Expert Systems With Applications 44 (2016) 187–197

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Information diffusion through social networks: The case of an online petition

Mohammad S. Jalali[a,b,*], Armin Ashouri[b], Oscar Herrera-Restrepo[b], Hui Zhang[c]

[a] Sloan School of Management, Massachusetts Institute of Technology, 30 Memorial Dr, Cambridge, MA 02142, USA
[b] Grado Department of Industrial and Systems Engineering, Virginia Tech, 7054 Haycock Rd, Falls Church, VA 22304, USA
[c] Grado Department of Industrial and Systems Engineering, Virginia Tech, 250 Durham Hall, Blacksburg, VA 24061, USA

ARTICLE INFO

Keywords:
Diffusion process
Online social networks
Petition
System dynamics modeling

ABSTRACT

People regularly use online social networks due to their convenience, efficiency, and significant broadcasting power for sharing information. However, the diffusion of information in online social networks is a complex and dynamic process. In this research, we used a case study to examine the diffusion process of an online petition. The spread of petitions in social networks raises various theoretical and practical questions: What is the diffusion rate? What actions can initiators take to speed up the diffusion rate? How does the behavior of sharing between friends influence the diffusion process? How does the number of signatures change over time? In order to address these questions, we used system dynamics modeling to specify and quantify the core mechanisms of petition diffusion online; based on empirical data, we then estimated the resulting dynamic model. The modeling approach provides potential practical insights for those interested in designing petitions and collecting signatures. Model testing and calibration approaches (including the use of empirical methods such as maximum-likelihood estimation, the Akaike information criterion, and likelihood ratio tests) provide additional potential practices for dynamic modelers. Our analysis provides information on the relative strength of push (i.e., sending announcements) and pull (i.e., sharing by signatories) processes and insights about awareness, interest, sharing, reminders, and forgetting mechanisms. Comparing push and pull processes, we found that diffusion is largely a pull process rather than a push process. Moreover, comparing different scenarios, we found that targeting the right population is a potential driver in spreading information (i.e., getting more signatures), such that small investments in targeting the appropriate people have 'disproportionate' effects in increasing the total number of signatures. The model is fully documented for further development and replications.

1. Introduction

Petitions are often formal letters or documents submitted to government organizations or public entities to convey requests on certain issues. Petitions represent the attitudes or opinions of petition initiators, as well as people who sign them. The initiators usually want to receive as many signatures as possible to raise awareness and maximize the impact of petitions. Traditional petitions collect handwritten signatures, but with the rise of the Internet and digital communications, online petitioning has become widespread. People regularly use email and social networks as platforms to launch their petitions due to their convenience, efficiency, and significant broadcasting power. For example, Care2, which was initiated in 1998, is a large petition website covering a broad spectrum of topics, such as animal rights, environment, politics, and human rights. Petitions on this site had received more than 373 million signatures as of September 2015. Many other sites provide similar services. People who want to launch petitions can follow simple processes to set up free online petitions and collect signatures. People interested in signing petitions visit the petition's webpage, fill in basic personal information, and submit the form. There is also an optional choice of sharing petitions with others through online social networks.

The spread of petitions in social networks raises various theoretical and practical questions: What is the diffusion rate? What actions can initiators take to speed up the diffusion rate? How does the behavior of sharing between friends influence the diffusion process? How does the number of signatures change over time? In order to address these questions, the mechanisms of the petition diffusion process need to be investigated and understood. In this research, we used system dynamics modeling to specify and quantify the core mechanisms of petition diffusion online. We used a case study to specify

* Corresponding author at: Sloan School of Management, Massachusetts Institute of Technology, 30 Memorial Dr, Cambridge, MA 02142, USA. Tel.: +1 617 253 2954.
E-mail addresses: jalali@mit.edu (M.S. Jalali), ashouri@vt.edu (A. Ashouri), oscar84@vt.edu (O. Herrera-Restrepo), corinnaz@vt.edu (H. Zhang).

http://dx.doi.org/10.1016/j.eswa.2015.09.014
0957-4174/

Table 1
Sharing, reminder, and forgetting variables.

Factor	Variables	Reference
Sharing	Satisfaction, length of relationship, novelty	De Bruyn and Lilien (2008)
Reminders	Familiarity, complexity, novelty	Tellis (1997)
Forgetting	Transience, absent-mindedness, misattribution, bias, socialization, interference	Percy (2004), Reitman (1971), Struben (2004)

the diffusion process of a petition, and based on empirical data, estimated the resulting dynamic model. Comparing different scenarios provided additional insights into the pragmatics of similar diffusion processes.

This article is organized as follows: Section 2 reviews the theoretical foundations of diffusion models; Section 3 presents the data and methods, including the case study and the developed model; Section 4 presents analysis, including model testing and calibration, comparison of our model with the Bass diffusion model, sensitivity analysis of the estimated model to different scenarios, and discussion on our modeling approach. The study is concluded in Section 5.

2. Theoretical foundations

2.1. Diffusion models

The phenomenon of diffusion has been widely studied due to its potential impact in various fields, such as epidemiology (Raj, Kuceyeski, & Weiner, 2012), marketing (Kim, Lee, & Ahn, 2006), and social behavior (Susarla, Oh, & Tan, 2012). Understanding contagion phenomena, product/service adoption and changing cultural features all depend on how people influence each other, which is the hallmark of diffusion models (Rogers, 2003). Better understanding can also enable decision makers to design policies that maximize benefits, minimize risks, and provide control over time. A basic diffusion process requires two main actors and one binding element. The transmitter (also called adopter and infectious) and receiver (also called potential adopter and susceptible) constitute the main actors, and various communication and contact channels establish the means to link them (Baran, 2010). Since the early 20th century (Kermack & Mckendrick, 1932), a vast literature has used such models to understand, predict, and control epidemics. In marketing applications, word of mouth (WOM) and advertisement are the key channels bridging early adopters and potential adopters of new products and services (Bass, 1969; Rogers, 1976). Various studies have looked at the spread of ideas and norms through diffusion processes (Centola, 2010).

The spread of online petitions on the Internet also follows a similar mechanism. Instead of purchasing products, signing petitions is the key choice that people may make in response to messages they receive from others. Potential adopters are defined as potential signatories, and the rate of diffusion can be measured by the number of new signatures per unit of time. Peer-to-peer word of mouth is one channel for spread of petitions. Parallel to the concept of advertising, first announcements followed by reminders may be considered, besides peer-to-peer mechanisms of spread.

2.2. Diffusion models in online social networks

Online social networks such as Facebook, Twitter and Google Plus are now integrated into people's daily lives. This magnifies the importance of studying mechanisms that affect the spread of information through these networks. Here we review diffusion models in two streams of research: marketing literature and expert and intelligent systems literature.

A large literature on marketing of products and services focuses on the effect of different factors, such as word of mouth, reminders, and forgetting on diffusion processes. These topics are mainly addressed from the business and psychological points of view (De Bruyn & Lilien, 2008; Percy, 2004; Tellis, 1997). De Bruyn and Lilien (2008) highlighted awareness and interest as the key factors in a diffusion process. Both of these factors influence the chances of adopting a product, service, or as in the case of this study, signing a petition, and as such should be further elaborated on. Tellis (1997) and Percy (2004) studied the influence of reminders and forgetting on awareness. Bentley and Earls (2008) discussed the nature of the diffusion process, whether it is a push process (a top-down form of diffusion) or a pull process (a bottom-up form of diffusion). De Bruyn and Lilien (2008) also proposed usefulness, trust, and customization as core elements of level of interest. Table 1 presents some of the relevant factors related to sharing, reminders, and forgetting. We benefit from these factors in our model.

A growing body of research in expert and intelligent systems also concentrates on information diffusion, which covers a wide range of theoretical and practical contributions. Here, we provide a brief overview of some of the studies and then discuss our contributions.

