An agent-based modelling approach to analyse the public opinion on politicians

Brouwers, Thijs M.A. ; Onneweer, John P.T.; Treur, J.

published in
Social Informatics
2018

DOI (link to publisher)
10.1007/978-3-030-01129-1_7

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Link to publication in VU Research Portal

citation for published version (APA)
Brouwers, T. M. A., Onneweer, J. P. T., & Treur, J. (2018). An agent-based modelling approach to analyse the public opinion on politicians. In O. Koltsova, D. I. Ignatov, & S. Staab (Eds.), *Social Informatics: Proceedings of the 10th International Conference on Social Informatics, SocInfo'18* (Vol. 1, pp. 102-116). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 11185 LNCS). Springer. https://doi.org/10.1007/978-3-030-01129-1_7

An Agent-Based Modelling Approach to Analyse the Public Opinion on Politicians

Thijs M. A. Brouwers, John P. T. Onneweer, and Jan Treur(✉)

Behavioural Informatics Group, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
tbrouwers95@gmail.com, johnonneweer@gmail.com, j.treur@vu.nl

Abstract. A politician's popularity can be measured by polls or by measuring the amount of times a politician is mentioned on the Internet in a positive or negative manner. This paper introduces an approach to an agent-based computational model to model a politician's popularity within a population that participates on the Internet over time. A particle swarm optimization algorithm is used for parameter tuning to identify the characteristics of all agents based on the analysis of public opinions on a politician found on the Internet. The properties of the network are verified by applying a social network analysis. A mathematical analysis is used to get more in depth understanding on the model and to verify its correctness.

Keywords: Opinion dynamics · Network model · Politician

1 Introduction

The process of opinion formation cannot be separated from the specific social network in which an individual is situated. An individual observes the behavior of and receives information from a small subset of society, consisting of friends, family, coworkers and peers, and a certain group of (opinion) leaders that he or she listens to and respects; all of them can have a substantial influence on the individual's opinion [9, 11]. Interactions will lead to dynamics in opinion formation [1]. The opinion dynamics in a network may lead to consensus among or polarization between the agents or, more general, to a certain fragmentation of the opinion patterns [5].

Information obtained within a social context can be interpreted differently among individuals. Some information will be more trusted than others and conjectures are formed about members and their intentions in the social network [1]. In a population or group, it seems that multiple groups, specifically groups of easily influenced individuals influence the opinion of an entire network more than powerful individuals alone [17].

Thierry Baudet is a much discussed politician in the Netherlands in 2017. He is the founder and leader of the political party Forum for Democracy. With this newly, right-wing and national conservative party, he managed to win two seats in the House of Representatives in the general election of 2017. After being elected, he made a number of striking statements and acts, causing a lot of fuss in the media. With these, sometimes controversial, statements and acts, he has made himself very popular with some,

S. Staab et al. (Eds.): SocInfo 2018, LNCS 11185, pp. 102–116, 2018.
https://doi.org/10.1007/978-3-030-01129-1_7

while others became fiercely against him. According to the polls of Maurice de Hond, the party would win fifteen seats at the end of 2017. It can be argued that the popularity in a network influences the popularity in the polls.

Within a social network, which is exposed to information as described above, a process of opinion formation takes place. In general, agents will, to a certain extent, take into account the opinions of other agents when forming their own opinion. They will neither simply share nor firmly disregard the opinions of others, which can be modelled by different weights on the opinions of others. There also may be a difference in how easily agents change their opinion, which can be modelled as well.

The goal of this paper is to present an agent-based computational model to simulate the overall opinion on Baudet over time. It takes into account agents representing persons who have access to and post on web pages. The computational model was designed by a Network-Oriented Modelling approach based on temporal-causal networks, which in previous work has turned out a useful and easy to use means to model social networks and their dynamics [16]. Manual sentiment analysis was performed to determine whether the contents of the web pages had a positive or negative tone on Baudet.

2 The Temporal-Causal Network Model

In this section, first the Network-Oriented Modeling approach used [16] is briefly explained. This Network-Oriented Modeling approach is based on temporal-causal networks. Causal modeling, causal reasoning and causal simulation have a long tradition in AI; e.g., [7, 8, 12]. The Network-Oriented Modeling approach based on temporal-causal networks described in [16] can be viewed on the one hand as part of this causal modeling tradition, and on the other hand from the perspective on mental states and their causal relations in Philosophy of Mind (e.g., [6]), and from the perspective of social networks where nodes affect each other. It is a widely usable generic dynamic AI modeling approach that distinguishes itself by incorporating a dynamic and adaptive temporal perspective, both on states and on causal relations. This dynamic perspective enables modeling of cyclic and adaptive networks, and also of timing of causal effects.

Temporal-causal network models can be represented at two levels: by a conceptual representation and by a numerical representation. These representations can be used to display the network graphically and for numerical simulations. The model usually includes a number of parameters for domain, person or social context-specific characteristics. Parameter tuning methods are available to estimate the values for such parameters.

A conceptual representation of a temporal-causal network model represents, in a declarative manner, states and connections between those states that represent impacts of states on each other, as assumed to hold for the addressed application domain. The states have activation levels that vary over time. Connections can have different strengths, and states can be affected by more than one other state. Moreover, states can have a different extent of flexibility; some states may be able to change fast, while other states are a more rigid and change more slowly. These three notions are covered by elements in the Network-Oriented Modeling approach based on temporal-causal networks, and are reflected in the conceptual representation of a temporal-causal network model [15, 16]:

- **Connection weight** $\boldsymbol{\omega}_{X,Y}$: Each connection from a state X to a state Y has a connection weight value $\boldsymbol{\omega}_{X,Y}$ representing the strength of the connection.
- **Combination function** $\mathbf{c}_Y(..)$**:** For each state a combination function $\mathbf{c}_Y(..)$ is used to combine the causal impacts of other states on state Y.
- **Speed factor** $\boldsymbol{\eta}_Y$: For each state Y a speed factor $\boldsymbol{\eta}_Y$ is used to represent how fast a state is changing upon causal impact.

