The Dynamics of Content Popularity in Social Media

Symeon Papadopoulos, Aristotle University of Thessaloniki, Greece, and Informatics & Telematics Institute, Greece

Athena Vakali, Aristotle University of Thessaloniki, Greece

Ioannis Kompatsiaris, Informatics & Telematics Institute, Greece

ABSTRACT

Social Bookmarking Systems (SBS) have been widely adopted in the last years, and thus they have had a significant impact on the way that online content is accessed, read and rated. Until recently, the decision on what content to display in a publisher's web pages was made by one or at most few authorities. In contrast, modern SBS-based applications permit their users to submit their preferred content, to comment on and to rate the content of other users and establish social relations with each other. In that way, the vision of the social media is realized, i.e. the online users collectively decide upon the interestingness of the available bookmarked content. This article attempts to provide insights into the dynamics emerging from the process of content rating by the user community. To this end, the article proposes a framework for the study of the statistical properties of an SBS, the evolution of bookmarked content popularity and user activity in time, as well as the impact of online social networks on the content consumption behavior of individuals. The proposed analysis framework is applied to a large dataset collected from digg, a popular social media application.

Keywords: Collaborative Technologies, Data Mining, Electronic Media, Online Behavior, Online Community, Resource Sharing, Web-Based Applications

INTRODUCTION

The startling success of Web 2.0 applications during the last years has reshaped our views on how information is generated and distributed to the masses. While traditionally the content items that appeared in the web pages of online publishers (e.g. news sites) were selected by a single authority or a small committee of experts, the advent of Web 2.0 applications, such as the Social Bookmarking Systems (SBS), has allowed mass participation in the content selection process.

In an SBS, users upload and save links to web pages (bookmarks) that they deem

interesting. These bookmarks are usually public: Once they are submitted by a user, the rest of the community is able to view them and if they consider them interesting, they may give a 'thumbs up', i.e. rate them. Thus, applications of this kind, originally meant to provide a convenient tool to organize one's bookmarks, have evolved to act as a content popularity (or 'interesting-ness') ranking mechanism. Examples of such systems are digg[1], propeller[2] and newsvine[3]. The proliferation and success of such applications has created online communities where the bookmarked content items undergo a perpetual rating process by the community. Typically, the most popular content items coming out of this process are displayed in a prominent place at the application's site. In that way, the vision of *social media* is realized.

The work presented in this article has been mainly motivated by the fact that the content rating processes, such as the ones taking place within social media applications, have not been sufficiently studied so far from the temporal and social point of view. For this reason, we established a framework for the study of such social media aspects and applied it for analyzing the content rating dynamics emerging in the digg application. More specifically, through the work described in this article, we make the following contributions:

- Introduce the *Diggsonomy* framework, inspired by the *Folksonomy* framework of (Mika, 2005) and (Hotho et al., 2006a), to enable the study of online content popularity;
- Collect and analyze the statistical properties of an extensive data set from digg;

- Provide insights to the temporal evolution of popularity and to the role of the social relations on the content rating process.

We consider the work presented here to be of significance for the following applications:

- **Administration of SBS and social media applications:** Since the operation of such a service involves a large investment in server and network equipment, better understanding of page popularity and thus request patterns could contribute to smoother service provision and/or reduced infrastructure costs by means of efficient data partitioning and page caching schemes.
- **Recommender system development:** Understanding of the temporal and social dynamics of content popularity can be exploited to improve the precision and relevance of recommendations for systems where these aspects are of particular significance (e.g. news articles, advertisements).

The rest of this article is structured as follows. Section 2 provides a short overview of existing work in the fields of mining web 2.0 data and analysis of related temporal and social phenomena. In Section 3 the proposed social media analysis and modeling framework is introduced. Further, the data collection and the associated experimentation as well as a discussion of the findings are provided in Section 4. Finally, Section 5 concludes the article.

Related Content

A Survey of Spatio-Temporal Data Warehousing
Leticia Gómez, Bart Kuijpers, Bart Moelans and Alejandro Vaisman (2009). *International Journal of Data Warehousing and Mining (pp. 28-55).*
www.irma-international.org/article/survey-spatio-temporal-data-warehousing/3895/

User-Centric Similarity and Proximity Measures for Spatial Personalization
Yanwu Yang, Christophe Claramunt, Marie-Aude Aufaure and Wensheng Zhang (2012). *Exploring Advances in Interdisciplinary Data Mining and Analytics: New Trends (pp. 128-146).*
www.irma-international.org/chapter/user-centric-similarity-proximity-measures/61172/

A Solution to the Cross-Selling Problem of PAKDD-2007: Ensemble Model of TreeNet and Logistic Regression
Mingjun Wei, Lei Chai, Renying Wei and Wang Huo (2008). *International Journal of Data Warehousing and Mining (pp. 9-14).*
www.irma-international.org/article/solution-cross-selling-problem-pakdd/1802/

Finding Persistent Strong Rules: Using Classification to Improve Association Mining
Anthony Scime, Karthik Rajasethupathy, Kulathur S. Rajasethupathy and Gregg R. Murray (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications (pp. 28-49).*
www.irma-international.org/chapter/finding-persistent-strong-rules/73432/

Mass Media Strategies: Hybrid Approach using a Bioinspired Algorithm and Social Data Mining
Carlos Alberto Ochoa Ortiz Zezzatti, Darwin Young, Camelia Chira, Daniel Azpeitia and Alán Calvillo (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications (pp. 1163-1188).*
www.irma-international.org/chapter/mass-media-strategies/73490/