Li et al. studied the efficiency of information diffusion under information overload on Facebook-like (Li & Sun, 2014) and Twitter-like (Li, Li, Wang, & Zhang, 2014) social networks based on the structure of the network and user behavior. They proposed a metric to measure information diffusion efficiency and analyzed its values on simulated social networks with different characteristics. Yang et al. (2015) studied the effects of users' social roles on information diffusion through social networks. They proposed a role-aware information diffusion model, which can be used to predict whether a user will repost a specific message at the micro-level and the scale of a diffusion process at the macro-level. Taxidou and Fischer (2014b) introduced a system for real-time analysis of information diffusion on Twitter. They also analyzed information diffusion on Twitter based on social graphs (star-shaped vs. complex) and types of influence (Taxidou & Fischer, 2014a). Cheng, Adamic, Dow, Kleinberg, and Leskovec (2014) defined temporal and structural features of posts as key predictors of cascade size in information diffusion. Their findings showed that the breadth, rather than depth, of a cascade is a better indicator of large cascades. Network's degree, PageRank and k-core were also studied as other cascade size predictors by Pei, Muchnik, Andrade, Zheng, and Makse (2014). They found k-core to be the only factor influencing information spread on social networks. Liu and Zhang (2014) proposed a dynamic susceptible-infected-recovered (SIR) model for information diffusion through social networks in which individuals can break links and reconnect to their second-order friends. Their proposed strategy increases the speed at which information spreads on social networks. Kim, Newth, and Christen (2014) analyzed behavioral patterns of news diffusion through mainstream news websites, social networks, and blogs in terms of activity, reactivity, and heterogeneity. They found that mainstream news websites are the most active, social networks are the most reactive, and blogs are the most persistent. Li, Qian, Jin, Hui, and Vasilakos (2015) studied the efficiency of information diffusion on social networks of microblogs by studying 10 million user profiles from Sina Weibo (a Chinese microblog) and 41.7 million profiles from Twitter. Liu, Xie, Hu, and Chen (2014) explored the effects of affinity of information with people on information cascade size. They also discussed the effects of affinity, average degree of the network and the probability of people losing their interest in the information on the size of information diffusion.

In our study, we considered push and pull processes from the marketing literature such that sending announcements to a target population subjected that population to the petition (a push process); while sharing of the petition by those who had signed it increased awareness from the bottom up (a pull process). The success of a petition depends on the strength of both mechanisms. We particularly explored the effect of sharing and reminding on the diffusion rate. To do so, we built a detailed dynamic model of the petition diffusion phenomenon and estimated it empirically. Our analysis provides information on the relative strength of push and pull processes and insights about awareness, interest, sharing, reminder, and forgetting mechanisms. The modeling approach also provides potential practical insights for those interested not only in designing and collecting signatures on petitions, but also spreading information using social networks. We used our calibrated model to assess the impact of potential strategies that petition organizers could take, such as reminders, to promote their project or campaign. Our specific contributions to the current body of the literature on expert and intelligent systems include: applying the system dynamics modeling approach; considering feedback mechanisms and delays in our model; focusing on marketing concepts (such as pull and push processes); and adding the forgetting mechanism to our diffusion model. We further discuss this in Section 4.5.

3. Data and methods

3.1. Case data

We secured access to the anonymous empirical dataset of an online petition. In more than two months, 4683 people signed this petition in response to a campaign organized by a concerned group. The relatively small size of the campaign allowed us to collect data on all the key announcements related to this petition across both email and Facebook platforms, reducing the common challenge of unobserved communication pathways in diffusion modeling. We called the petition organizers the "core group." The core group created a petition and broadcast it in order to get the support of the public by collecting signatures. The three channels used to promote the petition were: (1) word of mouth, (2) email, and (3) posting on a social network (Facebook in our study). Physical communication was negligible in this setting, and we had data from the core group's email and Facebook posts. The core group started the diffusion process by sending out a number of announcements. The initial target population, chosen by the core group, was the group's close connections and friends deemed potentially interested in the topic.

After people received announcements from the core group via email or a shared post on Facebook, they could respond in three ways: (1) being interested and signing the petition, (2) being uninterested and disregarding it or (3) being interested but not signing it and forgetting about it (they might sign it later). The factors that influence the choice of the three alternatives may include personal preferences on the petition topic, time availability, and relationship with the message sender. Moreover, a word-of-mouth process is also viable, as some of signatories may share it with their friends and encourage them to sign it.

During the diffusion process, one of the policy levers the core group could use was to send out reminders to people who had signed the petition, asking them to share the information with their friends, regardless of whether they had already shared the petition or not.

3.2. Model structure

We first discuss the key processes incorporated in the model and then present the formulations. To spread the word on a petition, the core group sends out initial announcements to an unaware target population, increasing the number of people with a pending message in their email inbox or on their Facebook page. Upon opening the message, this group flows into one of three subpopulations: people who sign the petition on first receipt, interested but forgetful people, and uninterested people. Two parameters, "interested fraction" and "forget rate," define what percentage of people choose each alternative. A fraction of people who sign might also share the petition with their contacts, increasing the exposure of the petition. Note that in contrast to traditional diffusion models, where adopters continue to influence potential adopters, we found it more accurate to use the adoption flow as the signal for sharing: few people save a petition email for long and send it to their friends repeatedly; instead, they often make a quick judgment at the time of signing on whether to share. At the same time, based on total signatures collected, the core group may decide to send reminders to a portion of the signatories, which increases the chance of sharing. An unaware population exposed to sharing becomes the next potential signatories.

Eventually, two streams of people, i.e., people who sign on first receipt and people who sign through sharing or reminders, contribute to total signatures. People who forget are returned to the unaware people pool. However, we assumed that the process ends for uninterested people.

Therefore, the model is presented in three major subsystems: (1) unaware population, (2) aware population, and (3) sharing and reminding subsystem. Fig. 1 provides an overview of the model structure. The equations for these subsystems are presented in Section 3.3. We fully documented the model following a set of minimum reporting requirements (Rahmandad & Sterman, 2012)—additional files, including data and details on replicating the model, optimization, and analyses are provided in the online supplementary materials

3.3. Model equations

3.3.1. Unaware population subsystem

This subsystem includes the process through which the unaware target population moves to the pool of people with a pending message. The unaware target population (U) receives announcements via: (1) email or shared posts on Facebook (initial announcement rate, A) from the core group and (2) through the sharing of the petition (rate of sharing, R_s) by those who have already signed the petition. Equations for this subsystem are presented in more detail in Table 2.

3.3.2. Aware population subsystem

This subsystem includes people who respond to a pending message. As discussed earlier, people who have a pending message are divided into three categories: (1) uninterested people, (2) interested but forgetful people, and (3) people who sign on first receipt (chance). The first group ignores the message through the *ignoring message* rate (I_m), the second group forgets about the petition through the *forgetting message* rate (f_m), and the third group signs the petition right after reading it through the *rate of signing on first chance* (R_c). The interested but forgetful subpopulation has a chance of signing the petition after receiving the shared information about the petition, represented in the sharing and reminding subsystem. Equations for the aware population subsystem are presented in Table 3 in more detail.

3.3.3. Sharing and reminder subsystem

This subsystem includes sharing the petition with the unaware population and sending reminders to the aware population. The petition is shared by two groups, signatories who share the petition right after signing it and signatories who share the petition after receiving reminders. People in this subsystem are eventually signatories who either share the petition or do not share it. Detailed equations for this subsystem are presented in Table 4.