A simple conceptual representation of the designed model is shown in Fig. 1. A single agent of the population represents a part of the Dutch population who think alike. To get a manageable network model it was decided to limit the number of agents which each are considered in some way to represent the different types of persons in the real world. Since there are 150 seats in the Dutch House of Representatives and one seat should represent a part of the Dutch population with the same political preference, a total of 150 agents is used in the proposed network. The connections represent how they affect each other's opinions. The opinion state represents the overall opinion of the network as a whole. Because it can be discussed that voting behaviour of a population relates to but may not be exactly equal to their overall opinion about the politician or party, another state is included that represents the voting behavior of the population. The state values represent the opinion on the Dutch politician called Baudet. A state value of 1 means that the agent has only positive opinions on Baudet, whereas a state value of 0 means there are no positive opinions on Baudet within that agent, only neutral or negative opinions. An arrow between two agents means that there is a connection from one agent to the other agent by which the former agent influences the latter agent.

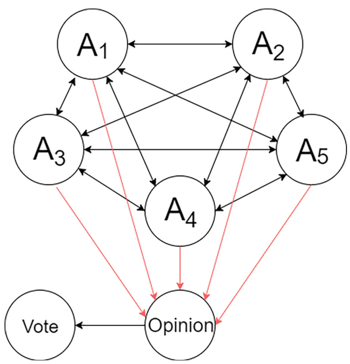

Fig. 1. A conceptual representation of the temporal-causal network model

A scale-free network is a network with a degree distribution that follows a power law. It has nodes with a degree that by far exceed the average degree of the network [2]. These highest-degree nodes are often called hubs, and are closely followed by other nodes with a smaller degree and so on. An important characteristic of a scale-free network is the clustering coefficient distribution, which decreases as the node degree increases. This distribution also follows a power law, which implies that low-degree nodes belong to dense subgroups [2]. Subgroups are connected to each other through hubs. In a social network, of the type used in this study, the subgroups are communities in which everyone knows everyone, and everybody has a few relationships to people outside the community. In such network, there are some people who are connected to a large number of communities within the network (e.g. celebrities, politicians).

However, in the proposed network the nodes have a high degree. The number of communities within the network thus will be small. The degrees are based on the activity of the members on the Internet. Some are online a lot, being influenced by others often. Others may be less online, resulting in fewer degrees. Because nowadays in life it is easy to propagate opinions on the Internet, with a high reach, we assumed that every member of the network has a high degree. Within the proposed scale-free network connections between agents are defined by their weight. A connection weight of 0 means there is no connection from one agent to another. Connection weights are numbers in the [0, 1] interval stand for a connection from one agent to another. This method implements that it is possible that agent X has a connection towards agent Y, while agent Y has no connection towards agent X. The connections are a basis for contagion between agents. All agents have an initial value which represents their initial opinion on Baudet. Over time agents will be influenced by other agents, by accessing the Internet. Here they will be influenced by the opinions of others, and influence others with their own opinion.

According to the Network-Oriented Modeling approach based on temporal-causal networks used, combination functions can have different forms. They specify for each state a way how multiple impacts on this state are aggregated. For this aggregation a variety of standard combination functions are available, among which sum and scaled sum combination functions as discussed below.

A conceptual representation of a model can be transformed systematically or even automatically into a numerical representation of the model as follows [15, 16]:

- at each time point t each state Y has a real number value $Y(t)$ in [0, 1]
  at each time point t each state X connected to state Y has an impact on Y defined as

$$\mathbf{impact}_{X,Y}(\boldsymbol{t}) = \boldsymbol{\omega}_{X,Y} X(t)$$

where $\boldsymbol{\omega}_{X,Y}$ is the connection weight

- The *aggregated impact* of multiple states X_i on Y at t is determined by:

$$\begin{aligned}\mathbf{aggimpact}_Y(\boldsymbol{t}) &= \mathbf{c}_Y(\mathbf{impact}_{X_1,Y}(\boldsymbol{t}), \ldots, \mathbf{impact}_{X_k,Y}(t)) \\ &= \mathbf{c}_Y(\boldsymbol{\omega}_{X_1,Y} X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y} X_k(t))\end{aligned}$$

where X_i are the states with connections to state Y

- The effect of $\mathbf{aggimpact_Y(t)}$ on Y is exerted *over time gradually*:

$$Y(t + \Delta t) = Y(t) + \boldsymbol{\eta}_Y[\mathbf{aggimpact}_Y(\boldsymbol{t}) - Y(t)]\Delta t$$
$$\text{or } \mathrm{d}Y(t)/\mathrm{d}t = \boldsymbol{\eta}_Y[\mathbf{aggimpact}_Y(\boldsymbol{t}) - Y(t)]$$

- Thus, the following *difference* and *differential equation* for Y are obtained:

$$Y(t + \Delta t) = Y(t) + \boldsymbol{\eta}_{\boldsymbol{Y}}\left[\mathbf{c}_{\boldsymbol{Y}}\left(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t)\right) - Y(t)\right]\Delta t$$
$$\mathbf{d}Y(t)/\mathbf{d}t = \boldsymbol{\eta}_Y[\mathbf{c}_{\boldsymbol{Y}}(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t)) - Y(t)]$$

Software environments are available in Matlab and Python that do this transformation and enable to run simulation experiments. For any set of values for the connection weights, speed factors and any choice for combination functions, each node gets a difference or differential equation assigned. For the model considered here, this makes a set of 152 coupled equations, that together, in mutual interaction, describe the model's behavior. Note that the parameters and in particular the speed factors enable to obtain realistic differences between individual agents in the model to tune the model to the characteristics of people in the real world.

