Approval-rating systems that never reward insincerity

Rob LeGrand and Ron K. Cytron

Abstract

Electronic communities are relying increasingly on the results of continuous polls to rate the effectiveness of their members and offerings. Sites such as eBay and Amazon solicit feedback about merchants and products, with prior feedback and results-to-date available to participants before they register their approval ratings. In such a setting, participants are understandably prone to exaggerate their approval or disapproval, so as to move the average rating in a favored direction.

We explore several protocols that solicit approval ratings and report a consensus outcome without rewarding insincerity. One such system is rationally optimal while still reporting an outcome based on the the usual notion of averaging. That system allows all participants to manipulate the outcome in turn. Although multiple equilibria exist for that system, they all report the same average approval rating as their outcome. We generalize our results to obtain a range of declared-strategy voting systems suitable for approval-rating polls.

1 Approval ratings and their aggregation

Approval ratings are one mechanism that communities can use to offer incentive and reward for good behavior or service. The prospect of feedback following a given interaction presumably increases the accountability of that interaction for all parties involved. Publication of approval ratings then enables appropriate consequences to follow from positive or negative experiences. In this paper we consider several forms of aggregation and we show that some methods can reward insincerity while others cannot. We next provide several examples of approval rating systems and formulate a general form of an approval rating poll.

1.1 Examples of approval rating polls

Subscribers and observers of media frequently learn of the results of *approval rating polls* that attempt to discern how strongly a participating electorate endorses a person or a position of interest. As an example, several web sites post varous forms of approval ratings for films and games. Specifically, Rotten Tomatoes [2] posts the results of two polls for each film:

- In effect, each review is turned into a 0 or 1 value, and the *Tomatometer* is the average of those values expressed as a percentage. Putative viewers might consult a film's Tomatometer value to determine whether they should see that film.
- Each critic can also rate a film's overall quality on a 1–10 scale. Rotten Tomatoes then publishes the average of all such ratings.

Finally, consider the electronic marketplace, in which participants are asked to rate the honesty and effectiveness of merchants and customers. Sites such as eBay poll their participants concerning how strongly they approve of the behavior of the marketplace members they encounter in transactions. Upon completion of a transaction, the involved parties are asked to rate each other. An aggregation of an indvidual's approval ratings is posted for public view, so that members can consider such information before engaging that individual in a transaction.

1.2 Formulation

We next define a general instance of an approval rating poll to facilitate presentation of our results.

- An electorate of n participants is polled. Based on the participants' response and the aggregation protocol at hand, the result of the poll will be published as a rational number in the interval $[0, 1]$.
- Each participant i has in mind a sincere preference rating r_i, where $0 \leq r_i \leq 1$, that can be construed as that participant's dictatorial preference. The tuple of all participants' sincere ratings is denoted by the vector $\vec{r}$. We further make the reasonable assumption that voter i's preferences are single-peaked and non-plateauing.[1]
- Finally, voter i participates in the poll by expressing a rating preference of v_i, which may or may not be the same as r_i. In fact, we are particularly interested in situations where $v_i \neq r_i$. For example, consider an eBay customer who undertakes a transaction with a highly approved merchant. If the customer becomes disgruntled with the merchant, then the customer's resulting rating of the merchant might be overly negative, precisely because of the merchant's otherwise high rating.

  The tuple of all expressed approval ratings is denoted by the vector $\vec{v}$.

This paper considers an approach that can account for, mitigate, or prevent the use of insincerity to increase a participant's effectiveness in an approval rating poll.

1.3 Aggregating approval ratings

The results of an approval rating poll are typically reported by an aggregation procedure that is disclosed *a priori*. In this section, we consider two popular aggregation schemes: average and median.

Average aggregation Here, the result of the approval rating poll is computed as the average of the participants' expressed approval ratings: $\bar{v} = \frac{\sum_{j=1}^{n} v_j}{n}$. While the Average aggregation function is sensitive to each voter's input, it has an important disadvantage: Voters can often gain by voting insincerely. For example, the 1983 film *Videodrome* has five critics' ratings on Metacritic [1]. If we assume that these critics rated the film sincerely (that each would prefer that the average rating of the film be his or her rating), we have $\vec{r} = [0.4, 0.7, 0.8, 0.8, 0.88]$. If these preferences are actually expressed sincerely in an Average aggregation context, then we have $\vec{v} = \vec{r}$ and the Average outcome is 0.716.

Consider voter 5, whose ideal outcome is $r_5 = 0.88$. That voter could achive a better outcome by *not* expressing the sincere preference $v_5 = 0.88$ and instead voting $v_5 = 1$. The resulting Average aggregation yields the outcome 0.74, which, being closer to 0.88, is preferred by voter 5 to 0.716.