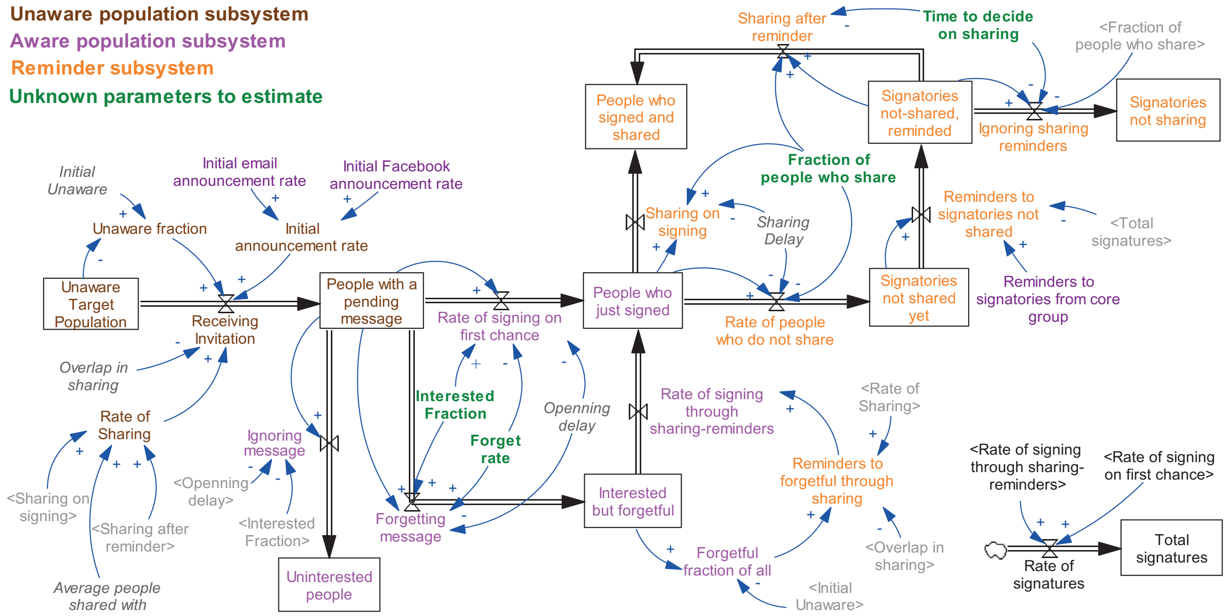

Fig. 1. Model structure.

Table 2
Equations for the unaware population subsystem.

Variable/parameter	Equation	Unit
Unaware target population	$U = \int_{t=0}^{T} -R\,dt + U_0$ Total population (initially equal to U_0) who could become aware of the petition by receiving invitations (R).	People
Receiving invitation	$R\begin{cases} A + w * R_s * (1-\nu), & U > 0 \\ 0, & U = 0 \end{cases}$ The rate at which people from the *unaware target population* are moved to *people with a pending message*.	People/day
Initial announcement rate	$A = A_f + A_e$ *Initial Facebook announcement rate* (A_f) and *initial email announcement rate* (A_e) are extracted from empirical data—see Fig. 2C and Fig. 2D, respectively.	People/day
Unaware fraction	$w = U/U_0$ The fraction of *unaware target population* to the initial *unaware target population*	Dmnl
Rate of sharing	$R_s = \mu * (S_{ar} + S_{os})$ The rate at which the unaware population becomes aware of the petition. It depends on *average people shared with* (μ), *sharing on signing* (S_{os}) (those who share right after signing the petition, see Table 4) and *sharing after reminder* (S_{ar}) (those who share after getting a reminder, see Table 4).	People/day
People with a pending message	$P_m = \int_{t=0}^{T} (R - f_m - R_c - I_m)dt$ Population who received the announcement email or a post on Facebook but have not read it yet. A fraction of this population will ignore the message (*ignoring message* (I_m), see Table 3), the rest will sign the petition right after reading it (*rate of signing on first chance* (R_c), see Table 3) or forget about the petition (*forgetting message* (f_m), see Table 3).	People
Initial unaware	$U_0 = 20,000$ U_0 is a constant representing the audience pool, initial unaware population, assumed to be 20,000[a].	People
Overlap in sharing	$\nu = 0.16$ On social networks such as Facebook, it is common for two persons' friend lists have some overlap. Hence, the overlapping fraction of friends, calculated from the empirical data, is deducted from *rate of awareness* using *overlap in sharing to* prevent the same person from being deducted twice from the *unaware population*.	Dmnl
Average people shared with	$\mu = 460$ Average number of friend connections who become aware of the petition through sharing.	Dmnl

[a] *Initial unaware* variable(U_0) and *interested fraction* (I_f) are both unknown and together form the size of the target population. The total number of signatories (as an input of the objective function in the estimation process) is correlated to the size of the target population (and not to (U_0) and I_f; therefore U_0 and I_f cannot be estimated simultaneously. We then assumed parameter U_0 as a constant and estimated parameter I_f Based on the data, at the end of the petition deadline (day 71), the total number of people who were aware of the petition was roughly 15,000. Given that some unaware people remained in the population on day 71, U_0 has to be larger than 15,000—we assumed $U_0 = 20,000$.

4. Analysis

4.1. Data

We used anonymous data from an online petition spread through Facebook and email. This empirical data accounts for 71 days between November 1, 2011 and January 11, 2012. The data includes the total number of daily signatures, the initial person who shared the petition and the pattern of sharing among friends. Also, the data contains the first announcing rates by email and Facebook as well as reminders sent by the petition's core group. Fig. 2 presents this data.

4.2. Model testing and calibration

We first conducted structure and behavior validity tests (Barlas, 1996). We then tested the formulations against different input values to ensure that they represented logical behaviors and were robust under extreme conditions (Morecroft, 1985; Sterman, 2000 Chap. 21) to build confidence in the model[1]. We then calibrated the model to the empirical data using maximum-likelihood estimation (MLE), which

[1] Simulation runs were executed in Vensim DSS™.

Table 3
Equations for the aware population subsystem.

Variable/parameter	Equation	Unit
Rate of signing on first chance	$R_c = (1 - f_r) * I_f * \frac{P_m}{d_o}$ *Rate of signing on first chance* is a fraction of *people with a pending message* (P_m) who are interested in the petition (*interested fraction,* (I_f) and will not forget about it (*1-forget rate,* f_r)). They read the message with some delay (*opening delay,* d_o).	People/day
Forgetting message	$f_m = f_r * I_f * \frac{P_m}{d_o}$ *Forgetting message* is the percentage of aware population who are interested (in signing the petition) but forget about the message. It is proportional to those who receive the message (*people with a pending message,* P_m), are interested to the petition (*interested fraction,* I_f) and read the message with some delay (d_o) but forget it (*forget rate,* f_r).	People/day
Interested but forgetful	$I_g = \int_{t=0}^{T} (f_m - R_r)dt$ Total population who are aware of the petition and are interested in signing it but forgot about it.	People
Rate of signing through sharing or reminders	$R_r = M_{fs}$ Rate of people who forgot about the petition but will sign it through sharing by other signatories (M_{fs}).	People/day
People who just signed	$P_s = \int_{t=0}^{T} (R_s + R_r - R_{ns} - E - S_{os})dt$ People who signed the petition but did not decide to share it.	People
Uninterested people	$Z = \int_{t=0}^{T} I_m dt$ People who are not interested.	People
Ignoring message	$I_m = P_m/d_o * (1 - I_f)$ *People with a pending message* (P_m) who are not interested in the petition (1-*interested fraction,* I_f). They read the message with some delay (*opening delay,* d_o).	People/day
Forgetful fraction of all	$f_a = I_g/U_0$ Fraction of the people who forget about the message to the initial unaware population.	Dmnl*
Opening delay	$d_o = 1$ Delay in opening the message. This is the assumed average delay.	Day
Interested fraction	I_f: *unknown parameter* Portion of people interested in the petition.	Dmnl
Forget rate	f_r: *unknown parameter* Portion of people interested in the petition but forget about it.	Dmnl

* Dimensionless

Table 4
Equations for the sharing and reminding subsystem.

