Omega 59 (2016) 15–25

Contents lists available at ScienceDirect

Omega

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

Social media optimization: Identifying an optimal strategy for increasing network size on Facebook

Michel Ballings [a,*], Dirk Van den Poel [b], Matthias Bogaert [b]

[a] Department of Business Analytics and Statistics, The University of Tennessee, 249 Stokely Management Center, 37996 Knoxville, TN, USA
[b] Department of Marketing, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium

ARTICLE INFO

Article history:
Received 28 August 2014
Accepted 29 April 2015
Available online 10 June 2015

Keywords:
Social media
Facebook
Strategy optimization
Predictive modeling
Prescriptive modeling
Random Forest
Genetic Algorithm
Data mining

ABSTRACT

This paper aims to create an expert system that yields an optimal strategy for increasing network size on Facebook. Data were obtained from 5488 Facebook users by means of a custom-built Facebook application. We computed a total of 426 variables. Using these data we estimated a predictive model of network size which is subsequently used in a prescriptive model. The former is estimated with Random Forest and the latter with a Genetic Algorithm. The results indicate that the proposed expert system can identify an optimal social media strategy. The system delivers concrete recommendations about, for example, the optimal time between status updates. The analysis reveals that network size can be increased by 61% if the optimal strategy is adopted. This study contributes to literature in the following two ways. First it devises a novel prescriptive social media expert system relying on an unprecedented variety of social media data. The results indicate that the system is effective and a viable strategic tool for increasing network size. Second it provides a list of the top drivers allowing future research to build similar systems efficiently.

1. Introduction

With 1.35 billion monthly active users and 864 million daily users on average [24], Facebook has become an important communication channel to interact with peers and customers [27,48] and distribute news in general. For advertisers, Facebook has evolved to become a chief advertising channel next to TV, print and radio. At the same time increasingly more questions are being raised about the platform's advertising effectiveness and how to increase it. One of those questions is ‘How to increase advertising reach on Facebook?’.

There are three main strategies to increase advertising reach, called post reach, on Facebook: (1) improve the effectiveness of a post (e.g., [16,18,39]), (2) increase network size (e.g., [40,43]), and (3) buy more reach. The former two strategies are called organic strategies and the latter strategy is called a paid strategy. While paid advertising has been heavily researched (e.g., [1,19,20,28]), literature on organic strategies is very scarce. The first strategy focuses on adapting the characteristics of a post (length, type, and timing) to increase the number of likes and comments it receives [39]. In turn, an increased number of likes and comments will result in Facebook showing the post to more Facebook users and hence increasing the post's reach. However, this does not necessarily increase the size of the network. A post may be very effective, but if the poster only has a very small network, posts will still have a small reach. The second strategy adopts a more holistic approach by analyzing the user's entire behavioral pattern on Facebook. In contrast to the first strategy, the focus is not on trying to increase the effectiveness of a post, on a post-by-post basis, but on improving the effectiveness of all of a user's posting behavior. Examples of this strategy are determining the optimal number of photos, videos, links and status updates to post, which pages a user should like or, how much time should go by before posting something new. Finally, the third strategy entails paying Facebook to increase a post's reach beyond its organic reach. Apart from paying Facebook, no further action is required from the poster.

This study focuses on the second strategy, increasing network size, because of the following two main reasons. First, we assume that practitioners are initially interested in increasing their posts' reach organically before they start paying for extra reach. Second, the network size strategy has been under-researched as opposed to the first strategy. While there is some extant literature on the second strategy, studying the associations between profile elements and network size [40,43], these studies do not give clear guidance on which actions are required to increase network size.

* Corresponding author.
E-mail addresses: Michel.Ballings@utk.edu (M. Ballings), Dirk.VandenPoel@UGent.be (D. Van den Poel), Matthias.Bogaert@UGent.be (M. Bogaert).

http://dx.doi.org/10.1016/j.omega.2015.04.017
0305-0483/

In this paper we attempt to fill this gap by devising a novel prescriptive social media expert system relying on an unprecedented variety of social media data. The system is aimed at providing recommendations to maximize one's network size. We provide a detailed description of the expert system's different components along with a list of the top variables, allowing future research to build similar systems efficiently.

The remainder of this paper is structured as follows. In the next section we provide a literature review on studies that explain network size on Facebook in order to highlight our contribution. Next, we elaborate on the methodology including the expert system, data, variables, analytical techniques and performance evaluation. In the penultimate section we discuss our findings and their implications. In the final section we address the limitations and avenues for future research.