Median aggregation (n odd) Another possible aggregation function computes a *median* of $\vec{v}$: $\tilde{v}$ is a value that satisfies $|\{i : \tilde{v} < v_i\}| \leq \frac{n}{2} \leq |\{i : \tilde{v} \leq v_i\}|$.[2] According to the *median*

[1] For the applications we describe, it is reasonable to assume that each voter i would prefer that the outcome be as near to the ideal r_i as possible. This single-peaked assumption makes possible the optimal strategy we describe in section 2.

[2] The above definition does not necessarily prescribe a unique outcome when n is even; we address this issue below.

voter theorem [4, 10], when n is odd, Median aggregation becomes the unique, Condorcet-compliant [13] rating system, yielding a result that is preferred by some majority of voters to every other outcome.

Unfortunately, Median aggregation can effectively ignore almost half of the voters—majority rule can mean majority tyranny. Given the tuple of votes $\vec{v} = [0, 0, 0, 1, 1]$, the 1-voters are effectively ignored when the median, 0, is chosen as the outcome. Majority tyranny could be quite undesirable for polls of this type, especially when the goal of aggregating ratings is to represent a satisfactory consensus for all voters. The Average outcome of the above tuple, 0.4, arguably provides such a much better consensus.

In contrast with Average aggregation, Median aggregation is nonmanipulable by insincere voters—at least when n is odd: a voter i can never improve the outcome from his or her point of view by voting $v_i \neq r_i$. (The treatment for an even number of voters and the proofs here and below are omitted for space; this material can be found in LeGrand [12, ch. 3].) Thus, Median aggregation does not reward insincerity for an odd number of participants.

Without losing nonmanipulability, the Median function can be generalized to give the outcome ${}^b\tilde{v}$ where $|\{i : {}^b\tilde{v} < v_i\}| \leq bn \leq |\{i : {}^b\tilde{v} \leq v_i\}|$ for any $0 \leq b \leq 1$ (in this notation, the b is intended as a parameter modifying the tilde symbol). If bn is an integer, there may be more than one $0 \leq \phi \leq 1$ that satisfies $|\{i : \phi < v_i\}| \leq bn \leq |\{i : \phi \leq v_i\}|$. In that case, define Φ as the set of all such ϕ. Then

$${}^b\tilde{v} \equiv \begin{cases} \min(\Phi) & \text{if} \quad b < \min(\Phi) \\ b & \text{if} \quad \min(\Phi) \leq b \leq \max(\Phi) \\ \max(\Phi) & \text{if} \quad \max(\Phi) < b \end{cases}$$

This order-statistic outcome equals $\max(\vec{v})$ when $b = 0$, the third quartile when $b = \frac{1}{4}$, the Median outcome when $b = \frac{1}{2}$, the first quartile when $b = \frac{3}{4}$ and $\min(\vec{v})$ when $b = 1$.

2 Rationally optimal strategy for Average aggregation

As shown in section 1, Average aggregation can reward insincerity. In this section, we develop a *rationally optimal* strategy: a computation by which a voter can achieve a result as close as possible to that voter's preferred outcome. As before, we assume an electorate in which n voters will express preferences. We begin by considering a rationally optimal ("best response") strategy from the perspective of a final, omniscient voter. We then consider the behavior of a system in which all voters use a rationally optimal strategy.

To facilitate exposition and analysis of our results, we begin by generalizing the scale on which preferences are expressed as follows. In an $[m, M]$-Average poll, voters are allowed to *express* preference ratings in the interval $[m, M]$, where $m \leq 0$ and $1 \leq M$. We continue to assume that *sincere* preference ratings are in the interval $[0, 1]$; the expanded range is therefore intended to allow voters more room to manipulate the outcome. We also assume that preferences are aggregated by computing the average of the voters' expressed preferences.

2.1 Strategy for a final, omniscient voter

Consider a $(-\infty, +\infty)$-Average poll in which voter v_n is the last voter to express an approval rating, and in which all other voters vote their sincere preference ratings: $(\forall i \neq n)\ v_i = r_i$. If voter n can see the expressed approval ratings of all voters, then the ideal outcome for voter n ($\bar{v} = r_n$) can be realized by voting $v_n = r_n n - \sum_{j \neq n} r_j$.

More generally, in an $[m, M]$-Average poll, voter n should express v_n to move the outcome as close to r_n as possible:

$$v_n = \min\left(\max\left(r_n n - \sum_{j\neq n} r_j, m\right), M\right) \tag{1}$$

The above is the *rationally optimal strategy* for voter n in an $[m, M]$-Average approval rating poll.

