sebastiano Vigna obtained his PhD in Computer Science at the Università degli Studi di Milano, where he is currently Associate Professor. His interests lie in the interaction between theory and practice. He has worked on highly theoretical topics such as computability on the reals, distributed computability, self-stabilization, and theoretical analysis of PageRank, but he is also (co)author of several widely used software tools ranging from high-performance libraries to a model-driven software generator, a search engine, a crawler, and a graph compression framework.