Travel and Tourism Research Association: Advancing Tourism Research Globally

2014 ttra International Conference

Opinion spreading of a tourism-related topic in an online travel forum

Eline De Cuypere
Department of Telecommunications and Information Processing, Ghent University

Koen De Turck
Department of Telecommunications and Information Processing, Ghent University

Dieter Fiems
Department of Telecommunications and Information Processing, Ghent University

Opinion spreading of a tourism-related topic in an online travel forum

Eline De Cuypere, Koen De Turck and Dieter Fiems
Department of Telecommunications and Information Processing, Ghent University
St-Pietersnieuwstraat 41, B-9000, Belgium
+32 92 64 34 12
eline.decuypere; koen.deturck; dieter.fiems@ugent.be.

Opinion spreading of a tourism-related topic in an online travel forum

Abstract

Nowadays, potential tourists base their purchase decisions on traveler's experiences, comments and opinions posted in online travel forums. The propagation of online tourism-related messages is therefore of major importance for the tourism industry. Hence, we propose a model for the dynamics of single opinion propagation as a result of messages posted in an online travel forum. In particular, opinion spreading is modeled as a Markovian non-standard epidemic model. Indeed, opinion propagation in social networks holds many qualitative similarities with infectious diseases spread. We evaluate the performance of such systems by an approximate efficient solution technique, illustrate our approach by means of numerical examples and conclude our work.

Keywords: Opinion propagation; Online travel forums; Epidemic model; MacLaurin series expansion; Performance analysis.

1. Introduction

The emergence of social media as an all-encompassing platform, where users can share ideas, pictures and experiences, has completely stirred up tourist behavior in their search of information. Indeed, the online Word-Of-Mouth (abbreviated as WOM) generated by postings in online travel forums are given more importance in the decision making process than supplier produced information (Morrison, 2002; Smith et al., 2007; Murphy et al., 2007; Hwang et al., 2011; Arsal et al., 2008). In particular, professional organizations e.g. travel agencies that may find benefit by providing information are found to be less reliable than online users who engage a WOM communication to share knowledge e.g. gained from a previous trip. Furthermore, the impossibility to test the tourism product before consumption makes potential tourists eager to search for information about other people's experiences in order to reduce uncertainty and to create some expectations (Zeithaml et al., 2006; Kotler, et al., 2010). A last important factor which motivates tourists to use social media platforms is the ease to access sources of information (McGrath, 2008). Hence, there is a very strong interest in understanding how an opinion influences potential tourists through a social media platform. Indeed, a better comprehension of the travel decision-making process online is key to success in the tourism industry.

There are several types of social media platforms where online users can interact and share information starting from blogs, chat rooms, forums, etc. Each type of platform fulfills a different need as it promotes a different way of interactions among the users. Forums, also known as message boards or discussion groups, provide relevant features for sharing opinion related to travel. Indeed, anyone can start a topic and reply to it at any time. Also, members are considered to be at equal level and the content is usually segmented by topic. Examples of online travel forums are losviajeros.com (2008) and the Lonely Planet's Thorn Free forum (1996).

In this work, we propose to study the propagation of a travel-related opinion in an online forum based on the dynamics of the spreading of diseases (Andersson and May, 1992; Hethcote, 2000). Clearly, opinion propagation in social networks holds many qualitative similarities with how infectious diseases spread. These disease models typically divide a population into classes that reflect the epidemiological status of individuals (e.g. susceptible, infected, recovered, etc.), who in turn transit between classes via mutual contact at given average rates. The application of these models on opinion propagation has been thoroughly studied in literature. Luo and Wuzhong (2012) developed an epidemiological model to predict the least time that is needed to propagate a message to a whole population in a small-world network. Liu and Zhang (2014) proposed to study dynamic social networks as an epidemiological model with fixed recovery times. Moreover, they applied the link rewiring strategy on the model to enhance information spreading efficiency. Gruhl et al. (2004) presented a model to predict the trend in authors' posting behavior in the blogosphere, on the assumption that they do not write multiple posts on a single topic. Bettencourt et al. (2006) applied several epidemiological models to empirical data on the advent and spread of Feynman diagrams through the scientific communities in the USA, Japan and USSR.