As an example, consider the $[0,1]$-Average system with sincere preferences from the *Videodrome* example above: $\vec{r} = [0.4, 0.7, 0.8, 0.8, 0.88]$. After all other voters express their sincere preferences, v_5's rationally optimal preference rating is given by

$$\begin{aligned} v_5 &= \min\left(\max\left(r_5 n - \sum_{j\neq 5} r_j, 0\right), 1\right) \\ &= \min\left(\max\left(0.88 \cdot 5 - (0.4 + 0.7 + 0.8 + 0.8), 0\right), 1\right) \ = \ 1 \end{aligned} \tag{2}$$

achieving an outcome $\bar{v}$ of 0.74. No other choice for v_5 would achieve an outcome $\bar{v}$ closer to $r_5 = 0.88$.

After voter n has voted using Equation 1, either voter n's ideal outcome r_n has been realized or voter n has moved the outcome as close to r_n as is immediately possible. Note also that $\bar{v} \in [0, 1]$ even though $v_n \in [m, M]$.

2.2 Equilibrium for n strategic voters

We have thus far allowed only voter n to use a rationally optimal strategy, requiring all other voters to express their sincere approval ratings. We now consider the properties of the more practical $[m, M]$-Average system in which each voter i uses a rationally optimal strategy to compute an expressed approval rating, based on i's sincere approval rating r_i and on the expressed votes of all other voters. When each voter i establishes v_i, other voters may wish to update their expressed approval ratings.

While there are many possible schemes that could accommodate iterative changes in expressed preferences, we examine the more general issue of reaching an equilibrium: each voter i has arrived at an expressed preference v_i such that the rationally optimal strategy recommends no change in v_i:

$$(\forall i)\ v_i = \min\left(\max\left(r_i n - \sum_{j\neq i} v_j, m\right), M\right) \tag{3}$$

So, at equilibrium, $(\forall i)\ (\bar{v} < r_i \wedge v_i = M) \vee (\bar{v} = r_i) \vee (\bar{v} > r_i \wedge v_i = m)$, and it follows that

$$(\forall i)\ \bar{v} < r_i \longrightarrow v_i = M \tag{4}$$

and

$$(\forall i)\ \bar{v} > r_i \longrightarrow v_i = m \tag{5}$$

Equation 4 says that for every i such that $\bar{v} < r_i$, $v_i = M$. So we can place a lower bound on the sum of all v_is by assuming all other v_is are at the minimum:

$$m \cdot |\{i : \bar{v} \geq r_i\}| + M \cdot |\{i : \bar{v} < r_i\}| \leq \sum_{i=1}^{n} v_i = \bar{v}n$$

Similarly, Equation 5 says that for every i such that $\bar{v} > r_i$, $v_i = m$. So we can place an upper bound on the sum of all v_is by assuming all other v_is are at the maximum:

$$\bar{v}n = \sum_{i=1}^{n} v_i \leq m \cdot |\{i : \bar{v} > r_i\}| + M \cdot |\{i : \bar{v} \leq r_i\}|$$

So we have

$$m \cdot |\{i : \bar{v} \geq r_i\}| + M \cdot |\{i : \bar{v} < r_i\}| \leq \bar{v}n \leq m \cdot |\{i : \bar{v} > r_i\}| + M \cdot |\{i : \bar{v} \leq r_i\}|$$

which implies [12, ch. 3]

$$|\{i : \bar{v} < r_i\}| \leq \frac{\bar{v} - m}{M - m} n \leq |\{i : \bar{v} \leq r_i\}|$$

Thus any average at equilibrium must satisfy the two equations

$$|\{i : \bar{v} < r_i\}| \leq \frac{\bar{v} - m}{M - m} n \tag{6}$$

and

$$\frac{\bar{v} - m}{M - m} n \leq |\{i : \bar{v} \leq r_i\}| \tag{7}$$

3 Multiple equilibria can exist

For some sincere-ratings vectors $\vec{r}$, multiple equilibria exist: there exist more than one $\vec{v}$ satisfying Equation 3. For example, if minimum vote $m = 0$, maximum vote $M = 1$ and $\vec{r} = [0.4, 0.7, 0.7, 0.8, 0.88]$ (a slight tweak to the *Videodrome* example), then any $\vec{v} = [0, v_2, v_3, 1, 1]$, where $v_2 + v_3 = 1.5$, satisfies Equation 3 and thus represents an equilibrium from which the optimal strategy would change no voter's vote.

In this case, at each possible equilibrium the outcome is $\bar{v} = 0.7$ (the ideal outcome of the two voters "conspiring" to keep it there). This is no coincidence; in general, it turns out that, even when multiple equilibria exist, the average at equilibrium is unique.

