International Journal of Data Warehousing and Mining
Volume 17 • Issue 1 • January-March 2021

Ranking News Feed Updates on Social Media:
A Review and Expertise-Aware Approach

Sami Belkacem, LSI Laboratory, University of Science and Technology Houari Boumediene, Algiers, Algeria
https://orcid.org/0000-0002-7259-9054

Kamel Boukhalfa, LSI Laboratory, University of Science and Technology Houari Boumediene, Algiers, Algeria
https://orcid.org/0000-0002-9746-579X

ABSTRACT

Social media are used by hundreds of millions of users worldwide. On these platforms, any user can post and share updates with individuals from his social network. Due to the large amount of data, users are overwhelmed by updates displayed chronologically in their newsfeed. Moreover, most of them are irrelevant. Ranking newsfeed updates in order of relevance is proposed to help users quickly catch up with the relevant updates. In this work, the authors first study approaches proposed in this area according to four main criteria: features that may influence relevance, relevance prediction models, training and evaluation methods, and evaluation platforms. Then the authors propose an approach that leverages another type of feature which is the expertise of the update's author for the corresponding topics. Experimental results on Twitter highlight that judging expertise, which has not been considered in the academic and the industrial communities, is crucial for maximizing the relevance of updates in newsfeeds.

KEYWORDS

Expertise, Newsfeed Updates, Ranking, Relevance, Social Media, Twitter

1. INTRODUCTION

Social media such as Facebook or Twitter are used by hundreds of millions of users worldwide and contribute to the concept of Big data (Kuang et al., 2016). Social data are known for their large volumes that can reach petabytes (10^{15} bytes), their variety (text, images, videos, music, etc.), and their velocity (arriving near real-time) (Kim et al., 2014). On social media, any user can share updates with individuals from his social network (Lahoti et al., 2017). Due to the large amount of data (Al-Sheikh and Hasanat, 2018), users find themselves overwhelmed by updates displayed chronologically in their news feed, from most recent to least recent (De Maio et al., 2017). For example, there are about 1500 new updates every day in the news feed of a typical Facebook user[1]. Moreover, several works have shown that most updates are irrelevant (Vougioukas et al., 2017). For example, Paek et al. (2010)

DOI: 10.4018/IJDWM.2021010102

asked 24 Facebook users from their Microsoft organization to assign relevance scores to news feed updates. The overall average score was close to 0. Hence, large data volume and irrelevance make it difficult for users to catch up with the relevant updates (Kuang et al., 2016). A standard Facebook user, for example, may miss relevant updates in the news feed even if he spends an average of 55 minutes a day on *Facebook* (Paek et al., 2010).

In several research approaches, ranking news feed updates in descending relevance order has been achieved based on the prediction of a relevance score between a beneficiary user and a new update in the news feed (Agarwal et al., 2015). In our previous work[2], we compared several approaches according to four main criteria: features that may influence the relevance of updates to beneficiary users, relevance prediction models, methods used to obtain training and evaluation data, and evaluation platforms. As regards the features that may influence relevance, we noticed the predominance of four types of features: (1) the relevance of the update content to the beneficiary's interests; (2) the social tie strength between the beneficiary and the update's author; (3) the author authority; and (4) the update quality. We believe that using these features when predicting the relevance of updates is necessary, but not sufficient. For example, updates on a specific topic authored by a novice user may not attract the attention as much as those posted by a recognized expert in his field. Indeed, updates posted by experts, also known as topical influencers (Zengin Alp and Gündüz Öğüdücü, 2018) or topical authorities (Lee et al., 2018), are often considered up-to-date, credible, and interesting (Wagner et al., 2012). Therefore, leverage the author's expertise could be fundamental to enable users to catch up with the valuable and trustworthy updates on specific topics.

Expertise information is usually not explicitly provided by social media users (Xu et al., 2017). Hence, existing methods rely on implicit expert finding, which aims at identifying users with the relevant knowledge on a given topic (Wei et al., 2016). The expert finding is a critical problem that has been studied in many applications such as viral marketing, recruiting talent, link and user recommendation, and question answering (Lee et al., 2018). On social media, the main techniques used to infer a user's expertise leverage other users' interactions with the textual content he posted (Li et al., 2013) as well as his past social behaviors, including: user-generated content, biographical information, social relationships, list memberships[3], etc. (Xu et al., 2017). Based on existing work, to the best of our knowledge, the author's expertise has not been used when predicting the relevance of updates to beneficiary users. In this paper, to extend our previous work[2], we first consider the approaches proposed in the area of ranking news feed updates according to the four criteria mentioned above. Then, we study the contribution of expertise to rank news feed updates and propose an approach that, in addition to other features used in related work, leverages the author's expertise that we infer from users' interactions with the textual content he posted.

The paper is structured as follows: Section 2 provides background on ranking news feed updates on social media, Section 3 discusses work carried out in this area, Section 4 describes the proposed approach which leverages the author's expertise, Section 5 presents the experiments we performed to evaluate our approach and study the contribution of expertise to rank news feed updates, and Section 6 concludes and proposes future work.

2. RANKING NEWS FEED UPDATES

In this section, we provide background information on ranking news feed updates on social media, including defining news feeds, exposing statistics that confirm the need for ranking and outlining the ranking process.

2.1. News Feed

According to Berkovsky et al. (2015), a user's news feed on social media (see Figure 1) is a list of information (posts, tweets, status, etc.) that allows him to follow updates about individuals from his social network. It includes text status updates, images, videos, etc. (Freyne et al., 2010). For example,