In contrast to previous research, this paper focuses on gaining insights in the dynamics of tourism-related opinion propagation in a travel online forum. Therefore, we propose to study opinion spreading as a Markovian non-standard SIR epidemic model. The abbreviation SIR stands for **S**usceptible (S) **I**nfected (I) and **R**ecovered (R). These are consequently the only possible states that an individual can be in when we discuss SIR type diseases. In the traditional SIR models, the only possible transitions between the states follow the order S-I-R. In particular, the SIR model suggests that if a healthy individual encounters a sick individual, there is a specific probability that the healthy individual will get infected. Also, there is a specific probability that an infected individual will recover from the disease. The propagation of an opinion in an online forum can be interpreted as behaving like a disease: users interested in a specific tourism-related topic decide to follow the discussion about this topic in an online forum (susceptible). After having read the messages, they may give their opinion by posting a message in this forum (infected). Then, these active users may become passive about the topic (recovered). In particular, they stop posting messages on the forum as they don't want to share their opinion any more or lose their interest in the topic.

The traditional SIR epidemic model requires however some modifications for the study of tourism-related opinion spread in an online forum. In particular, we allow each type of user to leave the discussion at any moment. Furthermore, inspired by existing online travel forums, we make the following additional assumptions, thereby producing a more realistic candidate model: (i) a new user on the forum can directly become passive about the tourism-related topic and (ii) the rates at which users come online to follow the discussion and change of type depend on the number and the type of users currently following the discussion.

The first assumption allows a new user on the forum to directly become inactive. This means that users on the forum can decide not to reply to the posted messages and eventually leave the discussion. In the second assumption, the number and the type of users are explicitly taken into account in the dynamics of opinion propagation. Indeed, the more messages are posted about a certain topic, the higher the probability other users following the topic will also start to write. Also, the total number of users following a topic impacts positively the number of users who will decide to follow the discussion. These assumptions are inspired by the content of the Lonely Planet's forum. In this forum, the total number of users who clicked on the title of a topic and of replies can be found. These numbers are given with the aim of positively influencing the number of users that will decide to follow the discussion and give their opinion.

The remainder of this paper is organized as follows. Section 2 describes the opinion propagation model at hand. In section 3, the balance equations are derived and the methodology is shortly explained. Our approach is illustrated by means of numerical examples and conclusions are drawn in section 4.

2. **Model description**

In this research, we model a single-opinion propagation system as the continuous-time Markov chain depicted in figure 1. In particular, we focus on the formation and spread of one opinion about a tourism-related topic and the influence of other opinions posted on this forum is assumed to be negligible. Also, we assume a forum where at most L users can be present. New users click on the link of this forum according to a Poisson process and remain on it for an exponentially distributed amount of time. Hence, the state of the system can be described by a vector $\mathbf{i} = (i_R, i_I, i_S)$ where i_R, i_I and i_S are respectively equal to the number of recovered, infected and susceptible individuals. Susceptible individuals represent users who have clicked on the link and start to read the posted messages. These individuals can have an opinion or not about the topic. Infected individuals represent users who have read the posted messages and actively participate to the conversation by posting a message. These individuals agreed on or have formed their opinion and want to spread it among the other users by replying on the posted messages. Recovered individuals represent users who have read the posted messages and passively participate to the conversation by not replying to it. These individuals have lost their interest in the topic, don't want to share their existing or formed opinion or are still neutral.