4 At most one equilibrium average rating can exist

We have seen that, given a length-n vector $\vec{r}$ of sincere ratings where $0 \leq r_i \leq 1$ for $1 \leq i \leq n$, any equilibrium $\vec{v}$ that results from every voter's using the optimal strategy will have a $\phi = \bar{v}$ that satisfies the inequalities

$$|\{i : \phi < r_i\}| \leq \frac{\phi - m}{M - m} n \tag{8}$$

and

$$\frac{\phi - m}{M - m} n \leq |\{i : \phi \leq r_i\}| \tag{9}$$

It turns out that at most one such ϕ exists for a given $\vec{r}$:

Theorem 4.1. *Given a vector $\vec{r}$ of length n where $0 \leq r_i \leq 1$ for $1 \leq i \leq n$,*

$$\begin{array}{l} |\{i : \phi_1 < r_i\}| \leq \frac{\phi_1 - m}{M - m} n \leq |\{i : \phi_1 \leq r_i\}| \quad \wedge \\ |\{i : \phi_2 < r_i\}| \leq \frac{\phi_2 - m}{M - m} n \leq |\{i : \phi_2 \leq r_i\}| \quad \longrightarrow \quad \phi_1 = \phi_2 \end{array}$$

(The proof considers two symmetric cases, $\phi_1 < \phi_2$ and $\phi_2 < \phi_1$, and shows by contradiction that each is impossible.)

5 At least one equilibrium always exists

It does little good to show that all equilibria will have equal averages if an equilibrium does not always exist. Fortunately, for any set of sincere preferred outcomes $\vec{r}$, there will always be at least one equilibrium $\vec{v}$ such that no voter i would choose to change v_i according to the optimal Average strategy defined above.

We can show that a particular procedure will always find an equilibrium. We use the *Videodrome* example ($\vec{r} = [0.4, 0.7, 0.8, 0.8, 0.88]$ with $m = 0$, $M = 1$) again for demonstration. This time, let us say initial votes are assumed to be, not sincere, but *zero* (the minimum allowed vote): $\vec{v} = [0, 0, 0, 0, 0]$. Then we again allow voters to revise their votes in order, from voter 5 down to voter 1. (This particular order will prove significant.) First, voter 5 deliberates:

$$v_5 = \min\left(\max\left(r_5 n - \sum_{j \neq 5} r_j, 0\right), 1\right) = \min\left(\max\left(0.8 \cdot 5 - (0 + 0 + 0 + 0), 0\right), 1\right) = 1$$

and changes v_5 to 1. The voters then in turn reason similarly and change v_4 to 1, v_3 to 1, v_2 to 0.5 and v_1 to 0. The resulting vote vector, $\vec{v} = [0, 0.5, 1, 1, 1]$, is indeed the same equilibrium found above in section 2.2, this time going through the voters only once.

This procedure inspires the following straightforward algorithm, which takes a $\vec{r}$ as input and outputs an equilibrium $\vec{v}$, assigning to each v_i exactly once. It orders the voters by decreasing r_i values, then uses the optimal strategy for each voter i in order, implicitly making the assumption that $v_j = m$ for $j > i$.

Algorithm 5.1. *FindEquilibrium*$(\vec{r}, m, M)$:

sort $\vec{r}$ so that $(\forall i \leq j)\ r_i \geq r_j$
for $i = 1$ to n do
  $v_i \leftarrow \min\left(\max\left(r_i n - \sum_{k<i} v_k - (n-i)m, m\right), M\right)$
return $\vec{v}$

Note that the algorithm assigns a value between m and M, inclusive, to each v_i exactly once, and that the assignment to v_i does not depend on the values of v_j where $j > i$. Therefore, after Algorithm 5.1 completes, it must be true that

$$(\forall i)\ v_i = \min\left(\max\left(r_i n - \sum_{k<i} v_k - (n-i)m, m\right), M\right)$$

but this is not quite enough to see that the resulting $\vec{v}$ is an equilibrium. To see that, we must show that an intermediate voter would not change his or her vote even after later voters have voted:

Theorem 5.2. *For any $\vec{r}$, where $0 \leq r_i \leq 1$ for $1 \leq i \leq n$, the vote vector $\vec{v}$ returned by Algorithm 5.1 satisfies*

$$(\forall i)\ v_i = \min\left(\max\left(r_i n - \sum_{k \neq i} v_k, m\right), M\right)$$

(The proof essentially shows that, because of the way that Algorithm 5.1 orders the voters, a certain kind of "partial" equilibrium is satisfied after each step of the algorithm, which implies that an equilibrium is found after the last step.)

So an equilibrium $\vec{v}$ must always exist for any input $\vec{r}$ and any $m \leq 0$ and $M \geq 1$.

We now know that, given some sincere-preference vector $\vec{r}$,

- at most one value ϕ satisfies Equations 8 and 9 (Theorem 4.1),
- any equilibrium $\vec{v}$ has average vote $\bar{v}$ satisfying Equations 6 and 7 (section 2.2), and
- at least one equilibrium $\vec{v}$ must exist (Theorem 5.2)

and so we can conclude that any ϕ that satisfies Equations 8 and 9 must equal the average vote $\bar{v}$ at all possible equilibria $\vec{v}$.

