Voter Model with Stubborn Agents: from Theoretical Solutions to Prediction of Political Elections

Antoine Vendeville, Benjamin Guedj & Shi Zhou

University College London
Inria, Lille – Nord Europe research centre, France

July 5, 2021

Papers

- A. Vendeville, B. Guedj & S. Zhou. Voter model with stubborn agents on strongly connected social networks. https://arxiv.org/abs/2006.07265
- A. Vendeville, B. Guedj & S. Zhou. Forecasting elections results via the voter model with stubborn nodes. *Applied Network Science* 6, 1 (2021). https://doi.org/10.1007/s41109-020-00342-7.

The voter model

- Social graph with n users.
- User i has opinion $x_i(t)$ at time t.

At the times of a Poisson process of parameter n, an user chosen uniformly at random adopts the opinion of a random neighbour of them.

Consensus [Aldous and Fill, 2002]

If the network is finite and strongly connected, almost surely everyone eventually agrees:

$$\mathbb{P}(\exists t,\ \forall i,j,\ x_i(t) = x_j(t)) = 1.$$

Our framework

- Complete graph.
- $N_1(t) :=$ quantity of opinion-1 users at time t.
- (s_0, s_1) stubborn agents: $0 \leqslant s_1 \leqslant n_1$ and $0 \leqslant s_0 \leqslant n - s_1$.

Our framework

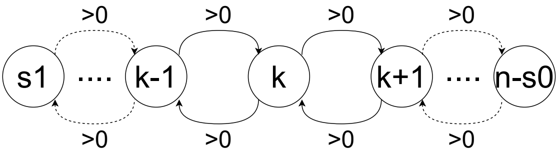
- Complete graph.
- $N_1(t) :=$ quantity of opinion-1 users at time t.
- (s_0, s_1) stubborn agents: $0 \leqslant s_1 \leqslant n_1$ and $0 \leqslant s_0 \leqslant n - s_1$.

Long-term behaviour

1. $s_0 > 0, s_1 = 0 \Longrightarrow$ almost surely $x_i(t) \to 0$ for all i.
2. $s_0 = 0, s_1 > 0 \Longrightarrow$ almost surely $x_i(t) \to 1$ for all i.
3. $s_0 > 0, s_1 > 0 \Longrightarrow$ equilibrium state.

Modelling

$N_1(t)$ describes a birth-and-death process over the state-space $\{s_1, \ldots, n - s_0\}$.



Transition rates:

$$\begin{cases} q_{k,k-1} = (k - s_1)(n-k)/(n-1) \\ q_{k,k+1} = k(n-k-s_0)/(n-1) \\ q_{k,k} = -q_{k,k-1} - q_{k,k+1}. \end{cases} \tag{1}$$

Distribution at time t

n_1: quantity of opinion 1 at initialisation.

The event $\{N_1(t) = k\}$ has probability

$$p_{n_1,k}(t) := [e^{tQ}]_{n_1,k}. \tag{2}$$

and

$$\mathbb{E}N_1(t) = \sum_{k=s_1}^{n-s_0} k p_{n_1,k}(t). \tag{3}$$

is the expected number of opinion-1 holders at time t.

Reference [Norris, 1997].

Equilibrium

Long-term average

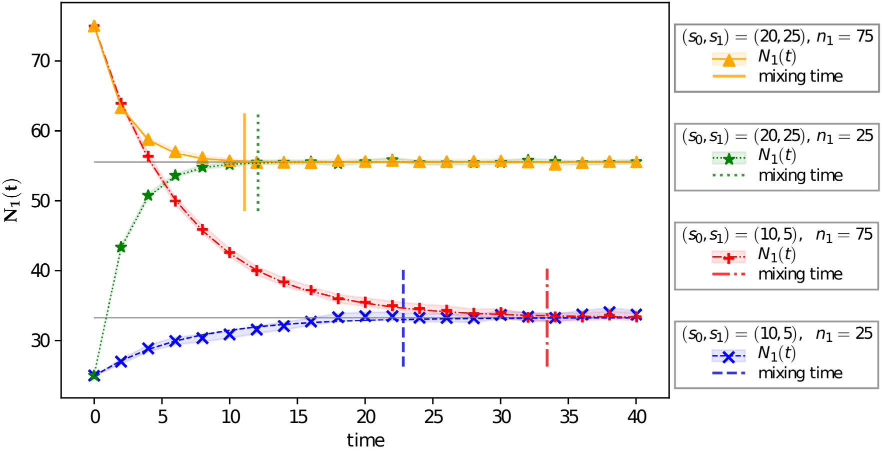
At equilibrium we expect a number

$$\mathbb{E}_\pi N_1(t) = n\frac{s_1}{s_0 + s_1}$$

of opinion-1 holders.

$s_1/(s_0 + s_1)$: average opinion at equilibrium.

Numerical simulations



$n = 100, \ \varepsilon = 10^{-2}, \ 1000$ simulations.

Application to the prediction of elections result

Objective: predict the result of general elections (UK) and presidential elections (US) based on previous ones.

