



Review

Modelling to identify influential bloggers in the blogosphere: A survey



Hikmat Ullah Khan ^{a, b}, Ali Daud ^{c, b, *}, Umer Ishfaq ^d, Tehmina Amjad ^b, Naif Aljohani ^c, Rabeeh Ayyaz Abbasi ^{c, e}, Jalal S. Alowibdi ^f

^a Department of Computer Science, COMSATS Institute of Information Technology, Wah, Pakistan

^b Department of Computer Science and Software Engineering, International Islamic University, Islamabad, Pakistan

^c Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

^d Department of Computer Science, COMSATS Institute of Information Technology, Attock, Pakistan

^e Department of Computer Sciences, Quaid-i-Azam University, Islamabad, Pakistan

^f Faculty of Computing and Information Technology, University of Jeddah, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:

Received 1 August 2016

Received in revised form

7 November 2016

Accepted 9 November 2016

Keywords:

Social web

Blog

Model

Influential bloggers

Blogosphere

ABSTRACT

The user participatory nature of the social web has revolutionized the use of the conventional web. The social web is an integral part of our daily life. Due to the resulting exponential growth of the social web, a number of research domains have emerged, involving research activities that aim to study human nature, to analyse human sentiments and emotions, and to find the impact of various users in the social networks. Recently, the research focus has shifted to identifying a user's influence on other users in a social network. In the recent literature, we find a number of models proposed to find the most influential users in the blogging community. In this paper, we review the models to find these influential bloggers. The existing models are classified into feature-based and network-based categories. The feature-based models consider the salient factors to measure bloggers' influence. The network models, on the other hand, consider the graph-based social network structure of the bloggers to identify those who have the most impact on fellow members. This survey introduces each model with its features, novel aspects, and the datasets used. In addition to the discussion about the model, a comparative analysis of the datasets is presented. We conclude by discussing applications of the relevant literature, exploring open research issues and challenges, and sharing possible future directions in this active area of research.

© 2016 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	65
2. Existing models	66
2.1. Feature based models	66
2.1.1. Non-temporal models	66
2.1.2. Temporal models	72
2.2. Network based models	73
2.2.1. Non-temporal model	73
3. Datasets	77
3.1. TUAW	77
3.2. Technorati	77
3.3. Digg	78
3.4. BlogCatalog	78
3.5. SLR CLub	78
3.6. Wretch blog	78

* Corresponding author. Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia.

E-mail addresses: Hikmat.ullah@ciitwah.edu.pk, hikmat.phdcs55@iiu.edu.pk (H.U. Khan), ali_msdb@hotmail.com, adfmohamad@ksa.edu.sa (A. Daud), umer.bravo@gmail.com (U. Ishfaq), tehminaamjad@iiu.edu.pk (T. Amjad), nraljohani@kau.edu.sa (N. Aljohani), rabbasi@kau.edu.sa (R.A. Abbasi), jalowibdi@uj.edu.sa (J.S. Alowibdi).

3.7.	Microsoft MSN space	78
3.8.	RSS Feed Aggregator	78
3.9.	Engadget	78
3.10.	Naver blog	78
3.11.	Everything2	78
3.12.	Portal Salon24	78
3.13.	PBM	78
3.14.	Yahoo! movies data	79
4.	Applications	79
4.1.	Finding trend-setters	79
4.2.	Viral marketing	79
4.3.	Political domain	79
4.4.	Social issues	79
4.5.	Emotion and sentiment analysis	80
5.	Challenges and research issues	80
5.1.	Quantity measure of influence	80
5.2.	Whom to choose: influential or active bloggers	80
5.3.	Lack of ground truth	80
5.4.	Temporal aspect in the blogosphere	80
5.5.	Missing links in datasets	80
5.6.	Identification of consistent influentials	80
5.7.	Sparse link structure	80
5.8.	Influence maximization	80
6.	Conclusion	81
	Acknowledgements	81
	References	81

1. Introduction

Social web platforms provide an opportunity for users to establish online social interactions by carrying out activities such as participating in social networking websites and multimedia content sharing websites, blogging, online shopping, etc. (Gruber, 2008). Social networking sites have become an integral part of our daily life, with a large portion of the world population interacting using various social networking websites such as Facebook, LinkedIn, and Twitter. Blogging is a popular way for people to express their views on a variety of social issues. The views expressed can be in the form of text or multimedia, and these posts are known as “blog posts”. The members of the blog who can create blog posts are known as bloggers, or blog writers. The pool of all the web logs of the social web is called “the blogosphere”. Blogs represent a democratic world, where people can freely express their views in blog posts. People can voice their agreement or discontent by writing comments (Kim, Jeong, & Leeb, 2010). One of the fascinating aspects of blogging is that it has facilitated the publishing of information on the web (Chi, 2008). Now, we are not only the information consumers as we use web to get information, but also the information producers as we may share information using social web channels.

In normal life, we consult various persons for their views and opinions about various topics. Likewise, bloggers consult others and seek their opinions about a number of social subjects, for instance, buying the products of a specific brand, selecting a movie, or advice on current fashion trends. These “others” whose advice or guidance is sought are known as influential bloggers (Agarwal, Liu, Tang, & Yu, 2012). Influential bloggers usually work in close collaboration. They create their own special interest groups where they initiate social or political discussions, spread awareness and gather support (Agarwal, Mahata, & Liu, 2014). The identification of influential bloggers has become significant due to the growing popularity of blogs. For instance, they may sway public opinion by helping to organize and mobilize social and political campaigns. It

has been argued in Agarwal et al. (2014) that the influential bloggers are the representatives of the virtual communities, thus they are in a position that they can influence public views and opinions this they can play a vital role. An organization can target such bloggers for their marketing purposed and can save commercial organizations millions of dollars in marketing and advertising expenses (Berthona, Pitt, Plangger, & Shapiro, 2012; Onishi & Manchanda, 2012). With the help of the influential bloggers, companies can also improve their understanding of changing market trends and offer better customer support (Aldhaheri & Christian, 2013). The influential users also render useful services in promoting citizen journalism. Even the mainstream media rely on the blogs for top stories and news reports (Lee et al., 2010).

The influence propagation in different social media platforms has been studied in a quantitative manner (Liu, Tang, Han, & Yang, 2012) by proposing a probabilistic model to discover the topics of discussion between friends. The influence propagation model doesn't only quantify the users' interest level towards the topic, but also measures the strength of influence between the social media friends. Domain specific influence on social media users has also been explored (Caiv & Chen, 2009). Classification based research (Trusov, Bodapati, & Bucklin, 2010) finds the influential bloggers on the basis of their content in their blog posts. It also figures out the influence over fellow online bloggers by maintaining the records of online social interactions as well. A few authors have worked on the innovative properties of the blogosphere and social media platforms. The features which encompass bloggers', blogging activity and characteristics of blog content are incorporated into the development of robust models, which provides improved results as compared to baseline methods to identify the influential bloggers (Eirinaki, Monga, & Sundaram, 2012; Li, Lai, & Chen, 2011; Qureshi, Younus, Saeed, & Touheed, 2010). Because this research domain is rapidly evolving, there is a constant need for a comprehensive overview of the most significant research developments and current research issues and challenges.

In this survey, we 1) classify the models for identifying

Download English Version:

<https://daneshyari.com/en/article/4937658>

Download Persian Version:

<https://daneshyari.com/article/4937658>

[Daneshyari.com](https://daneshyari.com)