6 Average-Approval-Rating DSV

We have seen that Algorithm 5.1, *FindEquilibrium*, always finds an equilibrium for any sincere-preference vector $\vec{r}$. We also know that any equilibrium $\vec{v}$ will have the same average $\bar{v}$ (and that $0 \le \bar{v} \le 1$). It follows that the average at equilibrium is unique and can be defined as a function:

Algorithm 6.1. *AverageAtEquilibrium*$(\vec{r}, m, M)$:
$\vec{v} \leftarrow \text{FindEquilibrium}(\vec{r}, m, M)$
return $\bar{v} = \frac{\sum_{i=1}^{n} v_i}{n}$

Even when $m < 0$ and/or $M > 1$, *AverageAtEquilibrium* will return an outcome between 0 and 1. In fact, the outcome returned will be within the range defined by the input vector of cardinal preferences:

Theorem 6.2. $(\forall m \le 0, M \ge 1)\ \min(\vec{r}) \le \text{AverageAtEquilibrium}(\vec{r}, m, M) \le \max(\vec{r})$.

6.1 Declared-Strategy Voting

In 1996, Lorrie Cranor and Ron K. Cytron [9] described a hypothetical voting system they called Declared-Strategy Voting (DSV). DSV can be seen as a meta-voting system, in that it uses voters' expressed preferences among alternatives to vote rationally in their stead in repeated simulated elections. The repeated simulated elections are run according to the rules of some underlying voting protocol, which can be any protocol that accepts any kind of ballots and uses them to choose one outcome. Cranor [8] explored using DSV with plurality, but DSV, as a meta-voting system, could conceivably work with any voting protocol for which a rationally optimal strategy can be described, such as Average aggregation.

6.2 A new class of rating systems

The Average and Median protocols necessarily take a vote vector $\vec{v}$ as input—voters' sincere preference information cannot be directly and reliably elicited, so $\vec{r}$ is not generally available. If the Average system is used and voters are rationally strategic (and are allowed to keep changing their votes until all decide to stop), the outcome can reasonably be expected to equal AverageAtEquilibrium$(\vec{r}, 0, 1)$. But instead of using Average on the vote vector $\vec{v}$ and relying on the voters to use optimally rational strategy when deciding on their votes v_i, AverageAtEquilibrium$(\vec{v}, 0, 1)$ can be calculated and taken as the outcome, implicitly and effectively using the DSV framework with Average as the underlying voting protocol. In fact, we are not limited to AverageAtEquilibrium$(\vec{v}, 0, 1)$; AverageAtEquilibrium$(\vec{v}, m, M)$ lies between 0 and 1 for any $m \le 0$ and $M \ge 1$ and so can serve as a rating system as well.

For illustration, we reuse the *Videodrome* example and assume sincere voters: $\vec{v} = [0.4, 0.7, 0.8, 0.8, 0.88]$. Suppose we want to take as the outcome of this election not the average vote $\bar{v}$ or the median vote $\tilde{v}$ but AverageAtEquilibrium$(\vec{v}, 0, 1)$. First we calculate FindEquilibrium$(\vec{v}, 0, 1)$, which we have seen in section 2.2 to be

$$\vec{w} = \text{FindEquilibrium}(\vec{v}, 0, 1) = [0, 0.5, 1, 1, 1]$$

Then we see that

$$\bar{w} = \frac{\sum_{i=1}^{5} w_i}{5} = \frac{0 + 0.5 + 1 + 1 + 1}{5} = 0.7$$

giving the outcome as 0.7, which equals neither the Average outcome ($\bar{v} = 0.716$) nor the Median outcome ($\tilde{v} = 0.8$).

Alternatively, we can let $m = -99$ and $M = 100$. Then the equilibrium we find turns out to be

$$\vec{w} = \text{FindEquilibrium}(\vec{v}, -99, 100) = [-99, -99, 2, 100, 100]$$

And then

$$\bar{w} = \frac{\sum_{i=1}^{5} w_i}{5} = \frac{-99 + (-99) + 2 + 100 + 100}{5} = 0.8$$

This time the effective power to determine the outcome fell to voter 3 rather than voter 2, giving the Median outcome of 0.8. (We will see that if $m + M = 1$ and $M - m$ is allowed to become large enough, the resultant outcome will equal the Median outcome.)

It turns out that in neither of these cases will any voter be able to gain from voting insincerely. This is no coincidence.

This Average-Approval-Rating (AAR) DSV system has three intuitively desirable properties: a kind of monotonicity (Theorem 6.3), immunity to Average-like strategy (Theorem 6.4) and a general nonmanipulability (Theorem 6.5). The first two will imply the third.

6.3 Monotonicity of AAR DSV

First, the monotonicity property: When some input votes are increased and none is decreased, the outcome never decreases.