Variable/parameter	Equation	Unit
Signatories not shared yet	$G_n = \int_{t=0}^{T} (R_{ns} - M_{sn})dt$ People who signed the petition but have not shared it yet; they may share the petition after receiving reminder (M_{sn}).	People
Rate of people who do not share	$R_{ns} = P_s * (1 - c)/d_s$ Fraction of people who signed the petition but do not share it. It is proportional to *people who just signed* (P_s) and *fraction of people who share* (c). d_s is assumed to be 1.	People/day
Reminders to signatories not shared yet	$M_{sn} = \begin{cases} M_{sc} * G_n/S & S > 0 \\ 0 & S = 0 \end{cases}$ Fraction of *signatories not shared yet* (G_n) who are reminded, proportional to *reminders to signatories from core* group (M_{sc}) and *total signatures* (S).	People/day
Signatories not-shared, reminded	$G_{nsr} = \int_{t=0}^{T} (M_{sn} - S_{ar} - I_{sr})dt$ Signatories who do not share the petition and are reminded to share it.	People
Sharing after reminder	$S_{ar} = c * G_{nsr}/t_s$ Fraction of *signatories not-shared, reminded* (G_{nsr}) who share. It is proportional to *fraction of people who share* (c), considering some delay for deciding (*time to decide on sharing,* t_s).	People/day
Ignoring sharing and reminders	$I_{sr} = (1 - c) * G_{nsr}/t_s$ Fraction of *signatories not-shared, reminded* (G_{nsr}) who do not share (proportional to *fraction of people who share,* c) even after receiving reminders—considering some delay for deciding (*time to decide on sharing,* t_s).	People/day
Signatories not sharing	$G_{ns} = \int_{t=0}^{T} (I_{sr})dt$ Signatories who never share the petition.	People
Sharing on signing	$S_{os} = P_s * c/d_s$ Fraction of *people who just signed* (P_s) and share the petition right after signing it. These people do not need reminders for sharing. It is proportional to *fraction of people who share,* c).	People/day
People who signed and shared	$P_{ss} = \int_{t=0}^{T} (S_{os} - S_{ar})dt$ Total signatories who shared the petition right after signing it or after receiving a reminder.	People
Reminders to forgetful through sharing	$M_{fs} = R_s * (1 - \nu) * f_a$ Fraction of people who forget about the petition and are reminded through Facebook sharing.	People/day
Rate of signatures	$R_{sg} = R_r + R_s$ Total rate of signatures including *rate of signing through sharing and reminders* (R_r) and *rate of signing on first chance* (R_s).	People/day
Total signatures	$S = \int_{t=0}^{T} (R_{sg})dt$ Total number of signatories.	People
Time to decide on sharing	t_s: *unknown parameter* Average delay in reading the reminder and deciding to share the petition.	Day
Fraction of people who share	c: *unknown parameter* Fraction of signatories who share the petition.	Dmnl

maximizes the likelihood of observing the data given the model parameters. Because of its parametric nature, MLE requires the definition of a likelihood function. Considering the behavior of the rate of signatures over time, we defined the likelihood function assuming a Poisson distribution for the daily number of signatures; when the likelihood function is intractable or very complicated, the method of simulated moments (MSM) is recommended instead of MLE (Jalali, Rahmandad, & Ghoddusi, 2013, 2015 Chap. 2). The Poisson distribution depicts the likelihood of a certain number of events occurring in a time and/or space interval (Johnson, Kotz, & Kemp, 1992). The events occur at an average rate (the only parameter of the distribution) and without dependency on the time of the previous event. Given the daily rate of signatures and assuming that the process of signing is random, the Poisson distribution allows us to know how likely it is to observe a daily number of signatures, conditional on an underlying average signature rate. The likelihood function based on the Poisson distribution is $\mathcal{L}(x_1, x_2, \ldots, x_n|\lambda) = \frac{\lambda^{\sum_1^n x_i} e^{-n\lambda}}{x_1! \ldots\ldots x_n!}$. Consequently, the log-likelihood function is $\ln(\mathcal{L}) = -n\lambda + (ln\lambda)\sum_1^n x_i - \ln(\prod_1^n x_i\,!)$. Therefore, the objective function[2] would be:

$$\text{Max} \ln(\mathcal{L}) = -nR_{sg} + (\ln(R_{sg}))\sum_1^n x_i - \ln\left(\prod_1^n x_i\,!\right)$$
where
$$n = 71\,(\text{number of observed data}---\text{daily signatories})$$
$$0 \le I_f,\ f_r,\ \text{and}\ c \le 1$$
$$0 \le t_s$$

[2] The functional form used in Vensim DSS is: $L = \ln(f) = -71 * rate\ of\ signatures + \ln(rate\ of\ signatures) * real\ data - GAMMA\ LN\ (real\ data + 1)$

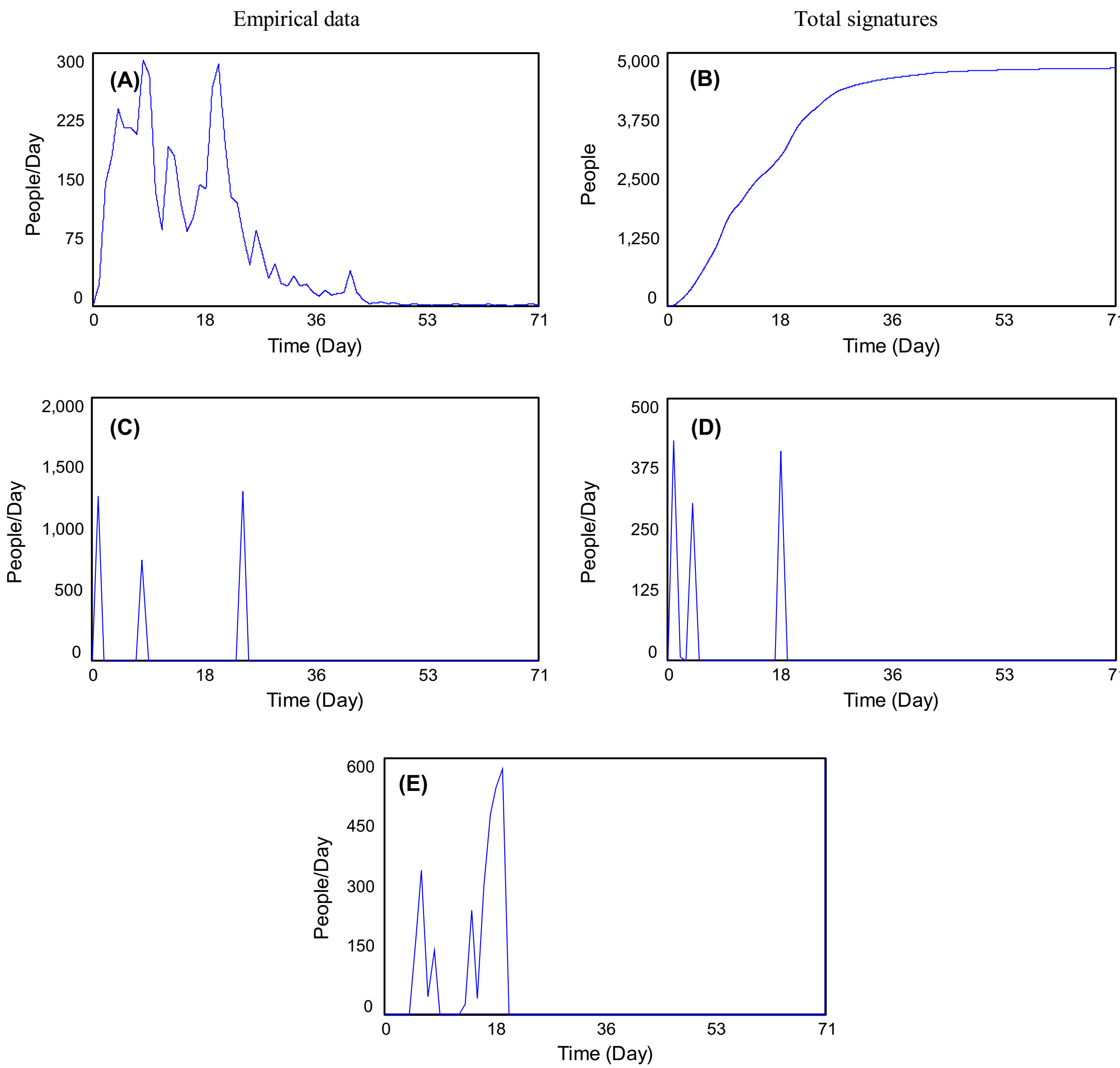

Fig. 2. Empirical data: (A) rate of signatures; (B) accumulated signatures; (C) initial Facebook announcement rate; (D) initial email announcement rate; (E) Reminders to signatories by core group.