We denote $Л = \{(i_R, i_I, i_S) \epsilon N^3 \mid s(\mathbf{i}) \leq L\}$ as the state space of the Markov chain and $S = \mid Л \mid$ as the size of the state space. The sum of all individuals at the location $s(\mathbf{i}) = \sum_{k \epsilon K}, i_k \leq L$ must be smaller than or equal to the maximum number of individuals L. Furthermore, we assume the arrival rate of new users $\lambda_S(\mathbf{i})$ to be state-dependent. Also, the probability for new users to become active $\alpha_{SI}(\mathbf{i})$, the probability for new users to become passive $\alpha_{SR}(\mathbf{i})$ and the probability of active users to become passive $\alpha_{IR}(\mathbf{i})$ are state-dependent. In particular, we consider the following rates to model the propagation of a tourism-related opinion in an online forum:

$\lambda_S(\mathbf{i}) = (1 + s(\mathbf{i})/L\ )\ \lambda_S,$
$\alpha_{SI}(\mathbf{i}) = (\alpha_{SI}^0 + \alpha_{SI}^1\ i_I/L\ )\ i_S,$
$\alpha_{SR}(\mathbf{i}) = \alpha_{SR}^0\ i_S,$
$\alpha_{IR}(\mathbf{i}) = \alpha_{IR}^0\ i_I,$

where λ_S, α_{SI}^0, α_{SI}^1, α_{SR}^0 and α_{IR}^0 are positive rates. The motivation behind these functions is the following. The higher the total number of users on the forum, the higher the rate at which new users will click on the link of this forum. Also, the higher the rate at which current users have posted a message, the higher the rate at which new users will also post a message. Furthermore, we add the constant term α_{SI}^0 to represent new users that have formed an opinion by themselves and post a message on the forum, i.e. independently on the number of active users. Finally, we assume that new users and active users may become passive with rate α_{SR}^0 and α_{IR}^0, respectively.

Having established the modeling assumptions and settled our notation we now focus on the analysis of the above described opinion propagation model.

3. Analysis

To obtain a number of interesting performance measures, we derive the balance equations of the model and propose an efficient numerical method to calculate the steady-state probability vector.

Let $\boldsymbol{\pi}(\mathbf{i})$ denote the steady-state probability distribution of the Markov chain which satisfies the following balance equations:

$$\boldsymbol{\pi}(\mathbf{i}) \left(\sum_{k\epsilon K} \mathbf{1}\{i_k > 0\}\, i_k\, \mu + \sum_{k\epsilon K} \mathbf{1}\{s(\mathbf{i}) \leq L\}\, (1 + s(\mathbf{i})/L)\, \lambda_S + \mathbf{1}\{i_S > 0\}\, (\alpha_{SI}^0 + \alpha_{SI}^1\, i_I/L)\, i_S \right.$$
$$\left. + \mathbf{1}\{i_S > 0\}\, \alpha_{SR}^0\, i_S + \mathbf{1}\{i_I > 0\}\, \alpha_{IR}^0\, i_I \right)$$
$$= \sum_{k\epsilon K} \boldsymbol{\pi}(\mathbf{i} + \mathbf{e}_k)\, (i_k+1)\, \mu\, \mathbf{1}\{s(\mathbf{i}) < L\} + \sum_{k\epsilon K} \boldsymbol{\pi}(\mathbf{i} - \mathbf{e}_S)\, (1 + (i_R + i_I + i_S - \mathbf{e}_S)/L)\, \lambda_S\, \mathbf{1}\{i_S > 0\}$$
$$+ \boldsymbol{\pi}(i_R, i_I - 1, i_S + 1)\, (\alpha_{SI}^0 + \alpha_{SI}^1\, (i_I - 1)/L)\, (i_S + 1)\, \mathbf{1}\{(i_R, i_I - 1, i_S + 1) \epsilon L\}$$
$$+ \boldsymbol{\pi}(i_R + 1, i_I, i_S + 1)\, \alpha_{SR}^0\, (i_S + 1)\, \mathbf{1}\{(i_R + 1, i_I, i_S + 1) \epsilon L\}$$
$$+ \boldsymbol{\pi}(i_R - 1, i_I + 1\ 1, i_S)\, \alpha_{IR}^0\, (i_I + 1)\, \mathbf{1}\{(i_R - 1, i_I + 1, i_S) \epsilon L\}\ ,$$

where $\mathbf{1}\{.\}$ refers to the indicator function of the event between brackets and $\mathbf{e}_k$ is a zero row vector with K elements and a single one entry at the *k*th position. Hence, the vectors $\mathbf{i} + \mathbf{e}_k$ and $\mathbf{i} - \mathbf{e}_k$ describe respectively the increase and the decrease of the *k*th element of the vector $\mathbf{i}$ by one.

