Performance Evaluation 99–100 (2016) 1–15

Contents lists available at ScienceDirect

Performance Evaluation

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

Opinion propagation in bounded medium-sized populations

Eline De Cuypere [a], Koen De Turck [b], Sabine Wittevrongel [a], Dieter Fiems [a,*]

[a] Department of Telecommunications and Information Processing, Ghent University, St-Pietersnieuwstraat 41, B-9000 Gent, Belgium

[b] Laboratoire des Signaux & Systèmes, Centrale Supélec, 3, Rue Joliot-Curie, 91192 Gif sur Yvette, France

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form 28 October 2015
Accepted 14 March 2016
Available online 24 March 2016

Keywords:
Opinion propagation
Epidemic model
Maclaurin-series expansion
Fluid model
Markov chain
Performance analysis

ABSTRACT

We study the dynamics of opinion propagation in a medium-sized population with low population turnover. Opinion spreading is modelled by a Markovian non-standard Susceptible–Infected–Recovered (SIR) epidemic model with stochastic arrivals, departures, infections and recoveries. The system performance is evaluated by two complementary approaches: a numerical but approximate solution approach which relies on Maclaurin-series expansions of the stationary solution of the Markov process and a fluid limit approach. Both methods are evaluated numerically. Moreover, convergence to the fluid limit is proved, and explicit expressions for the fixed points of the differential equations are obtained for the case of linearly increasing infection and arrival rates.

1. Introduction

Given the rapid growth of companies in the internet sector that base their revenue model on advertisement (such as Google, Facebook, etc.) [1] and the ascent of social networks in particular, the study of opinion spreading is a trending topic, and there is a strong interest in understanding how new opinions spread through a community. Apart from these economic considerations, the analysis of opinion spreading can improve our comprehension of social relations among individuals, both online and offline.

This paper studies opinion propagation by drawing parallels with the spreading of diseases [2,3]. Indeed, opinion propagation bears some similarity to the spread of an infectious disease, and particularly to Kermack and McKendrick's classical compartmental SIR model for such propagation [4]. The acronym SIR stands for susceptible (S), infectious (I) and recovered (R), and refers to the possible states that an individual can be in, the possible transitions between these states following the order S $\rightarrow$ I $\rightarrow$ R. In particular, the SIR model assumes that if a healthy individual encounters an infected individual, there is a chance that the healthy individual gets infected. An infected individual then recovers from the disease after some time, making him/her immune for the infection. This process can be directly reformulated in terms of the propagation of opinions on a particular topic: a susceptible or non-opinionated individual has yet to form an opinion on the topic, whereas infected or opinionated individuals do have such an opinion. Susceptible individuals may form an opinion when they encounter infected individuals, like an infection is carried over to a healthy individual upon contact. Finally, individuals lose their interest in the topic after some time and stop spreading their opinion. These individuals have recovered.

The SIR epidemic model has been predominantly applied to study disease, see e.g. [3,5–7], but there is some prior work in the areas of rumour and opinion propagation as well. Concerning the latter, Zhao extends the SIR model for opinion

* Corresponding author.
E-mail address: dieter.fiems@UGent.be (D. Fiems).

http://dx.doi.org/10.1016/j.peva.2016.03.001

spreading in social networks by including a hibernator state in which individuals temporary interrupt infecting others [8]. Moreno et al. study SIR like spreading of rumours explicitly accounting for the network structure [9]. Bettencourt et al. [10] draw parallels between epidemics and idea diffusion by applying several epidemiological models to empirical data. Woo et al. [11] show the plausibility to describe the mechanism of violent topic diffusion in web forums by a SIR model while Fan et al. [12] propose an extended SIR model for opinion dynamics in which individuals can have a positive or negative opinion about a topic. Finally, some preliminary results on series expansion techniques (cf. infra) for stochastic SIR models for opinion propagation were presented in [13].