The model included eight parameters. Among these, four parameters were associated with directly observable physical processes and an approximate value for them was available (See Tables 1 and 2). This left us with four parameters that needed to be adopted from the literature, estimated statistically, or calculated through other methods. *Interested fraction* (I_f) divided the population (i.e., those who were aware of the petition) into two groups, those who were interested in the petition and those who were not. *Forget rate* (f_r) separated those who signed the petition (immediately) from those in the interested group who forgot to sign. *Time to decide on sharing* (t_s) is the delay between receiving a reminder and taking action and deciding whether to share. *Fraction of people who share* (c) is the ratio of signatories who shared the petition with their friends. Struben (2004) used a value of 0.5 for f_r, which we can use as a candidate value. Parameter c can also be estimated from empirical data to be 0.0071[3]. Yet given the significant differences between our model and Struben's, and uncertainty in the estimate of c due to the small sample size for empirical data, we were potentially interested in statistically estimating these two parameters as well.

We note that MLE favors more complex models with more free parameters, which might cause overfitting. To control for this potential problem, we performed a tradeoff analysis (between the goodness of fit of the model and the complexity of the model due to the number of parameters) using the Akaike information criterion (AIC) (Akaike, 1974) and relative likelihood (Burnham, Anderson, & Burnham, 2002). The AIC penalizes the number of free parameters (n) such that $AIC = 2n - 2\ln(f)$ where f is the likelihood function, and lower values of AIC are preferred. Table 5 presents a comparison of four estimated models, including statistical estimation of different combinations of the free parameters. The adjusted coefficients of determination ($\bar{R}^2$) for models 1 and 2 are very close, but based on AIC, model 1 is preferred with the minimum AIC (886), implying that it minimizes the information loss among the four models. Based on AIC,

[3] Parameter c was estimated from the empirical data, which contained the ratio of people who became aware of the petition, those who signed and those who shared the petition, their friends as well as their mutual friends on Facebook. However, given the uncertainty in the estimation of c due to the small sample size for empirical data, we estimated it in the model.

Table 5
Model comparison based on Akaike information criterion.

Model	Unknown parameters to estimate	Number of unknown parameters	$\bar{R}^2$	ln(L)*	AIC**	Relative likelihood
Model 1	I_f, f_r, t_s, c	4	0.860	−439	886	1
Model 2	I_f, t_s, c	3	0.856	−451	908	1.5E−05
Model 3	I_f, f_r, t_s	3	0.807	−568	1143	1.5E−56
Model 4	I_f, t_s	2	0.805	−597	1198	1.8E−68

* L is the likelihood function.
** The model with minimum AIC is preferred.

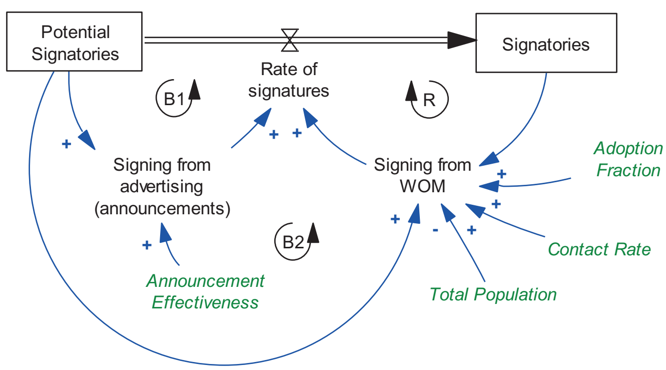

Fig. 3. Bass diffusion model, adapted from Sterman (2000).

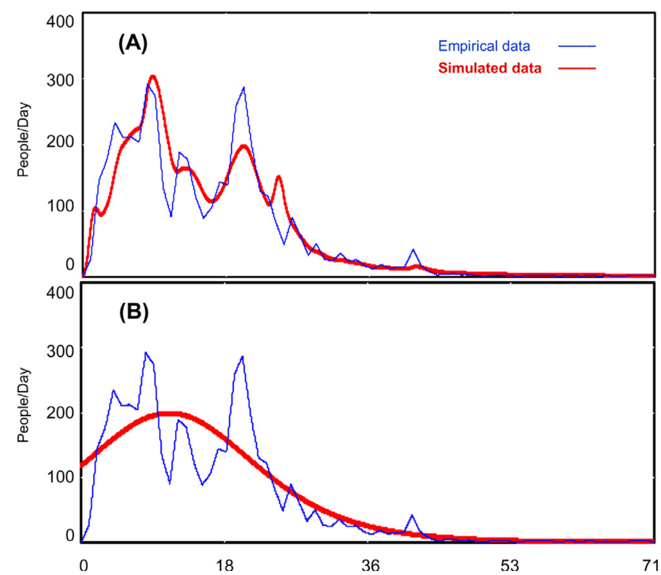

Fig. 4. Simulated data vs. empirical data; (A) signatures per day in model 1; (B) signatures per day in the Bass model.

model 2 would be the next model to consider; however, based on a relative likelihood measure ($\exp(AIC_{min} - AIC_i)/2$), where $i = 1,\ldots,4$), model 2 is 1.5E-5 times as probable as model 1 for minimizing information loss. This loss is due to having one fewer parameter than model 1. If this difference was close to 1, we could consider model 2 as a viable alternative, but given the significant loss of information, model 1 is the only preferred model among the four.

4.3. Comparison with the Bass diffusion model

We also compared model 1 with the Bass diffusion model (Bass, 1969), which is a popular model for the diffusion of innovations. The Bass model was originally used for forecasting sales of new products. The basic idea in the Bass model is that potential adopters become aware of innovations through external information. To operationalize the Bass model, two independent mechanisms affect the adoption rate: Positive feedback, which is usually considered to be word of mouth (internal influence); and external information, considered to be advertising (external influence). When an innovation is initially introduced, the adoption rate is affected only by external influence, advertising; then it is changed by both advertising and word of mouth. Even though the model works well in various examples, it has several restrictive limitations. Both word of mouth and advertising increase adoption without any delays. In many different examples, including our case study, people do not always adopt the innovation immediately (e.g., not all people sign the petition on the first chance). In our model, we considered the situation in which people may forget about the petition and sign it after some delay (i.e., after receiving a reminder) or not sign it at all. Besides that, the Bass model assumes a homogenous population such that there is no heterogeneity among people in adopting the innovation and everyone will eventually adopt it; innovation discard is also not considered (Sterman, 2000). However, in our model, not everyone has to be a signer. Fig. 3 presents the Bass diffusion model (for an online petition case); see Chapter 9 in Sterman (2000) for more discussion about the model.

We calibrated four constants of the proposed Bass model (green parameters in Fig. 3): *adoption fraction, contact rate, total population* and *announcement effectiveness* (the advertising effectiveness), based on our empirical data. Given that everyone in the Bass model would eventually sign the petition, the total population would be considered as equal to the total number of empirical signatures; hence, we only estimated the other three parameters. Table 6 presents the comparison between model 1 and the Bass model. Based on AIC and relative likelihood, it is implied that model 1 potentially minimized information loss compared to the Bass model.

Fig. 4 shows best fit for simulated and empirical data (signatures per day) in model 1 (Fig. 4A) and in the Bass model (Fig. 4B). Although the fit of the Bass model seems to capture the overall increases and decreases in the number of signatures, that is due to the unrealistic assumption that the whole population (as the potential signatories) will eventually sign the petition, i.e., the total population is equal to the total number of signatures from the empirical data. On the other hand, even with an assumed population size of 20,000, model 1 better captures increases and decreases in the number of daily signatures, due to considering other realistic underlying mechanisms (such as the forgetting mechanism) in the model.

4.4. Confidence intervals and sensitivity analysis

Statistical parameter estimation may include various errors. Struben, Sterman, and Keith (2015) indicate that errors in parameter estimation can be due to sampling errors, measurement errors, violations of hypotheses, or specification errors (when models are wrong in ways that matter). We expected that due to the use of electronic data in our study, measurement error would not be a major concern. To partially address the risks of hypotheses and specification errors, we performed behavior reproduction (Barlas, 1989), unit consistency and extreme condition tests (Sterman, 2000)—extreme conditions such as: if no one was interested in the petition, the total number of signatures should remain zero; or if everybody who

Table 6
Model 1 vs. Bass model.

