

An overview of user profiling in area of online social networks

Yang Qi, ITMO University, Saint-Petersburg

Supervisor – Filchenkov A.A., ITMO University.

Introduction

With the popularization of social media around the world, exemplified by Facebook and Twitter in the US, and Sina Weibo in China, the result has been a massive accumulation of user-generated data. This data contains users' personal information, original posts, retweets, comments, and other information. User-generated data contain several particularities. First, while microblog text posted by users is limited to 140 characters, it contains valuable information. Second, the data are gathered in real-time and can be transmitted rapidly. Microblogs are open platforms and data can be obtained freely. Due to these particularities, a great deal of research interest has been sparked in certain aspects of microblogs; for example, applying user-generated data in prediction research.

Since an user can use the social networking sites for various purposes like instant messaging, media posting etc., it helps in connecting with friends, family from a distance apart. Students staying away from parents, spouse working in different geographical location depend on these sites to be closer with their loved ones. Now a days, people generally depend on the data generated from the internet for their personal purpose as well. In this way the usage of social networking sites have become part and parcel of everyone's life especially for youth. Users also post their day-to-day activities and concerns, opening an avenue for emerging data mining research.

The prediction of online users' demographic characteristics has extensive applicability and practical significance. For instance, researchers can carry out precise online marketing based on predicted gender. In social media networks, the majority of users are unwilling to disclose their gender, age, and other demographic characteristics. However, it is possible to predict user characteristics based on the content they generate in social media networks. Using these methods, social media users' personal information can be accurately predicted which allows for improved personalized recommendation systems and personalized marketing

Objective of this work

User Profiling is the process of Extracting, Integrating and Identifying the keyword based information to generate a structured Profile and then visualizing the knowledge out of these findings. User profiling helps personalizing a system to work according to user. Therefore user profiling or personalization is one of the major concepts used for accessing the user relevant information, which can be used to solve the difficult problems of recommender system like classification and ranking of items in accordance with an individual's interest.

Advances in Online Social Networks is creating huge data day in and out providing lot of opportunities to its users to express their interest and opinion. Identifying users is challenging and requires digging huge knowledge out of the data being flown in the social media. This work gives an insight to profile users in online social networks. User Profiles are established based on the behavioral patterns, correlations and activities of the user analyzed from the aggregated data using techniques like clustering, behavioral analysis, content analysis and face detection.

Depending on application and purpose, the mechanism used in profiling users varies. Further study on other mechanisms used in profiling users is under the scope of future endeavors.

For future researchers, profiling users based on their activities in line with current affairs in society, friend's activities, search engine log data are essential for fraud detection, determining influential user, personalization, e-marketing and advertising etc. The need for research on user data and their activities in online multi social networks simultaneously is essential for accurate profiling of any social network user.

Reference:

1.Itoh, Masahiko, Daisaku Yokoyama, Masashi Toyoda, Yoshimitsu Tomita, Satoshi Kawamura and Masaru Kitsuregawa. "Visual Exploration of Changes in Passenger Flows and Tweets on Mega-City Metro Network,"IEEE Transactions on Big Data, vol. 2, no. 1, pp. 85-99, 2016

2.Davis Jr, C.A., et al.: Inferring the location of twitter messages based on user relationships. Transactions in GIS **15**(6), 735–751 (2011)

3.Dougnon, R.Y., Fournier-Viger, P., Nkambou, R.: Inferring user profiles in online social networks using a partial social graph. In: Barbosa, D., Milios, E. (eds.) Canadian AI 2015. LNCS (LNAI), vol. 9091, pp. 84–99. Springer, Heidelberg (2015)

4.Jurgens, D.: That's what friends are for: inferring location in online social media platforms based on social relationships. In: Proc. of the 7th International AAI Conference on Weblogs and Social Media, pp. 273–282. AAAI Press (2013)

Author _____

Supervisor _____