Theorem 6.3. *If $\vec{v} = [v_1, v_2, \ldots v_n]$ and $\vec{v}' = [v'_1, v'_2, \ldots v'_n]$ where $(\forall i)$ $v_i \leq v'_i$, then* AverageAtEquilibrium$(\vec{v}, m, M) \leq$ AverageAtEquilibrium$(\vec{v}', m, M)$.

(The proof is by contradiction.)

6.4 AAR DSV is immune to Average-style strategy

Another desirable property of AAR DSV is that its outcome is unaffected by voters' using Average-style strategy, trying to move the outcome in the desired direction by moving their votes in that direction.

Theorem 6.4. *If $\vec{v} = [v_1, v_2, \ldots v_n]$ and $\vec{v}' = [v'_1, v'_2, \ldots v'_n]$ where, for all $1 \leq i \leq n$,*

- $v'_i \leq v_i$ *if* AverageAtEquilibrium$(\vec{v}, m, M) > v_i$
- $v'_i = v_i$ *if* AverageAtEquilibrium$(\vec{v}, m, M) = v_i$
- $v'_i \geq v_i$ *if* AverageAtEquilibrium$(\vec{v}, m, M) < v_i$

then AverageAtEquilibrium$(\vec{v}', m, M) =$ AverageAtEquilibrium$(\vec{v}, m, M)$.

(The proof relies on Theorem 4.1.)

6.5 AAR DSV never rewards insincerity

For any voting system, it is desirable to show that a voter can never gain a better outcome by voting insincerely than by voting sincerely, however sincerity is defined. It turns out that, when AverageAtEquilibrium$(\vec{v}, m, M)$ is selected as the outcome, no voter i can gain an outcome closer to the ideal r_i by voting $v_i \neq r_i$ instead of $v_i = r_i$, guaranteeing a strong nonmanipulability property to AAR DSV:

Theorem 6.5. *If $\vec{v} = [v_1, v_2, \ldots v_n]$ where $v_1 = r_1$ and $\vec{v}' = [v'_1, v'_2, \ldots v'_n]$ where $v'_1 \neq r_1$ and $(\forall i > 1)$ $v'_i = v_i$, then* $|\text{AverageAtEquilibrium}(\vec{v}, m, M) - r_1| \leq |\text{AverageAtEquilibrium}(\vec{v}', m, M) - r_1|$.

(The proof consists of four cases and relies on Theorems 6.3 and 6.4.)

7 Evaluation of AAR DSV systems

To simplify the evaluation of AAR DSV systems, we re-parameterize them by defining

$$\Phi_{a,b}(\vec{v}) \equiv \lim_{x \to a^+} \text{AverageAtEquilibrium}\left(\vec{v}, b - \frac{b}{x}, b + \frac{1-b}{x}\right)$$

(The limit is needed for the $a = 0$ case; as a approaches 0, $\Phi_{a,b}(\vec{v})$ approaches the ${}^b\tilde{v}$ outcome defined in section 1.3.)

Any system that uses the outcome function $\Phi_{a,b}(\vec{v})$ where $0 \leq a \leq 1$ and $0 \leq b \leq 1$ has the property that no voter can gain by voting insincerely. But it does not follow that any values of a and b give equally desirable outcomes.

One approach to evaluating this continuous range of nonmanipulable systems is to take the Average system as a benchmark and determine which $\Phi_{a,b}$ function comes nearest, on average, to giving the Average outcome. Given a vote vector $\vec{v}$, we can calculate the Average outcome $\bar{v}$ and the outcome $\Phi_{a,b}(\vec{v})$ for many a, b combinations. For any particular a and b, we can calculate the squared error from $\bar{v}$: $\text{SE}_{a,b}(\vec{v}) = (\Phi_{a,b}(\vec{v}) - \bar{v})^2$. If $\mathbf{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3 \ldots \vec{v}_N\}$ is a vector of N vote vectors, then we can find the root-mean-squared error from Average, weighted by the number of ratings in each vote vector $\vec{v}_i$:

$$\text{RMSE}_{a,b}(\mathbf{V}) = \sqrt{\frac{\sum_{i=1}^{N} |\vec{v}_i| \cdot \text{SE}_{a,b}(\vec{v}_i)}{\sum_{i=1}^{N} |\vec{v}_i|}}$$

Given some "training" vector $\mathbf{V}$ of vote vectors, we would like to choose a and b to minimize $\text{RMSE}_{a,b}(\mathbf{V})$.