Model	Unknown parameters to estimate	Number of unknown parameters	$\bar{R}^2$	ln(L)*	AIC**	Relative likelihood
Model 1	I_f, f_r, t_s, c	4	0.860	−439	886	1
Bass	*adoption fraction, contact rate,* and *announcement effectiveness*	3	0.759	−597	1199	8.1E−69

* L is the likelihood function.
** The model with minimum AIC is preferred.

received an announcement was fully interested in the petition and no one forgot to sign, the total number of signatures should not exceed the assumed initial unaware population. These tests provided some confidence in the model, but generalization beyond similar petition diffusion settings is not guaranteed. The estimation of confidence intervals helps in addressing sampling errors; any dataset provides a limited sample from the real phenomena. So assuming that the data measurements and the model specification are correct (Struben et al., 2015), we can use confidence intervals to bound the "true" value of the underlying parameters.

We used the likelihood ratio (LR) method to estimate confidence intervals for estimated parameters (Venzon & Moolgavkar, 1988). The likelihood ratio method compares the likelihood of the estimated parameters $\hat{\theta}$ to that of an alternative set $\theta*$ and assumes that $-2(\ln(\mathcal{L}(\hat{\theta})) - \ln(\mathcal{L}(\theta^*)))$ follows a chi-square distribution (χ^2_{df}) where the degree of freedom, *df*, is the number of independent parameters involved in the confidence region estimation (Struben et al., 2015). In our case, we calculated the summation of log-likelihood values over time with optimized parameters as $\ln(\mathcal{L}(\hat{\theta}))$, then changed (increased and decreased) the value of each parameter in consecutive simulations and calculated summation of the log-likelihood ($\ln(\mathcal{L}(\theta^*))$) until it hit $\ln(\mathcal{L}(\hat{\theta})) - \frac{1}{2}\chi^2_{1,95\%}$. This process can be done through Monte Carlo simulations by generating uniform random numbers within a small neighborhood of parameters $\hat{\theta}$. Table 7 presents the estimated parameters as well as the 95% confidence intervals around the parameters to estimate their uncertainty. The confidence intervals were tight in all cases, adding a layer of confidence to the reliability of the estimated values.

Based on the estimated parameters, around 27% of the total population (i.e., assumed 20,000) was interested in the petition, from which 40% signed the petition (immediately) after becoming aware and 60% forgot about it. Among those signatories, only 2% shared the petition with their friends, either right after signing the petition or after being reminded by the core group. For the latter group, it usually took around 15 days to read the reminder and share the petition.

In addition to estimating confidence intervals, sensitivity analysis is often used to study the sensitivity of model outputs to uncertainties in model parameters (Moxnes, 2005). We conducted sensitivity analysis of the model outputs (total number of signatures) by changing the estimated parameters. Given that forget rate (f_r) and time to decide on sharing (t_s) were not directly under the control of the petition designers, we were more interested in studying how the uncertainty in the number of signatures depended on the interested fraction (I_f) and fraction of people who share (c). Parameter I_f directly relates to the choice of target population and the design of the petition. The more appealing a petition is and the better it matches the population, the higher the chances of getting signatures from the target population. Parameter c can also be changed by various interventions of the core group to encourage sharing. In our setting, over a quarter of targeted individuals signed the petition; however, only 2% shared it, showing a major gap between the interest and sharing mechanisms. We explored how different values for these two parameters changed the diffusion dynamics. Fig. 5 presents the changes in the total number of signatures based on these two parameters. We ran 5000 different simulation scenarios based on variations in c and I_f. These parameters were randomly drawn from a uniform distribution [0,1], and

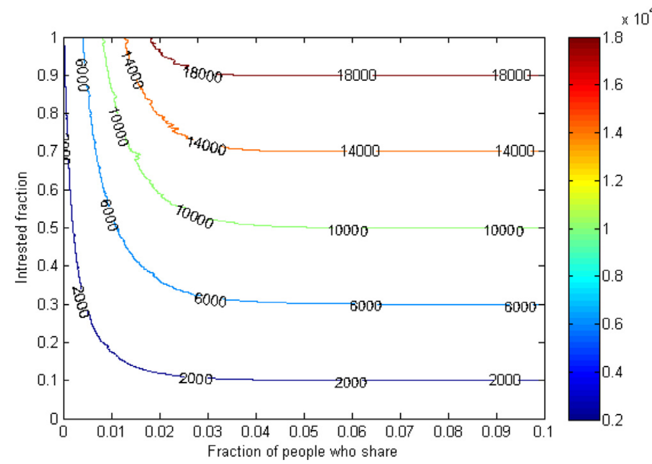

Fig. 5. Total number of signatures based on *fraction of people who share* (c) and interested fraction (I_f).

Latine Hypercube sampling (McKay, Beckman, & Conover, 1979) was used to ensure that the entire range of each parameter was explored.

Color coding presents the total number of signatures. Five examples are illustrated based on the model assumption that the total population is 20,000.

Example: at $I_f = 0.6$ and $c = 0.02$, the model predicts the total number of signatures to be around 10,000—color coded as green.

Fig. 5 shows that I_f was a major driver in increasing the number of signatures. If the interested fraction was low, such that the targeted population was not highly interested in the petition (e.g., $I_{fr} < 0.1$), with increases in the fraction of people who shared the petition, the total number of signatures would not exceed 2000 signatures (10% of the total possible signatures). However, when the right population was targeted, the total number of signatures potentially increased; e.g., with $I_{f_r} = 0.3$, a small increase in the fraction of people who shared the petition would bring the total number of signatures to 6000 (30% of the total possible signatures). This emphasizes the importance of targeting the right population, rather than sharing strategies after the petition is sent out.

We were also interested in the sensitivity of total signatures to the pushing process (through sending announcements) and pulling process (through sharing the petition). The model was estimated based on empirical announcement rates (where total first announcements by email included 1126 individuals and total first announcements through Facebook targeted 3302 people—see Fig. 2C and D). We changed the announcement rates without changing their timing, i.e., scaling the three announcements on Facebook and the three by email. Therefore, we defined a new fraction (announcement magnitude), which changed the total number of first announcements from zero to the maximum of 20,000 (assumed population in the model) and matched the empirical pattern at a value of 0.22. Our analysis showed that the fraction of people who shared the petition played a more important role in increasing the total signatures than the number of announcements. Let us review two scenarios. If the total number of announcements (4428) and the fraction of people who share

Table 7
Estimated model parameters.

Parameter	Unit	Estimate	95% Confidence interval	
			Lower bound	Upper bound
Forget rate (f_r)	Dmnl*	0.5997	0.5710	0.6278
Fraction of people who share (c)	Dmnl	0.0229	0.0219	0.0238
Interested fraction (I_f)	Dmnl	0.2675	0.2606	0.2744
Time to decide on sharing (t_s)	Day	15.3418	13.2380	16.9980

* Dimensionless

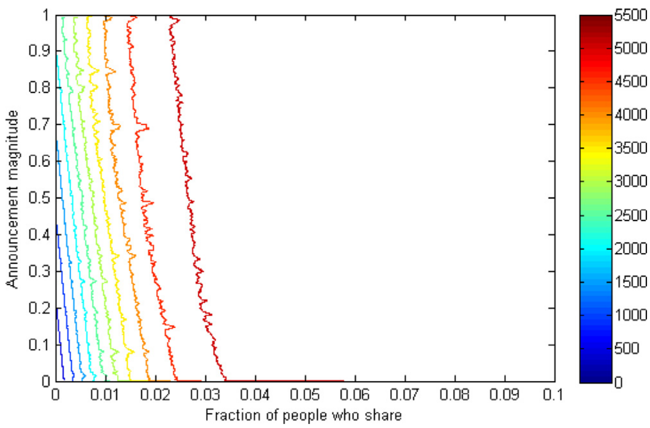

Fig. 6. Total number of signatures based on *fraction of people who share* (c) and announcement magnitude (this magnitude changes the first announcement rates).

(0.023) do not change, the total number of signatures is 4685. Changing the total number of first announcements to 20,000 (announcement magnitude of 1), the total number of signatures increases to 5014. That is, even reaching every possible person in the population (which requires enormous resouces and in practice is most likely impossible), the total signatures increase by less than 10%. On the other hand, by implementing interventions to increase the fraction of people who share the petition slightly from 0.023 to 0.029, the total number of signaures is 5014. Fig. 6 presents comparisons of announcements vs. sharing, where the total possible signatures can be reached by a small increase in the fraction of people who share.