While the former system of equations is easily solved for a low maximum number of individuals, the state space explodes for even reasonable L and a direct solution is computationally infeasible. To cope with this inherent state space explosion, we propose an approximate numerical algorithm which relies on MacLaurin series expansions of the steady-state probability vector in μ. More information about this solution approach can be found in the paper of De Turck et al. (2012) and De Cuypere et al. (2013). In the paper of De Turck et al. (2012), cases for which numerical computation of the steady-state vector is possible through a MacLaurin series expansion are given. The approach is illustrated by the practical example of a kitting process. In the paper of De Cuypere et al. (2013), a traditional Markovian SIR model for opinion propagation is developed and solved with this novel solution approach.

In the last section, we illustrate our approach by means of numerical examples and conclude our work.

4. Results & Conclusion

In this work, we aim to gain insights in the dynamics of opinion propagation about a tourism-related topic in an online forum. In particular, we propose to quantify the impact of different parameter settings on interesting performance measures. Note that the considered parameter values and performance measures are chosen only by way of illustration and are not to be considered as limiting.

Figure 2 depicts the mean number of active users EQ_I and passive users EQ_R versus the refusing rate α_{SR}^0. We assume the infecting rates α_{SI}^0 and α_{SI}^1 and the recovery rate α_{IR}^0 to be equal to 3. The maximum number of individuals L equals 20 and we assume different values of arrival rates as depicted in the figure. As can be seen in this figure, the mean number of active users decreases and the mean number of passive users increases as the refusing rate increases. Furthermore, the mean number of active and passive users decreases as the arrival rate of new users increases. Indeed, the higher the rate at which a new user arrives, the higher the probability new users will become active or passive users. However, the impact of the arrival rate on both performance measures decreases as the value increases. Intuitively, this is because we assume a maximum number of users on the forum. Figure 3 depicts the mean number of new users EQ_I versus the rate at which a user leave the forum μ. We assume the same parameter values as in figure 2 and the arrival rate to be equal to 3. We depict the mean number of new users for different values of the infecting rate α_{SI}^1. As can be seen by the figure, the lower the rate at which a user leaves the forum, the smaller the mean number of new users. Indeed, the smaller μ, the higher the probability new users eventually become active or passive users. Also, the mean number of new users decreases as the infecting rate α_{SI}^1 increases. Indeed, the higher the rate at which a new user becomes active, the lower the average number of new users in the forum.

To conclude, the contribution of this work is twofold. Firstly, we proposed a new model for the dynamics of a single-opinion propagation of a tourism-related product in an online forum. Secondly, we applied an efficient numerical algorithm which relies on MacLaurin series expansions to calculate the steady-state probability vector. This allows us to obtain interesting performance measures so that we can assess the impact of the parameter settings on opinion spreading. Future work will focus on expanding the system to a multiple opinion propagation model and on fitting the parameter settings with data sampled from existing or new online forums.