For all agents a scaled sum function $\mathbf{ssum}_\lambda$(.,.) is used as combination function. The opinion state is defined by a sum function **sum**(.,.) [16]:

$$\mathbf{c}_Y(V_1, \ldots, V_k) = \mathbf{sum}(V_1, \ldots, V_k) = V_1 + \ldots + V_k$$
$$\mathbf{c}_Y(V_1, \ldots, V_k) = \mathbf{ssum}_\lambda(V_1, \ldots, V_n) = (V_1 + \ldots + V_k) / \lambda$$

where λ is the scaling factor. The voting state uses the identity function **id(.)**, defined as **id**(V) = V. In our network, where all connection weights are assumed $\geq$ 0, the scaling factor λ is defined as the sum of the incoming weights for any agent. For example, the scaling factor of agent 1 $\lambda = \omega_{A_1}$ is defined as the sum of the incoming weights of all other agents: $\omega_{A_1} = \omega_{B_1,A_1} + \ldots + \omega_{B_k,A_1}$. The combination function for agent 1 is

$$\mathrm{c}_{\mathrm{A}_1}(V_1, \ldots, V_k) = \mathbf{ssum}_{\omega_{\mathrm{A}_1}}(V_1, \ldots, V_k) = (V_1 + \ldots + V_k)/\omega_{\mathrm{A}_1}$$

The difference and differential equation are as follows:

$$\mathrm{A}_1(t + \Delta t) = \mathrm{A}_1(t) + \boldsymbol{\eta}_{\mathrm{A}_1}[\mathbf{ssum}_{\omega_{\mathrm{A}_1}}(\omega_{B_1,\mathrm{A}_1}B_1(t), \ldots, \omega_{B_k,\mathrm{A}_1}B_k(t)) - \mathrm{A}_1(t)]\Delta t$$
$$\mathbf{d}\mathrm{A}_1(t)/\mathbf{d}t = \boldsymbol{\eta}_{\mathrm{A}1}\left[\mathbf{ssum}_{\omega_{\mathrm{A}_1}}\left(\omega_{B_1,\mathrm{A}_1}B_1(t), \ldots, \omega_{B_k,\mathrm{A}_1}B_k(t)\right) - \mathrm{A}_1(t)\right]$$

The **sum**(.,.) combination function is used for the opinion state. For example, the difference and differential equation for state $\mathrm{A}_{\mathrm{opinion}}$ are as follows:

$$\mathrm{A}_{\mathrm{opinion}}(t + \Delta t) = \mathrm{A}_{\mathrm{opinion}}(t) + \boldsymbol{\eta}_{\mathrm{A}_{\mathrm{opinion}}}\left[\mathbf{sum}\left(\omega_{\mathrm{A}_1,\mathrm{A}_{\mathrm{opinion}}}\mathrm{A}_1(t), \ldots, \omega_{A150,\mathrm{A}_{\mathrm{opinion}}}\mathrm{A}_1(t)\right) - \mathrm{A}_{\mathrm{opinion}}(t)\right]\Delta t$$
$$\mathbf{d}\mathrm{A}_{\mathrm{opinion}}(t)/\mathbf{d}t = \boldsymbol{\eta}_{\mathrm{A}_{\mathrm{opinion}}}[\mathbf{sum}(\omega_{\mathrm{A}_1,\mathrm{A}_{\mathrm{opinion}}}\mathrm{A}_1(t), \ldots, \omega_{A150,\mathrm{A}_{\mathrm{opinion}}}\mathrm{A}150(t)) - \mathrm{A}_{\mathrm{opinion}}(t)]$$

For the voting state they are

$$\mathrm{A}_{\mathrm{voting}}(t+\Delta t) = \mathrm{A}_{\mathrm{voting}}(t) + \boldsymbol{\eta}_{\mathrm{A}_{\mathrm{voting}}}[\omega_{\mathrm{A}_{\mathrm{opinion}},\mathrm{A}_{\mathrm{voting}}}\mathrm{A}_{\mathrm{opinion}}(t) - \mathrm{A}_{\mathrm{voting}}(t)]\Delta t$$
$$\mathbf{d}\mathrm{A}_{\mathrm{voting}}(t)/\mathbf{d}t = \boldsymbol{\eta}_{\mathrm{A}_{\mathrm{voting}}}[\omega_{\mathrm{A}_{\mathrm{opinion}},\mathrm{A}_{\mathrm{voting}}}\mathrm{A}_{\mathrm{opinion}}(t) - \mathrm{A}_{\mathrm{voting}}(t)]$$

3 Social Network Analysis

Gephi version 0.9.2. [3] has been used to analyse the social network, which is shown in Fig. 2. The network consists of 152 nodes and 7059 edges. The average degree is 46.441, with a highest in-degree of 150 for the Opinion node. Within the agents, node 36 has the highest in-degree of 108. This means this agent has 108 incoming connections. Using the notion of centrality of nodes within a network, the relative importance of a node in the network can be determined. In this social network the centrality of the node would represent the popularity of the agent on the web. Node 5 has the highest betweenness centrality of 2152.369, after which node 141 follows with 1998.363, node 14 with 1832.571 and node 78 with 1224.405. High values for betweenness centrality are most likely popular agents who influence the other connecting agents the most. The number of communities in this network is 7 and the modularity is 0.83. The Opinion and Vote node are communities on itself. There are also 4 nodes which have very few to zero in-degrees, each forming a separate community. All other nodes are part of the remaining community.