This approach requires a concrete source of vote-vector data or a distribution for generating such. The website Metacritic [1] offers ideal data for our purposes: Reviews for over 4000 films are summarized into ratings between 0 and 100. For example, one film[3] has the seven ratings 70, 70, 80, 80, 88, 88 and 100, which are easily converted into the vote vector $\vec{v} = [0.7, 0.7, 0.8, 0.8, 0.88, 0.88, 1]$. Converting all films on Metacritic the same way gives us a large vector $\mathbf{V}$ of vote vectors.[4]

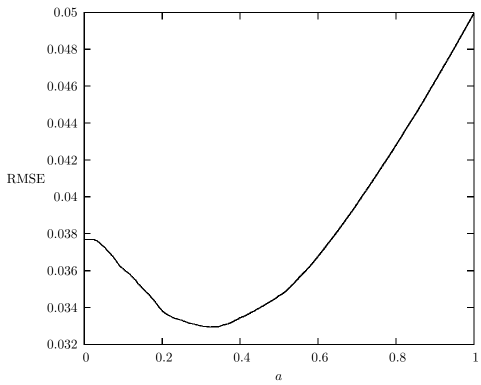
Since there are two parameters, a and b, it is somewhat impractical to try all combinations. But it may be desired to fix $b = 0.5$ to ensure a kind of symmetry: If $(\forall i)$ $v'_i = 1 - v_i$,

[3] The 1978 film *Animal House*.

[4] We use the data for the 4581 films mined from Metacritic on Thursday, 3 April 2008, that had at least three critics rate them. Note that we are implicitly assuming that the rating data are sincere; unfortunately, we know of no large data set gathered using a nonmanipulable rating protocol such as ours, so we must hope that most critics are more interested in maintaining their professional reputations than in optimizing a film's Metacritic rating.

then $(\forall a)\ \Phi_{a,0.5}(\vec{v}') = 1 - \Phi_{a,0.5}(\vec{v})$, so electorates that prefer low and high outcomes are treated symmetrically. Fixing $b = 0.5$ and trying all 10001 evenly spaced values of a, we find that $a = 0.3240$ (Figure 1) gives the minimum RMSE for the Metacritic data.

Figure 1: RMSE, varying a and fixing $b = 0.5000$



Having fixed $b = 0.5$ and found the value of a that minimizes RMSE (0.3240), we can now fix $a = 0.3240$ and find the value of b that minimizes RMSE, then fix b again accordingly and continue in a hill-climbing fashion until we find a stable minimum. In practice, the procedure is guaranteed to halt because the RMSE decreases at each step for which either a or b changes.

Using this procedure on the Metacritic data and testing 10001 evenly spaced values of a or b at each step, whether we start with $a \in \{0, 0.25, 0.5, 0.75\}$ or with $b \in \{0, 0.25, 0.5, 0.75\}$,[5] we find a local RMSE minimum (approximately 0.03242) at $a = 0.3647$, $b = 0.4820$; such a system is equivalent to running an Average election with rationally optimal voters and allowing votes between $m \approx -0.8396$ and $M \approx 1.9023$.

Other preference domains may have very different properties and thus different ideal values for a and b.

8 Related and future work

In this paper we have applied the DSV framework of Cranor and Cytron [9] to create a large class of nonmanipulable rating systems, assuming only that each voter has a single-peaked preference function over the bounded, one-dimensional outcome space. The single-peaked assumption allows us to avoid the negative implications of the Gibbard-Satterthwaite theorem [11, 15], making it possible to find nonmanipulable protocols that have no dictator.

Most relevant to our work is Moulin's [14] result. He characterized the set of all non-manipulable protocols that resolve a vector of real inputs into one real outcome, showing that any such protocol is equivalent to adding some fixed set of $n-1$ points to the n input points and taking as the outcome the median of the combined set. It turns out that our AverageAtEquilibrium$(\vec{v}, m, M)$ is equivalent to adding the (evenly spaced) points

$$m + \frac{1}{n}(M-m), m + \frac{2}{n}(M-m), \ldots m + \frac{i}{n}(M-m), \ldots m + \frac{n-1}{n}(M-m)$$

[5] Note that when $a = 1$, the outcome is simply AverageAtEquilibrium$(\vec{v}, 0, 1)$ and does not depend on the the b parameter, so different values of b cannot be compared. When b is set to 1, RMSE turns out to be minimized at $a = 1$.

to the input points (which fall between 0 and 1) and taking the median of that set as the outcome, so our set of AAR DSV systems is indeed a subset of Moulin's set of nonmanipulable protocols.

In future work we plan to explore higher-dimensional outcome spaces. The Median system can be perhaps most naturally generalized to $d > 1$ dimensions by finding the point t that minimizes $\sum_{i=1}^{n} \operatorname{dist}(t, v_i)$, where $\operatorname{dist}(t, v_i) = \left(\sum_{j=1}^{d} (t_j - v_{ij})^2\right)^{1/2}$, the Euclidean distance between t and v_i. t is known as the Fermat–Weber point [18, 7]. When $d > 1$, unlike in the one-dimensional case, it usually has a single optimum point even when n is even (the only exception is an even number of collinear points). Unfortunately, there is no computationally feasible exact algorithm to calculate the Fermat–Weber point in general [3], but numerical approximation is quite easy [17, 6].