References

1. Roy M. Andersson and Robert M. May (1992). Infectious diseases of humans. Oxford University Press, UK.
2. Irem Arsal, Sheila Backman, and Elizabeth Baldwin (2008). “Influence of an Online Travel Community on Travel Decisions”. In Information and Communication Technologies in Tourism. Pp 82-93.
3. Eline De Cuypere, Koen De Turck, Sabine Wittevrongel, and Dieter Fiems (2013). “Markovian SIR model for opinion propagation”. Proceedings of the 25th International Teletraffic Congress, Shanghai.
4. Koen De Turck, Eline De Cuypere, Sabine Wittevrongel and Dieter Fiems (2012). “Algorithmic approach to series expansions around transient Markov chains with applications to paired queueing systems”. Proceedings of the 6th International Conference on Performance Evaluation Methodologies and Tools, Valuetools, Cargèse.
5. Herbert W. Hethcote (2000). “The mathematical of infectious diseases”. SIAM Review, 42:599-563.
6. Yeong-Hyeon Hwang, Dev Jani and Ho Kyun Jeong (2011). “Analyzing international tourists' functional information needs: A comparative analysis of inquiries in an on-line travel forum”, Journal of Business Research 66 (2013):700–705.
7. Losviajeros.com (2008), URL: <http://www.losviajeros.com>. (24/04/2014).
8. Chuang Liu and Zi-Ke Zhang (2014), “Information spreading on dynamic social networks”. Communications in Nonlinear Science and Numerical Simulation, 19(4):896-904.
9. Philip Lu, John T. Bowen, and James C. Makens (2010). Marketing for Hospitality and Tourism (5th edition). Boston, MA: Pearson Education.
10. Dan Luo, and Zhou Wuzhong. (2012). “Characteristics of the tourism-related perception relation”. Advances in Computer, Communication, Control and Automation Lecture Notes in Electrical Engineering, 121:337-344.
11. Laurie Murphy, Gianna Mascardo and Pierre Benkendorff (2007). “Exploring word of mouth influence on travel decisions: friends and relatives vs. other travelers”. International Journal of Consumer Studies, 31:517–27.
12. Alastair M. Morrison (2002). Hospitality and Tourism Marketing (3rd ed.), Delmar.
13. Donnavieve Smith, Satya Menon and K. Sivakumar (2007). “Online peer and editorial recommendations, trust and choice in virtual markets”. Journal of Interactive Marketing. 19 (3) 15.
14. Planet’s Thorn Free forum (1996), URL: <http://www.thornfree.lonelyplanet.com>. (24/04/2014).
15. Valerie A. Zeithaml, Mary Jo Bitner and Dwayne Gremler (2006). “Services marketing: Integrating customer focus across the firm”. 4th edition London: McGraw Hill.