For gathering *real world data* on the popularity of Baudet on the web, all search results on google.com with 'thierry baudet' as search term were analyzed and classified manually. The web pages analyzed were of all types. The majority concerned web pages of news agencies and blogs where people could give their opinion.

Based on the content of the web page an overall sentiment of that particular page was given: positive, neutral or negative. Evaluation of sentiment was based on signal words and the positive or negative connection that was given to those words. For example in the sentence: "He was brave to say such words in such a public." gives a positive sentiment based on the word 'brave'. In total, 920 web pages dating from January 2017 to December 2017 were analyzed for positive, neutral or negative sentiment. Potential classification errors could have had influence on the sentiment analysis, but as the analysis was done by two persons in accordance with each other, this risk was minimised. Therefore, potential errors are expected not to affect the overall course of the sentiment. Furthermore, poll outcomes from the monthly political poll of Maurice de Hond in 2017 were gathered to analyze voting behaviour.

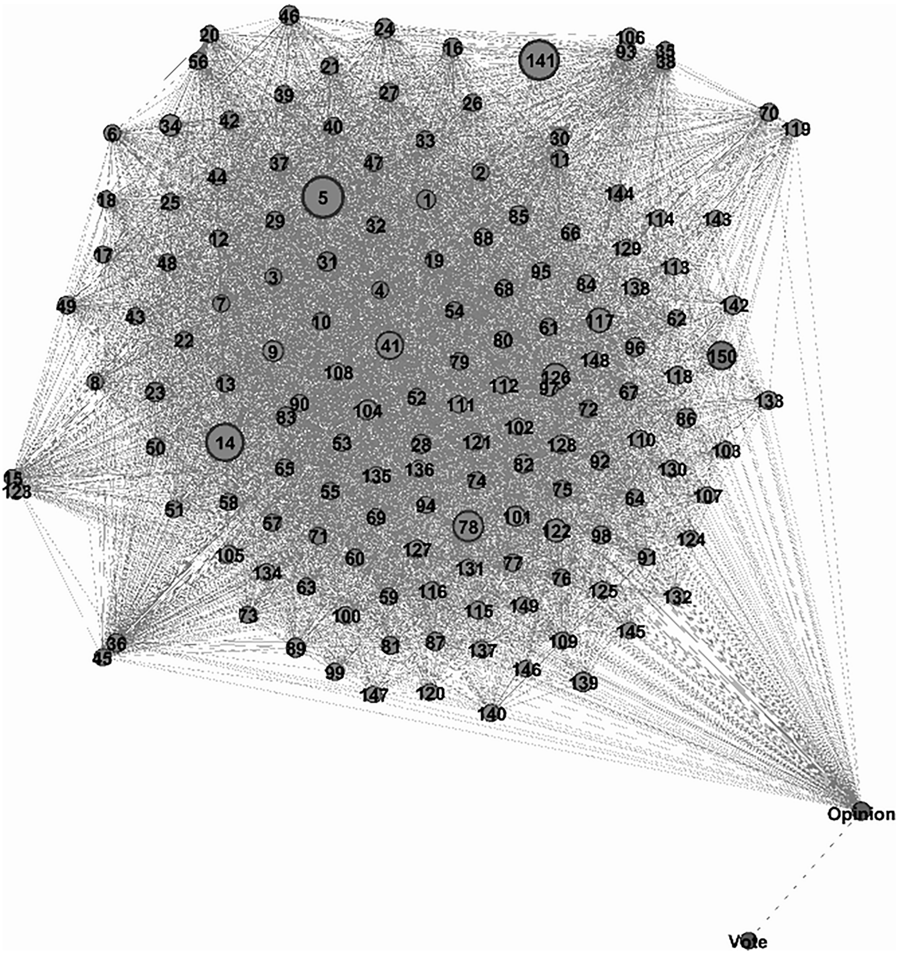

Fig. 2. Nodes size according to their betweenness centrality. The bigger the node, the higher their betweenness centrality value.

4 Validation Based on Real World Data

Matlab v2017 was used to simulate the numerical representation of the model. As the model uses continuous state values between 0 and 1, the empirical opinion data was aggregated by the percentage of positive mentions relative to total amount of mentions in each month. As the maximum amount of seats in the House of Representatives in The Netherlands is 150, the empirical vote data was aggregated by dividing it by 150. The initial value of the Opinion state was 0.22 and of the Vote state 0.01, derived from the month of January 2017 of the empirical data. For the assignment of the initial values of the agents, the initial value of the opinion state is taken into account. It is assumed that, initially, 22 percent of all agents had a positive opinion on Baudet. Therefore a random value above the average opinion is assigned to these agents. The remaining 78 percent of the agents were assigned a random value below the average opinion. A step size Δt of 0.1 was used, with a maximum t of 12. To compare the empirical data against the model data, 11 time points were chosen. The opinion and vote values at time point 2 to 12 of the empirical data are compared with time points 11, 21, 31, 41, 51, 61, 71, 81, 91, 101 and 111 of the Opinion and Vote state of the model.