2. Literature review

To the best of our knowledge only two articles have been published that investigate the drivers of Facebook network size using Facebook data. Lampe et al. [40] collected data from 30,773 Facebook users through web crawling. They found that the amount of information in profiles is weakly related to number of friends. Conversely, the presence of information, versus no information at all, was found to be strongly related. Professional/educational variables, demographic variables (e.g., gender) and general Facebook account variables (length of Facebook membership, and recency of last update) had a significant effect on the number of friends. In addition, the analyses showed that whether a user shared geographic data (hometown and residence) is a better driver of number of friends, than whether he or she shared demographic data (e.g., contact information, relationship status, and birthday). The latter is in turn a better driver than whether he or she shared personal data such as interests and opinions. Lewis et al. [43] also used web crawling to extract data from 1,215 Facebook users. In accordance with Lampe et al. [40] they found that general Facebook account variables (length of Facebook membership, and recency of last update) had a significant influence on number of friends. In addition, the demographic variable race/ethnicity played an important role.

Extant literature is limited in that the authors only go as far as estimating descriptive models. Those models include a relatively small number of variables (see Appendix A for an overview of the variables used in extant literature) and typically investigate descriptor coefficients. Because literature on network size in Facebook is so scarce we draw on additional literature to determine which other variables are likely to be top drivers and should therefore be included as candidate variables in the predictive modeling process. The result of this review is covered in the following paragraph.

Every time a user (the initiator) changes something on his or her profile or makes an update, it is propagated to friends' News Feeds, creating an opportunity for reaction (e.g., liking or commenting) from those friends. A friend's reactions will in turn be propagated to the initiator's friends-of-friends subsequently creating awareness of the initiator. This could possibly lead friends-of-friends to click through to the initiator's profile and make a friend request. Hence, every change in profile elements and every update can be conceived of as advertising and friends' interactions can be conceived of as word-of-mouth. Friends' interactions could therefore be of key importance in predicting network size. In advertising literature, a chief strategic dimension is repetition [41]. Advertising performance has been shown to improve when advertisements are repeated within a given time interval [41,50,35]. More specifically, and in accordance with mere exposure theory [58], there is a positive relationship between exposure frequency and brand measures such as brand awareness [21], brand recall [54], perceived brand quality [50], number of associations in memory [4] and brand recognition [35]. This means that more changes by a Facebook user will result in more interactions with those changes by friends, which in turn will improve the initiator's visibility with friends-of-friends. It is plausible that this visibility is translated into a higher number of friend requests. In contrast to profile elements, updates change much more frequently. For example, the profile element ‘gender’ will never change while the field ‘status’ will change frequently and therefore updates are most likely more important for predicting network size than profile elements.

Because they are likely to be important we made sure that both updates (status updates, photo uploads, photo user tags, video uploads, album uploads, likes, tags, events, link uploads, check-ins, and notes made) and friends' interactions (likes and comments) with a user's updates are included as candidate variables. In addition to including updates and friends' interactions with these updates we include all types of profile elements (see Appendix A for the full list).

3. Methodology

3.1. Expert system

The expert system (Fig. 1) starts by extracting live data from Facebook users with their authorization. Next, data are transformed yielding the variables in Appendix A. Using Random Forest [6] we then estimate a predictive model of network size. This model is used to assess the importance of specific variables. In the prescriptive engine, the top n drivers are selected to be optimized by a Genetic Algorithm. The number of variables is a parameter to be set by the user. The Genetic Algorithm includes the predictive model in the objective function. The prescriptive engine will output an optimal strategy (i.e., the optimal values of the top n drivers). To increase network size, the user-admin executes that strategy which in turn

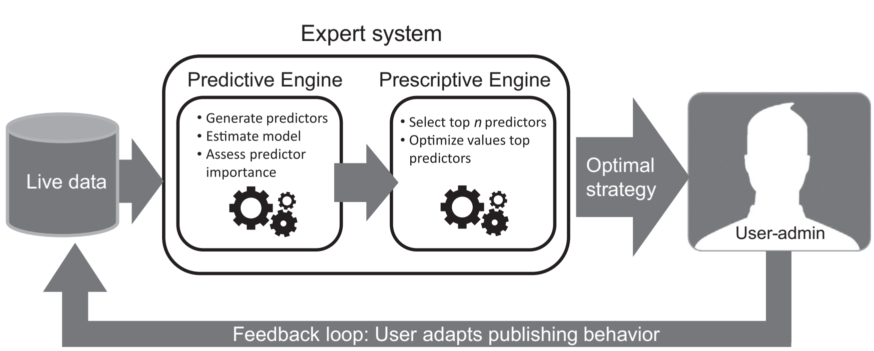

Fig. 1. Expert system.