Figures

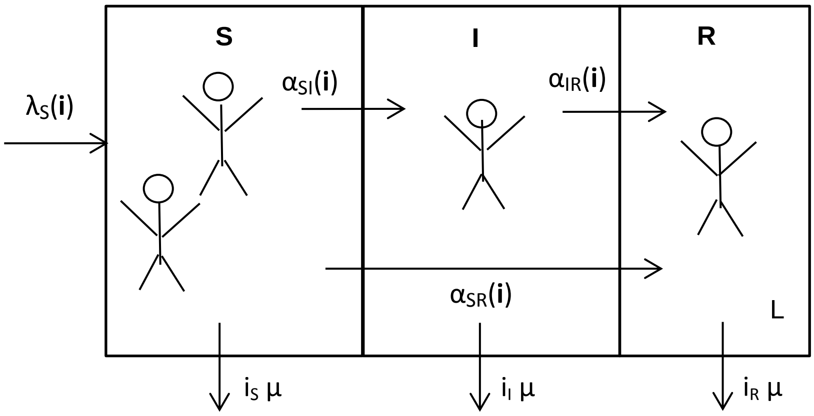

Figure 1: Tourism-related single-opinion propagation model

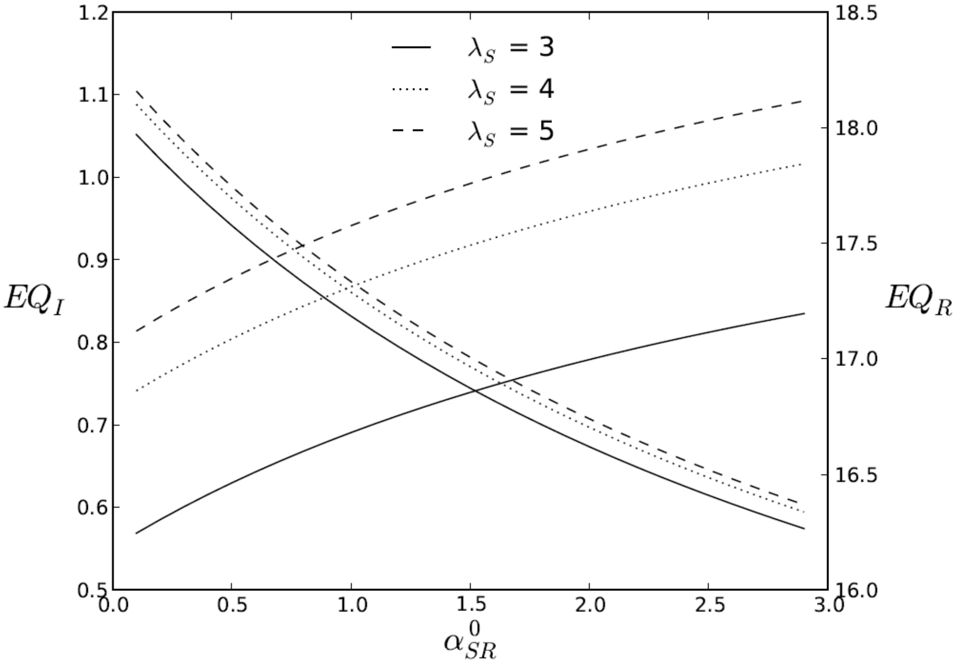

Figure 2: Mean number of active (left) and passive (right) users.

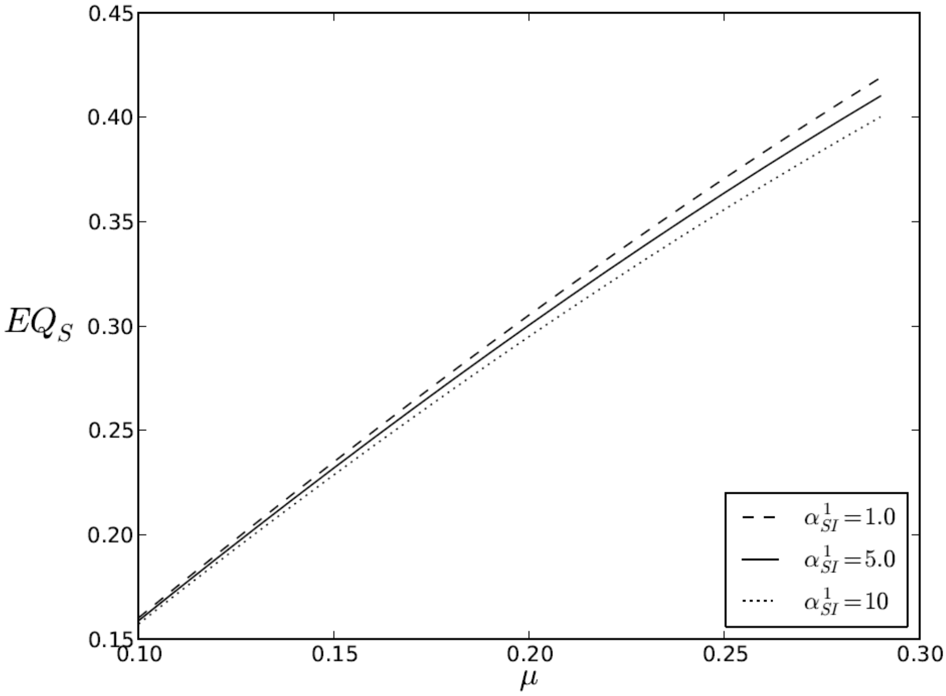

Figure 3: Mean number of new users.