To make the model fitting with the real world data, a large number of parameters has been tuned by Particle Swarm Optimization (PSO) [13]. The speed factors of all agents, the initial values of the Opinion state and Vote state, the connection weights between all agents and the assignment of initial values to the agents were tuned; see Table 1.

Table 1. Parameters tuned by optimization algorithm

Parameter	Notation	Interval
Connections between all agents	$\omega_{i,j}$	[0, 1]
Speed factors for all agents	η_i	[0, 1]
Speed factor for opinion state	$\eta_{opinion}$	[0, 1]
Speed factor for vote state	η_{vote}	[0, 1]
Assignment of initial values	150*	[0, 1]*

* the 150 initial values for the agents were first selected and during the tuning process allocated to the 150 agents

This makes a total of 22802 parameters that were tuned by the Particle Swarm Optimization algorithm.

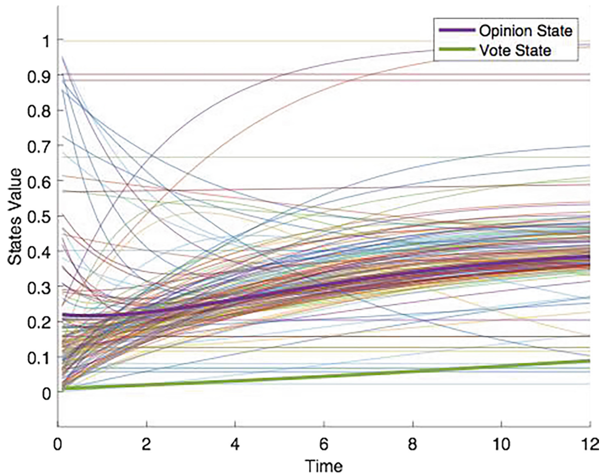

Fig. 3. Simulation results after tuning. The lines not indicated in the legend represent the opinion towards Thierry Baudet of the 150 agents in the model.

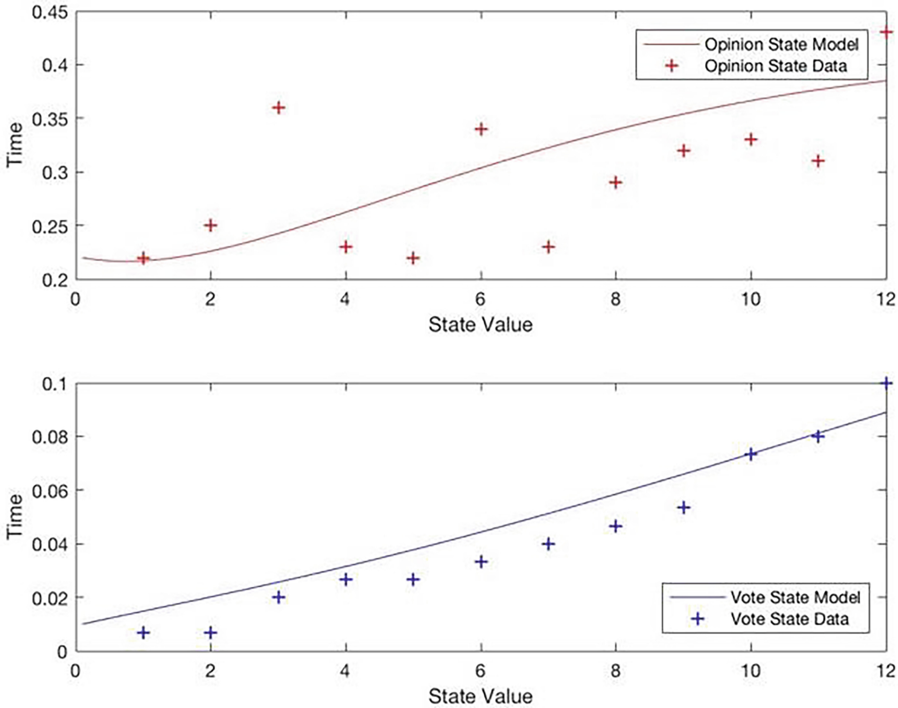

Fig. 4. Opinion and Vote State: model and empirical data values.

As fitness function for the Particle Swarm Optimization algorithm the sum of squares SSR of the differences between empirical and model data was used. The goal of the algorithm was to minimize this difference by tuning the parameters. After tuning, the sum of squares SSR was 7.257×10^{-5}. This gives an average deviation of $\sqrt{(SSR/11)} = 0.0026$. The simulation after tuning is shown in Fig. 3. Figure 4 shows the values of the Opinion and Vote state of the model and empirical data.

5 Verification by Mathematical Analysis

In this section, it is discussed how a Mathematical Analysis of the behavior of the model was performed. Doing this, the model was verified by comparing simulation outcomes to analysis results. Moreover, more in depth understanding of the outcomes was obtained. A formula has been derived for the trend of the sum of all agent state values, in the form of the derivative of that sum; see Theorem 1. The following Lemma 1 can easily be verified.

Lemma 1. For a scaled sum combination function $\mathbf{ssum}_\lambda(..)$ with scaling factor $\lambda = \omega_A = \omega_{B_1,A} + \ldots + \omega_{B_k,A}$ the outcome of the function is a weighted average of the incoming state values:

$$\mathbf{ssum}_{\omega_A}\left(\omega_{B_1,A}B_1(t),\ \ldots,\ \omega_{B_k,A}B_k(t)\right) = \left(\omega_{B_1,A}/\omega_A\right)B_1(t) + \ldots + \left(\omega_{B_k,A}/\omega_A\right)B_k(t)$$

where the sum of the weights $\omega_{Bi,A}\ /\ \omega_A$ for the state values is 1.