Ten different examples are illustrated in which the color coding represents the total number of signatures. An announcement magnitude of “1” means that every person in the population is reached by the announcements. The figure shows that a small increase in the proportion of people who share (on the scale of 0.01 or even smaller) has a huge effect on getting more signatures.

4.5. Modeling discussion

Besides the comparison against the Bass model in Section 4.3, here we discuss the advantages and limitations of our modeling approach in a general sense. Compared to several studies in the literature (e.g., Bazgana, Chopin, Nichterlein, & Sikora, 2014; Cao, Wua, Wang, & Hub, 2011; Kempe, Kleinberg, & Tardos, 2003; Kim, Newth, & Christen, 2013; López-Pintado, 2008; Saito, Kimura, Ohara, & Motoda, 2013), our approach looks at social networks from an aggregate perspective rather than detailed network nodes and their connectivity. This allows us to study the overall information diffusion process based on mechanisms which govern its behavior. To some extent, this also helps us model individual processes that shape the information diffusion process. Some examples of these processes are: the ability (1) to become aware of an online petition, (2) to forget sharing an online petition, and (3) to share a petition with other people. In studies dealing with detailed networks and computational issues, the way the mechanisms above affect and shape the diffusion process are often overlooked or oversimplified through probabilistic functions or constant values. Instead, the interest is in maximizing the diffusion of information in the network through the identification of influential nodes exhibiting high degrees of connectivity (e.g., more friends) within and between multiple social networks (e.g., Kempe et al., 2003; Kim et al., 2013). The problem then overlooks the need to include and understand the individual decision-making mechanisms undertaken by the nodes to spread information. Furthermore, the purpose of numerous detailed-network and computation-oriented studies is to increase the velocity and accuracy of the algorithms used to identify influential nodes (e.g., Cao et al., 2011; Kempe et al., 2003; López-Pintado, 2008). In terms of the simulation techniques used to study the diffusion phenomenon, the current literature includes various techniques, such as system dynamics (e.g., Maier & Milling, 2009), agent-based modeling (e.g., Cao et al., 2011), and analytical modeling (e.g., Kempe et al., 2003). Both agent-based and analytical modeling allow the study of social networks composed of individual nodes (e.g., people); system dynamics groups those nodes into aggregate “packs.” Both disaggregate and aggregate approaches have advantages and limitations. For example, in using system dynamics we focused on the mechanisms driving the behaviors of spreading information in social networks; however, we treated people through “packed” groups where each person has a similar level of importance. These “packed” groups with equal individual importance are not considered in agent-based (e.g., Cao et al., 2011) or analytical modeling (e.g., López-Pintado, 2008) studies. Thus, our use of system dynamics provides a complementary point of view for studying information diffusion in social networks. Meanwhile, it is also important to highlight how our modeling work contributes to the study of diffusion processes in system dynamics literature. Most system dynamics studies on diffusion focus on innovation. The diffusion of innovations relates to the velocity at which new/substitution products/services are adopted over time. Innovations may also include ideas, information, and technologies, but the majority of the surveyed studies focus on the first rather than the other two. Studies by Maier (1998) and Mehdi and Kunsch (2008), among others, provide a sense of the prediction and structural purpose pursued in most of the available studies. The prediction aims to increase the capacity to foresee the future adoption of, for example, new products. To achieve more valid predictions, the model parameters are defined based on expert knowledge and/or calibration of historical data. The structural purpose relates to the exploration and understanding of the mechanisms that govern the behavior of the innovation adoption. In this regard, our research approach does not differ from the studies above. However, instead of prediction, our approach aimed to achieve greater resemblance, in trend and magnitude, between the data resulting from our proposed model and the empirical data from the social network selected as case study. In doing so, we applied MLE to maximize the likelihood of observing the case study data under certain parameters entering our model, which allowed achievement of a potential resemblance between the simulated data and the empirical data. This also provided us with a better understanding of the ‘push’ and ‘pull’ dynamics governing the information diffusion process. Examples of the use of MLE in information diffusion studies include Myers and Leskovec (2010), Saito et al. (2013) and Gomez-Rodriguez, Leskovec, and Krause (2012)—these authors use MLE in disaggregate approaches such as agent-based and analytical modeling.

5. Conclusion

In this study, we presented a dynamic model to quantify the core mechanisms of petition diffusion, including interest, awareness, forgetting, sharing and reminding. Based on our modeling and case study, we compared different strategies for increasing the number of signatures in an online petition, e.g., increasing the fraction of people who share (pull process) vs. sending announcements (push process), and targeting the right population. Comparing the push (sending announcements) and pull (sharing by the signatories) processes, we found that spread is largely a pull process rather than a push process, which is consistent with the findings of Bentley and Earls (2008). In our analysis, we observed that even if everybody is reached by announcements—which is practically impossible—it might not be as effective as small increases in the fraction of people who share. It should be noted that implementing interventions to increase the fraction of people who share information is more practical, e.g., through incentives, than reaching more people by announcements, which requires extensive resources. Moreover, we discuss targeting the right population as a potential driver in getting more signatures. If the petition initiators do not choose the right population, such that people with low interest in the petition are targeted, sending more announcements and intervening to increase the sharing rate by the signatories (which are potentially resource-intensive strategies) might not be very effective. On the other hand, if the right population is targeted, small investments to increase the sharing rate will have ‘disproportionate’ effects in increasing the total number of signatures.

This study contributes to the expert and intelligent systems literature by modeling the diffusion process of an online petition using a system dynamics approach. Compared with prior studies in the literature that focused on network structure (e.g., Luo, Du, Liu, Xuan, & Wang, 2015), social media user influence (e.g., Li et al., 2014), and friend recommendations (e.g., Chen, Zeng, & Yuan, 2013), this study emphasized: (1) the behavior and strategies involved in online petition diffusion, such as population targeting, petition sharing, and sending reminders; and (2) the aggregated perspective of the social network system. This study also incorporated statistical analysis to provide quantifiable evidence supporting the strategic recommendations, i.e., targeting the right population and increasing sharing rate through incentives. Comparisons were made between the proposed model and other diffusion models and simulation approaches; and similarities and differences were discussed.

This study is subject to some limitations. The case study was based on a small sample of data. We also assumed that the total population was fixed over time. Given that our case study dealt with a small time horizon, i.e., 71 days, the rate of changes in the total population was negligible; however, in studies with longer time horizons, this change should be taken into account. Given our focus on online social networks, we assumed no physical interactions between people. We also assumed that all signatories shared in favor of the petition (favorable word of mouth). Unfavorable word of mouth is common for products and innovations; it can be observed for petitions as well, especially for political ones. Our work also mainly describes the model as a flow process with few but relevant feedback loops.

Future studies could benefit from using larger datasets and further validating our findings. More dynamic factors involved in the diffusion process can also be considered, such as a growing population of online users or an increasing coverage of users’ friend groups over time. Future dynamic modeling studies can also incorporate more feedback loop mechanisms. Moreover, emotions and attitudes towards the petition as well as the effects of social influence (e.g., Jalali (2014)) could also be captured; different attitudes may either facilitate or hinder the diffusion process. It would also be interesting to have a panel of several heterogeneous social networks (e.g., Facebook, Twitter, Google Plus, etc.) and study the similarities and differences of petition diffusion patterns on various platforms.