The Fermat–Weber point does not change when a point v_i is moved farther away from t in the direction of the vector from t to v_i [16], so, in a sense, direction matters but not distance. Because of this property, a naïve Average-style strategy for manipulating this Fermat–Weber system fails, and any successful manipulation would have to move a sincere vote in some other direction. Unfortunately, an insincere voter can indeed manipulate the Fermat–Weber point to move closer to his or her ideal outcome [12, ch. 3]. In fact, Zhou [19] showed that no protocol with effective outcome space of dimension $d > 1$ is generally nonmanipulable when voters can have any single-peaked (concave) preference function.

The Average system is easily generalized to higher-dimension hypercubes by taking the average of each coordinate, effectively calculating the centroid, the center of mass given a set of unit masses. This generalization is equivalent to finding the point t that minimizes $\sum_{i=1}^{n} \operatorname{dist}(t, v_i)^2$. The resulting system is equivalent to conducting d separate and independent Average elections, and the results above for strategic behavior under the one-dimensional Average system apply to the "election" for each coordinate. In particular, if each voter has *separable* preferences [5] (preferences in one dimension are independent of preferences in all other dimensions), conducting a d-dimensional AAR DSV election is equivalent to conducting d parallel one-dimensional AAR DSV elections, and so gives a nonmanipulable system. Such a preference-function space is not *abundant* by Zhou's definition.

The one-dimensional space between 0 and 1 can be generalized in other ways than into hypercubes. For example, the outcome space could be the d-dimensional simplex (for example, $\{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$), which could describe the division of a limited resource among several uses (such as a committee allocating a fixed sum among budget items). Unfortunately, even when all voters' preferences are separable, AAR DSV systems may be manipulable—in a sense, dimensions are interdependent for the outcome space itself. It may be, however, that no voter can move the outcome to one which is closer to ideal on one dimension without moving it further on some other dimension. We plan to investigate this "dominance"-nonmanipulability.

We would like to thank Steven Brams and the anonymous reviewers for their valuable comments and suggestions on this work.

References

[1] Metacritic. `http://www.metacritic.com/`.

[2] Rotten Tomatoes. `http://www.rottentomatoes.com/`.

[3] Chandrajit L. Bajaj. The algebraic degree of geometric optimization problems. *Discrete & Computational Geometry*, 3:177–191, 1988.

[4] Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34, February 1948.

[5] Kim C. Border and J. S. Jordan. Straightforward elections, unanimity and phantom voters. *The Review of Economic Studies*, 50(1):153–170, January 1983.

[6] Prosenjit Bose, Anil Maheshwari, and Pat Morin. Fast approximations for sums of distances, clustering and the Fermat-Weber problem. *Computational Geometry: Theory and Applications*, 24(3):135–146, April 2003.

[7] Jack Brimberg. The Fermat-Weber problem revisited. *Mathematical Programming*, 71:71–76, 1995.

[8] Lorrie Faith Cranor. *Declared-Strategy Voting: An Instrument for Group Decision-Making*. PhD thesis, Department of Computer Science, Washington University, St. Louis, Missouri, December 1996.

[9] Lorrie Faith Cranor and Ron K. Cytron. Towards an information-neutral voting scheme that does not leave too much to chance. In *Midwest Political Science Association Annual Meeting*, pages 1–15, April 1996.

[10] Anthony Downs. *An Economic Theory of Democracy*. Harper, New York, New York, 1957.

[11] Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(3):587–601, May 1973.

[12] Rob LeGrand. *Computational Aspects of Approval Voting and Declared-Strategy Voting*. PhD thesis, Department of Computer Science and Engineering, Washington University, St. Louis, Missouri, April 2008.

[13] Samuel Merrill III. *Making Multicandidate Elections More Democratic*. Princeton University Press, Princeton, New Jersey, 1988.

[14] Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, January 1980.

[15] Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, April 1975.

[16] Christopher G. Small. A survey of multidimensional medians. *International Statistical Review*, 58(3):263–277, December 1990.

[17] Y. Vardi and Cun-Hui Zhang. A modified Weiszfeld algorithm for the Fermat-Weber location problem. *Mathematical Programming*, 90(3):559–566, 2001.

[18] Alfred Weber. *Alfred Weber's Theory of the Location of Industries*. University of Chicago Press, Chicago, Illinois, 1929. Translated from the German by C. J. Freidrich.

[19] Lin Zhou. Impossibility of strategy-proof mechanisms in economies with pure public goods. *The Review of Economic Studies*, 58(1):107–119, January 1991.

Rob LeGrand Ron K. Cytron
Department of Computer Science & Engineering
Washington University
One Brookings Drive, Campus Box 1045
St. Louis, Missouri 63130-4899
Email: `{legrand,cytron}@cse.wustl.edu`