Theorem 1. In a network based on the scaled sum combination function with scaling factor $\lambda = \omega_X$ the following holds:

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_B \left[\sum_{A,A\neq B} (\eta_A\omega_{B,A}/\omega_A - \eta_B\omega_{A,B}/\omega_B\right]B(t)$$

and

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_B \left[\sum_{A,A\neq B} \left(\eta_A\omega_{B,A}/\omega_A\right) - \eta_B\right]B(t)$$

with ω_X the sum of all $\omega_{C,X}$ for C not equal to X.

Proof. This requires some algebraic rewriting. First replace each connection weight $\omega_{B,A}$ by $\omega_{B,A}\ /\ \omega_A$ so that it results in $\sum_{B\neq A} \omega_{B,A} = 1$. Then rewrite as follows:

$$\begin{aligned}
\mathbf{d}\sum_A A(t)/\mathbf{d}t &= \sum_A \mathbf{d}A(t)/\mathbf{d}t \\
&= \sum_A \eta_A\left[\sum_{B\neq A} \omega_{B,A}B(t) - A(t)\right] \\
&= \sum_A \eta_A\left[\sum_{B\neq A} \omega_{B,A}B(t) - \sum_{B\neq A} \omega_{B,A}A(t)\right] \\
&= \sum_A \eta_A \sum_{B\neq A} \omega_{B,A}[B(t) - A(t)\ ] \\
&= \sum_{A,B,B\neq A} \eta_A\omega_{B,A}[B(t) - A(t)]
\end{aligned}$$

Thus it is obtained

$$\begin{aligned}
\mathbf{d}\sum_A A(t)/\mathbf{d}t &= \sum_{A,B,B\neq A} \eta_A\omega_{B,A}[B(t) - A(t)] \\
\mathbf{d}\sum_B B(t)/\mathbf{d}t &= \sum_{A,B,B\neq A} \eta_B\omega_{A,B}[A(t) - B(t)]
\end{aligned}$$

Add these two to obtain

$$\begin{aligned}
&\mathbf{2\,d}\sum_A A(t)/\mathbf{d}t = \sum_{A,B,B\neq A} \eta_A\omega_{B,A}[B(t) - A(t)] + \sum_{A,B,B\neq A} \eta_B\omega_{A,B}[A(t) - B(t)] \\
&= \sum_{A,B,B\neq A} \eta_A\omega_{B,A}[B(t) - A(t)] + \eta_B\omega_{A,B}[A(t) - B(t)] \\
&= \sum_{A,B,B\neq A} \left[\eta_A\omega_{B,A}\,B(t) - \eta_A\omega_{B,A}A(t) + \eta_B\omega_{A,B}A(t) - \eta_B\omega_{A,B}B(t)\right] \\
&= \sum_{A,B,B\neq A} \left[(\eta_A\omega_{B,A} - \eta_B\omega_{A,B})B(t) - (\eta_A\omega_{B,A} - \eta_B\omega_{A,B})A(t)\right] \\
&= \sum_{A,B,B\neq A} \left[(\eta_A\omega_{B,A} - \eta_B\omega_{A,B})B(t)\right] - \sum_{A,B,B\neq A} \left[(\eta_A\omega_{B,A} - \eta_B\omega_{A,B})A(t)\right] \\
&= \sum_{A,B,B\neq A} \left[(\eta_A\omega_{B,A} - \eta_B\omega_{A,B})B(t)\right] - \sum_{A,B,B\neq A} \left[(\eta_B\omega_{A,B} - \eta_A\omega_{B,A})B(t)\right] \\
&= \sum_{A,B,B\neq A} (\eta_A\omega_{B,A} - \eta_B\omega_{A,B})B(t) - (\eta_B\omega_{A,B} - \eta_A\omega_{B,A})B(t) \\
&= \sum_{A,B,B\neq A} \left[(\eta_A\omega_{B,A} - \eta_B\omega_{A,B}) - (\eta_B\omega_{A,B} - \eta_A\omega_{B,A})\right]B(t) \\
&= \sum_{A,B,B\neq A} 2(\eta_A\omega_{B,A} - \eta_B\omega_{A,B})\,B(t)
\end{aligned}$$

Therefore it is found

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_{A,B,B\neq A} (\eta_A \omega_{B,A} - \eta_B \omega_{A,B})\, B(t)$$

This can be rewritten into

$$\begin{aligned}\mathbf{d}\sum_A A(t)/\mathbf{d}t &= \sum_B \sum_{A,A\neq B} (\eta_A \omega_{B,A} - \eta_B \omega_{A,B})\ B(t)\\ &= \sum_B \left[\sum_{A,A\neq B} (\eta_A \omega_{B,A} - \eta_B \omega_{A,B})\right] B(t)\\ &= \sum_B \left[\sum_{A,A\neq B} \eta_A \omega_{B,A} - \eta_B \sum_{A,A\neq B} \omega_{A,B}\right] B(t)\\ &= \sum_B \left[\sum_{A,A\neq B} \eta_A \omega_{B,A} - \eta_B\right] B(t)\end{aligned}$$

In terms of the original connection weights this is

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_B \left[\sum_{A,A\neq B} \left(\eta_A \omega_{B,A}/\omega_A - \eta_B \omega_{A,B}/\omega_B\right)\right]\ B(t)$$

or

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_B \left[\sum_{A,A\neq B} \left(\eta_A \omega_{B,A}/\omega_A\right) - \eta_B\right]\ B(t)\ \blacksquare$$

The latter formula explains that the sum increases when the speed factor is low for states with a high state value:

[*B(t)* high $\Rightarrow \eta_B$ low] $\Rightarrow$ the terms in the sum with high *B(t)* have a higher coefficient

$$\Rightarrow \mathbf{d}\sum_A A(t)/\mathbf{d}t > 0$$

Indeed, the outcome of the tuning process shows a tendency of relatively lower speed factors for states with high initial value; see Fig. 5. The general idea behind this is that due to slower changing of nodes with higher values, their impact over time on the whole population will be stronger than the impact of nodes that change relatively fast due to which they soon adapt to the other nodes and then lose their influence.