SIR models are not the only models for studying opinion propagation in literature. For example, the threshold model starts from a random directed graph where each node selects a random threshold [14]. Opinion propagation then evolves deterministically: a node becomes active (gets an opinion) if the fraction of its active neighbours exceeds its threshold. In contrast to the threshold model, the dynamics of the voter model are stochastic: each node changes its state to the state of a random neighbour [15]. The Sznajd model [16] assumes more complex interactions between nodes: a node and a neighbouring node are selected at random. If the neighbouring node is undecided, it adopts the opinion with some probability. If both have the same opinion, they try to convince the other neighbours. Finally, if they have different opinions, nothing happens. Specifically focusing on tweet propagation, Yang et al. [17] propose a factor graph model based on the analysis of the factors influencing the user's retweet behaviour. Kawamoto et al. [18] model the information diffusion as a random multiplicative process, with a particular focus on retweet behaviour. In [19], the traditional Susceptible–Infected–Susceptible (SIS) epidemic model is studied in order to predict retweeting trends.

In this paper, we focus on a compartmental Markovian SIR model for opinion spreading in bounded medium-sized populations. While the total population of internet users easily qualifies as large, the size of an online community – say, of people contributing to an online forum or of people tweeting and retweeting some hash tag – is often not that large. Moreover, these communities hardly remain constant over time, with individuals joining and leaving all the time. We study such communities assuming that the number of individuals is bounded, there is a maximum number of individuals in the community. Such epidemic processes on bounded medium-sized populations are interesting from a mathematical point of view as well. While our numerical results reveal that a fluid limit approach yields accurate results when individuals only remain for a limited time (cf. Fig. 7), such an approach cannot be used to assess the system when individuals remain for a long time. For small populations, this is not problematic as small population sizes translate into small state spaces of the corresponding Markov processes such that the steady-state probability vector is easily calculated. In contrast, when the bound (and therefore also the state space) is larger – we use the term medium-sized populations – direct calculation of the stationary vector is computationally infeasible and performance needs to be assessed by other means when individuals remain for a long time. This is the subject of the present paper.

The contributions of this paper are twofold. First, we investigate the performance of opinion propagation in a Markovian framework by an approximate solution technique for Markov processes which relies on Maclaurin-series expansions of the steady-state probability vector. This technique was recently applied to study kitting processes [20]; a kitting process is a type of multi-buffer queueing system in which service is synchronised between the different buffers and temporarily blocked if one of the buffers is empty [21]. The epidemic process under consideration generalises the Markovian SIR process in various ways. We assume that the population size is bounded, but individuals join and leave the population over time to account for the dynamic formation of online communities. Moreover, the assumptions on infection and recovery rates are relaxed and individuals are allowed to move from susceptible to recovered directly as community members not necessarily want to spread the opinion to others. Secondly, in addition to the Maclaurin-series approach, we consider a fluid limit of the Markov process at hand and formally prove convergence. Fluid limits are a popular mathematical technique (see e.g. [22–28] and the references therein) which (when a good scaling is found) allow for focusing on the salient features of the stochastic process while discarding 'second-order fluctuations' around this main trend. In the present paper, it helps to make the link with more standard deterministic SIR models. We like to mention that the fluid scaling under study (arrival rates and location capacity are sent to infinity), differs significantly and therefore complements the Maclaurin-series expansion limit (which holds for low departure rates). We thus aim to view this computationally cumbersome Markov model from different limiting cases, and gain new insights by combining them. We also note that the derivation of the fluid limit as performed in this paper also lends itself naturally to refinements in the form of diffusion results, but this is considered to be outside of the scope of the current paper.

The remainder of this paper is organised as follows. Section 2 introduces the opinion spreading model at hand as well as some particular examples, discussed further on. In Section 3, the balance equations are derived and the numerical series expansion approach is explained. Next, we find a fluid limit for the epidemic Markov process in Section 4. To illustrate both approaches, Section 5 considers various numerical examples. Finally, conclusions are drawn in Section 6.

2. Model description

We consider an opinion propagation system as depicted in Fig. 1. There are at most L individuals in the community, each individual either being recovered (r), infected (i) or susceptible (s) (this particular ordering instead of the traditional s, i, r will prove useful for the Maclaurin analysis of Section 3). Let $X_k(t)$ be the number of individuals of type $k \in \mathcal{K} = \{r, i, s\}$ at time t, and let $\mathbf{X}(t) = (X_r(t), X_i(t), X_s(t)) \in \mathcal{L} = \{(x_r, x_i, x_s) \in \mathbb{N}^3 | x_i + x_r + x_s \leq L\}$. For any $\mathbf{x} \in \mathcal{L}$, x_k is the number of