Acknowledgments

We would like to express our sincere gratitude to Hazhir Rahmandad, who encouraged us to pursue the topic and provided us with great feedback. We also thank participants at the 2015 System Dynamics conference for their comments and suggestions.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eswa.2015.09.014.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.
Baran, M. (2010). Diffusion of innovations in the system thinking approach. *Management Business Innovation, 6*, 17–24.
Barlas, Y. (1989). Multiple tests for validation of system dynamics type simulation models. *European Journal of Operations Research, 42*, 59–87.
Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. *System Dynamics Review, 12*, 183–210.
Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science, 15*, 215–227.
Bazgana, C., Chopin, M., Nichterlein, A., & Sikora, F. (2014). Parameterized approximability of maximizing the spread of influence in networks. *Journal of Discrete Algorithms, 27*, 54–65.
Bentley, A., & Earls, M. (2008). *Forget influentials, herd-like copying is how brands spread*, 19–22.
Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference : a practical information-theoretic approach* (2nd ed.). New York: Springer.
Cao, T., Wu, X., Wang, S., & Hu, X. (2011). Maximizing influence spread in modular social networks by optimal resource allocation. *Expert Systems with Applications, 38*, 13128–13135.
Centola, D. (2010). The spread of behavior in an online social network experiment. *Science, 329*, 1194–1197.
Chen, L., Zeng, W., & Yuan, Q. (2013). A unified framework for recommending items, groups and friends in social media environment via mutual resource fusion. *Expert Systems with Applications, 40*, 2889–2903.
Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M., & Leskovec, J. (2014). Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web* (pp. 925–936). ACM.
De Bruyn, A., & Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing, 25*, 151–163.
Gomez-Rodriguez, M., Leskovec, J., & Krause, A. (2012). Inferring networks of diffusion and influence, *ACM Transactions on Knowledge Discovery from Data (TKDD): 5* (pp. 1–37).
Jalali, M., Rahmandad, H., & Ghoddusi, H. (2013). Estimation of unknown parameters in system dynamics models using the method of simulated moments. *Proceedings of the Winter Simulation Conference* (pp. 1212–1222). IEEE.
Jalali, M., Rahmandad, H., & Ghoddusi, H. (2015). Using the method of simulated moments for system identification. *Analytical methods for dynamic modelers*. MIT Press.
Jalali, M. S. (2014). How individuals weigh their previous estimates to make a new estimate in the presence or absence of social influence. *Social computing, behavioral-cultural modeling and prediction* (pp. 67–74). Springer.
Johnson, N. L., Kotz, S., & Kemp, A. W. (1992). *Univariate discrete distributions* (2nd ed.). Wiley.
Kempe, D., Kleinberg, J., & Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 137–146). ACM.
Kermack, W. O., & McKendrick, A. G. (1932). Contributions to the mathematical theory of epidemics – II. The problem of endemicity. In *Proceedings of the Royal society of London. Series A: 138* (pp. 55–83).
Kim, J., Lee, D. J., & Ahn, J. (2006). A dynamic competition analysis on the Korean mobile phone market using competitive diffusion model. *Computers & Industrial Engineering, 51*, 174–182.
Kim, M., Newth, D., & Christen, P. (2013). Modeling dynamics of diffusion across heterogeneous social networks: news diffusion in social media. *Entropy, 15*, 4215–4242.
Kim, M., Newth, D., & Christen, P. (2014). Trends of news diffusion in social media based on crowd phenomena. In *Proceedings of the companion publication of the 23rd international conference on World wide web companion* (pp. 753–758). International World Wide Web Conferences Steering Committee.

Li, J., Peng, W., Li, T., Sun, T., Li, Q., & Xu, J. (2014). Social network user influence sense-making and dynamics prediction. *Expert Systems with Applications, 41*, 5115–5124.
Li, P., Li, W., Wang, H., & Zhang, X. (2014). Modeling of information diffusion in Twitter-like social networks under information overload. *The Scientific World Journal, 2014* 8 pages. Article ID 914907.
Li, P., & Sun, Y. (2014). Modeling and performance analysis of information diffusion under information overload in Facebook-like social networks. *International Journal of Communication Systems, 27*, 1268–1288.
Li, Y., Qian, M., Jin, D., Hui, P., & Vasilakos, A. V. (2015). Revealing the efficiency of information diffusion in online social networks of microblog. *Information Sciences, 293*, 383–389.
Liu, C., & Zhang, Z.-K. (2014). Information spreading on dynamic social networks. *Communications in Nonlinear Science and Numerical Simulation, 19*, 896–904.
Liu, H., Xie, Y., Hu, H., & Chen, Z. (2014). Affinity based information diffusion model in social networks. *International Journal of Modern Physics C, 25* 12 pages. Article ID 1440004.
López-Pintado, D. (2008). Diffusion in complex social networks. *Games and Economic Behavior, 62*, 573–590.
Luo, S., Du, Y., Liu, P., Xuan, Z., & Wang, Y. (2015). A study on coevolutionary dynamics of knowledge diffusion and social network structure. *Expert Systems with Applications, 42*, 3619–3633.
Maier, F., & Milling, P. (2009). System dynamics analysis of the diffusion of innovations. *Encyclopedia of complexity and applied systems science* (pp. 1956–1982). Springer.
Maier, F. H. (1998). New product diffusion models in innovation management – a system dynamics perspective. *System Dynamics Review, 14*(14), 285–308.
McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*, 239–245.
Mehdi, E., & Kunsch, P. L. (2008). Simulation of micro-CHP diffusion by means of system dynamics. *Energy Policy, 36*, 2308–2319.
Morecroft, J. D. W. (1985). Rationality in the analysis of behavioral simulation models. *Management Science, 31*, 900–916.
Moxnes, E. (2005). Policy sensitivity analysis: simple versus complex fishery models. *System Dynamics Review, 21*(22), 123–145.
Myers, S. A., & Leskovec, J. (2010). On the convexity of latent social network inference. In *Proceedings of the 24th conference on neural information processing systems (NIPS'10)* (pp. 1–9). NIPS'10.
Pei, S., Muchnik, L., Andrade, J. S., Jr, Zheng, Z., & Makse, H. A. (2014). Searching for superspreaders of information in real-world social media. *Scientific Reports, 4* Article ID 5547.
Percy, L. (2004). Advertising and the seven sins of memory. *International Journal of Advertising, 23*, 413–427.
Rahmandad, H., & Sterman, J. D. (2012). Reporting guidelines for simulation-based research in social sciences. *System Dynamics Review, 28*, 396–411.
Raj, A., Kuceyeski, A., & Weiner, M. (2012). A network diffusion model of disease progression in dementia. *Neuron, 73*, 1204–1215.
Reitman, J. S. (1971). Mechanisms of forgetting in short-term memory. *Cognitive Psychology, 2*, 185–195.
Rogers, E. M. (1976). New product adoption and diffusion. *Journal of Consumer Research, 2*, 290–301.
Rogers, E. M. (2003). *Diffusion of innovations* (5th .ed.). New York: Free Press.
Saito, K., Kimura, M., Ohara, K., & Motoda, H. (2013). Detecting changes in information diffusion patterns over social networks. *ACM Transactions on Intelligent Systems and Technology, 4*(3) 1–23. Article ID 55.
Sterman, J. (2000). *Business dynamics: systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill.
Struben, J. (2004). Technology transitions: identifying challenges for hydrogen vehicles. In *Proceedings of the 22nd international conference of the system dynamics society* (pp. 25–29).
Struben, J., Sterman, J. D., & Keith, D. (2015). Parameter and confidence interval estimation in dynamic models: maximum likelihood and bootstrapping methods. In H. Rahmandad, R. Oliva, & N. Osgood (Eds.), *Analytical methods for dynamic modelers*. MIT Press.
Susarla, A., Oh, J.-H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: evidence from YouTube. *Information Systems Research, 23*, 23–41.
Taxidou, I., & Fischer, P. M. (2014). Online analysis of information diffusion in Twitter. In *Proceedings of the companion publication of the 23rd international conference on World wide web companion* (pp. 1313–1318). International World Wide Web Conferences Steering Committee.
Taxidou, I., & Fischer, P. M. (2014). Rapid: a system for real-time analysis of information diffusion in twitter. In *Proceedings of the 23rd ACM international conference on information and knowledge management* (pp. 2060–2062). ACM.
Tellis, G. J. (1997). Effective frequency: one exposure or three factors? *Journal of Advertising Research*, 75–80.
Venzon, D. J., & Moolgavkar, S. H. (1988). A method for computing profile-likelihood-based confidence intervals. *Applied Statistics, 37*, 87–94.
Yang, Y., Tang, J., Leung, C., Sun, Y., Chen, Q., Li, J., & Yang, Q. (2015). RAIN: social role-aware information diffusion. In *Proceedings of the 29th AAAI conference on artificial intelligence (AAAI'15)*.