Based on the connection weights, and speed factors generated by the optimization algorithm, and the simulated state values of the model, the equation

$$\mathbf{d}\sum_A A(t)/\mathbf{d}t = \sum_B \left[\sum_{A\ ,A\neq B} \left(\eta_A \omega_{B,A}/\omega_A\right) - \eta_B\right] B(t)$$

from Theorem 1 has been checked. In Fig. 5 the outcome is shown. It turns out that the standard error of the estimate √(SSR/N) is 0.0140, which is good; this is such a small difference that it is hardly visible in Fig. 6. This contributes verification outcomes for both the formula and the implementation of the model and provides reliable evidence for the implemented model to do what is expected.

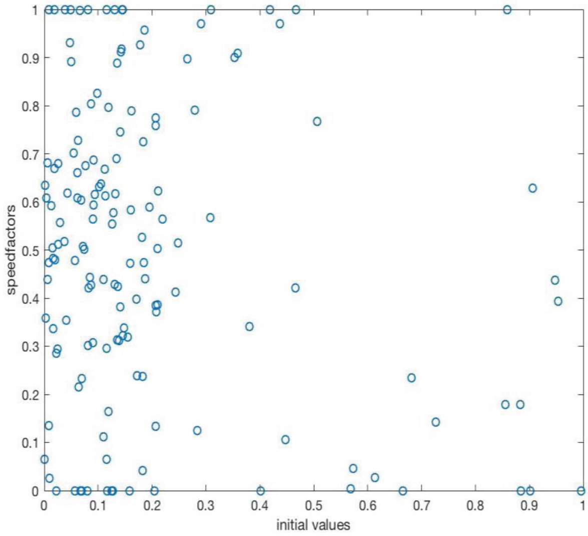

Fig. 5. Speed factors versus initial state values after tuning

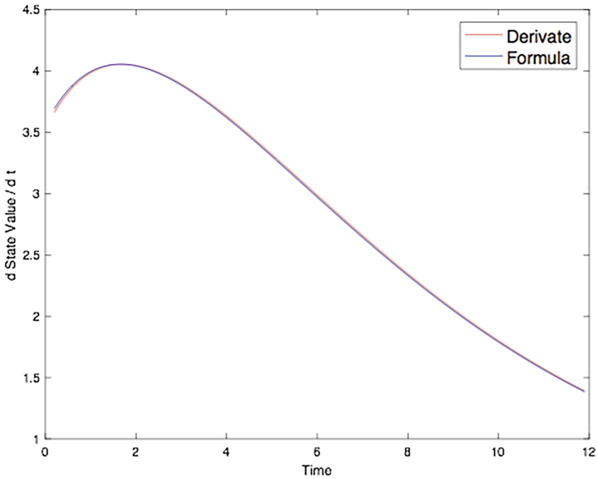

Fig. 6. The outcome of **d** $\sum_A A(t)$ / **d**t in red and $\sum_B [\sum_{A,A\neq B} (\eta_A\ \omega_{B,A}/\omega_A) - \eta_B]\ B(t)$ in blue (Color figure online)

6 Mathematical Analysis of Equilibria

When simulated for longer time periods, the model reaches an equilibrium state. In such an equilibrium state, the state values usually turn out equal. This has been analysed mathematically as well.

Definition

(a) The network model is in *equilibrium* at t if $\mathbf{d}Y(t)/\mathbf{d}t = 0$ for all states.
(b) The network is called *weakly symmetric* if for all nodes A and B it holds $\omega_{A,B} = 0 \Leftrightarrow \omega_{B,A} = 0$ or, equivalently: $\omega_{A,B} > 0 \Leftrightarrow \omega_{B,A} > 0$.
(c) The network is called *fully symmetric* if $\omega_{A,B} = \omega_{B,A}$ for all nodes A and B.

Lemma 2. Let a network be given based on the scaled sum combination function with scaling factor $\lambda = \omega_A = \omega_{B1,A} + \ldots + \omega_{Bk,A}$, then the following hold:

a) If for some node A at time t for all nodes B with $B(t) > A(t)$ it holds $\omega_{B,A} = 0$, then $A(t)$ is decreasing at t: $\mathbf{d}A(t) / \mathbf{d}t \leq 0$.
b) If, moreover, a node B exists with $B(t) < A(t)$ and $\omega_{B,A} > 0$, then $A(t)$ is strictly decreasing at t: $\mathbf{d}A(t)(t) / \mathbf{d}t < 0$.