Method: assume result of previous elections are a realisation of the voter model with stubborn nodes, estimate parameters and predict future elections via Theorem 1.

Data: [UK] percentage of popular votes for Labour and Conservative parties in each general election since 1922.
[US] percentage of popular votes for Democratic and Republican parties in each presidential election since 1912.

Method

Choose $a \in \{\text{Labour, Conservative, Democratic, Republican}\}$.
x_i^a: percentage of votes for party a in the i^{th} election, rounded to the nearest integer.

$$\rightarrow \text{How to predict } x_i^a \text{ knowing } x_0^a, \dots, x_{i-1}^a?$$

Method

Choose $a \in \{\text{Labour, Conservative, Democratic, Republican}\}$.
x_i^a: percentage of votes for party a in the i^{th} election, rounded to the nearest integer.

$$\rightarrow \text{How to predict } x_i^a \text{ knowing } x_0^a, \dots, x_{i-1}^a?$$

Forecast of election i

Assume $x_0^a, \dots, x_{i-1}^a$ to be a realisation of the voter model with stubborn agents with $n = 100$ nodes. If we can estimate the values of the parameters (s_0, s_1) we get the theoretical distribution of x_i^a.

Parameters estimation

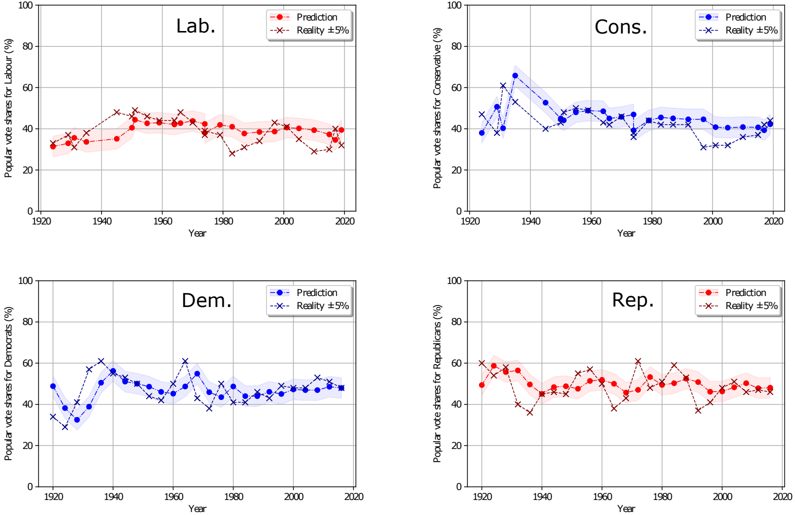
Maximum-likelihood $(i \geqslant 3)$

$$(s_0^\star, s_1^\star) = \underset{s_0, s_1}{\operatorname{argmax}} \sum_{j=1}^{i-2} \log \left(p^{(s_0,s_1)}_{x_j^a, x_{j+1}^a}(t_{j+1} - t_j) \right) \tag{4}$$

$p^{(s_0,s_1)}_{x_j^a, x_{j+1}^a}(t_{j+1} - t_j)$ probability that the voter process with parameters (s_0, s_1) equals x_{j+1}^a at time t_{j+1} knowing it equaled x_j^a at time t_j.

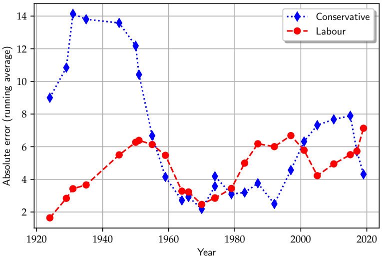
Searches for parameters that maximise the probability of the (known) results for elections $1, \ldots, i-1$.

Predictions result

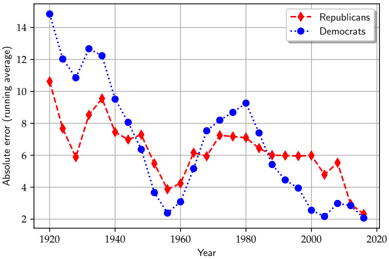


Predictions accuracy

United Kingdom



United States



Mean average error in predictions: 4.74%

Predict last result: 5.03%

Tumasjan et al. [2011] 1.65%, Saleiro et al. [2016] 0.63%

Discussion and future research

- Other methods have better results...
- ... but $(s_0^\star, s_1^\star)$ is valuable information on the political landscape.
- Opinion manipulation: add/remove stubborn nodes into a community to reach target average opinion $\mathbb{E}\pi/n$.
- Non-complete user graphs?

References

D. J. Aldous and J. A. Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

J. R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997. doi: 10.1017/CBO9780511810633.

P. Saleiro, L. Gomes, and C. Soares. Sentiment aggregate functions for political opinion polling using microblog streams. In *Proceedings of the Ninth International Conference on Computer Science &; Software Engineering*, C3S2E '16, page 44–50, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450340755. doi: 10.1145/2948992.2949022.

A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe. Election forecasts with twitter: How 140 characters reflect the political landscape. *Soc. Sci. Comput. Rev.*, 29(4):402–418, 2011. doi: 10.1177/0894439310386557.