Proof: (a) Using Lemma 1, from the expressions for $\mathbf{c}_A(\ldots)$ it follows that $\mathbf{c}_A(\ldots) \leq A(t)$, and therefore $\mathbf{d}A(t)/\mathbf{d}t \leq 0$, so $A(t)$ is decreasing at t. b) In this case $\mathbf{c}_A(\ldots) < A(t)$ and therefore $\mathbf{d}A(t)/\mathbf{d}t < 0$, so $A(t)$ is strictly decreasing. ■

Theorem 2. Suppose a weakly symmetric network is based on the scaled sum combination function with scaling factor $\lambda = \omega_X$. Then in an equilibrium state all connected states have the same value.

Proof: Suppose in an equilibrium state at t some nodes A and B exist such that $A(t) \neq B(t)$ and $\omega_{B,A} > 0$ and $\omega_{A,B} > 0$. Take a node A with this property with highest value A. Then for all nodes C with $C(t) > A(t)$ it holds $\omega_{C,A} = 0$, and there exists a B with $B(t) < A(t)$ and $\omega_{B,A} > 0$. Now apply Lemma 2b) to this node A. It follows that $\mathbf{d}A(t)/\mathbf{d}t < 0$, so A is not in equilibrium at t. This contradicts the assumption. Therefore all nodes that are connected have the same state value in this equilibrium. ■

This Theorem 2 explains what is observed in the simulations. Connected states converge to the same value, but isolated states can keep their original value.

7 Discussion

Persons within a social network, for example, family or co-workers, influence each other [11]. Also, some persons within a network tend to influence others more then other persons, those so-called ‘opinion leaders’ can have a strong influence on the opinions within those network [9]. This paper presented a computational network model to analyse the popularity of a politician within a population that participates on the Internet. The approach enabled to identify the characteristics of agents in the network based on the analysis of public opinions on that politician found on the Internet. The model was designed as a scale-free agent network [2], according to the Network-Oriented Modeling approach presented in [16]. The model was tuned to the aggregated public opinion, using a Particle Swarm Optimization algorithm [13]. The model was verified by social network analysis and by mathematical analysis.

There are studies that model the outcome of elections based on behavior of people on social media [4, 10, 14], but as far as the authors know not how in a social network an opinion towards a specific politician changes over time. Therefore there is no comparison available for the proposed model that does this based on the sentiment on the web.

Often in a model for the social contagion principle, all the state values converge without showing an upward or downward general trend. However, it was found that by using low speed factors for states with a high value, the mean of all states first show an upward trend and after a longer time converge to a value substantially above an expected mean of all initialization state values. This phenomenon was also verified and explained by the mathematical analysis. Apparently, as in most of the literature the speed factors for all agents are set equal (or there even is not such a concept in the model able to make differences in speed), this phenomenon is not often showing up. But from an agent-based modeling perspective it is quite natural to consider personalized agent characteristics, in particular for the speed factors.

Connecting the probability to vote to the overall public opinion by a sum function may be a bit too simplistic view on voting behavior, for example, when a certain politician is popular in a group, this does not mean other politicians would not be even more popular. In future research variations of the model should include more in-depth modelling on voting behavior based on the overall public opinion dynamics.

This paper showed that dynamics of an overall opinion of a network can be modelled, and that in a population where the overall opinion is negative, positive agents may be less influenced or are less likely to change their opinion than negative agents.

References

1. Acemoglu, D., Ozdaglar, A.: Opinion dynamics and learning in social networks. Dyn. Games Appl. **1**(1), 3–49 (2011)
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
3. Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: ICWSM, vol. 8, pp. 361–362 (2009)
4. DiGrazia, J., McKelvey, K., Bollen, J., Rojas, F.: More tweets, more votes: Social media as a quantitative indicator of political behavior. PLoS ONE **8**(11), e79449 (2013)
5. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence models, analysis, and simulation. J. Artif. Soc. Soc. Simul. **5**(3) (2002)
6. Kim, J.: Philosophy of Mind. Westview Press, Boulder (1996)
7. Kuipers, B.J.: Commonsense reasoning about causality: deriving behavior from structure. Artif. Intell. **24**, 169–203 (1984)

8. Kuipers, B.J., Kassirer, J.P.: How to discover a knowledge representation for causal reasoning by studying an expert physician. In: Proceedings of the 8th International Joint Conference on Artificial Intelligence, IJCAI 1983, William Kaufman, Los Altos, CA, pp. 49–56 (1983)
9. Lazarsfeld, P.F., Gaudet, H., Berelson, B.: The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign. Columbia University Press, New York (1965)
10. Metaxas, P.T., Mustafaraj, E., Gayo-Avello, D.: How (not) to predict elections. In: Proceedings of the Third International Conference on Social Computing, SocialCom 2011, pp. 165–171. IEEE (2011)
11. Nickerson, D.W.: Is voting contagious? evidence from two field experiments. Am. Polit. Sci. Rev. **102**(01), 49–57 (2008)
12. Pearl, J.: Causality. Cambridge University Press, New York (2000)
13. Poli, R., Kennedy, J., Blackwell, T.: Particle swarm optimization. An overview. Swarm Intell. **1**, 33–57 (2007)
14. Sang, E.T.K., Bos, J.: Predicting the 2011 Dutch senate election results with twitter. In: Proceedings of the workshop on semantic analysis in social media, pp. 53–60. Association for Computational Linguistics (2012)
15. Treur, J.: Dynamic modeling based on a temporal–causal network modeling approach. Biologically Inspired Cogn. Architectures **16**, 131–168 (2016)
16. Treur, J.: Network-Oriented Modelling: Addressing Complexity of Cognitive, Affective and Social Interactions. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45213-5
17. Watts, D.J., Dodds, P.S.: Influentials, Networks, and Public Opinion Formation. J. Consum. Res. **34**(4), 441–458 (2007)