

# Chapter 17

## Finding Similar Users in Facebook

**Pasquale De Meo**

*University of Messina, Italy*

**Emilio Ferrara**

*University of Messina, Italy*

**Giacomo Fiumara**

*University of Messina, Italy*

### **ABSTRACT**

*Online social networks are rapidly asserting themselves as popular services on the Web. A central point is to determine whether two distinct users can be considered similar, a crucial concept with interesting consequences on the possibility to accomplish targeted actions like, for example, political and social aggregations or commercial promotions. In this chapter, the authors propose an approach in order to estimate the similarity of two users based on the knowledge of social ties (i.e., common friends and groups of users) existing among users, and the analysis of activities (i.e., social events) in which users are involved. For each of these indicators, authors draw a local measure of user similarity, which takes into account only their joint behaviours. After this, the chapter considers the whole network of relationships among users along with local values of similarities and combine them to obtain a global measure of similarity. Applying the Katz coefficient, a popular parameter introduced in Social Science research, carries out such a computation. Finally, similarity values produced for each social activity are merged into a unique value of similarity by applying linear regression.*

DOI: 10.4018/978-1-61350-444-4.ch017

## INTRODUCTION

Online social networks like Facebook, My Space, YouTube or LinkedIn are rapidly emerging as one of the most popular services on the Web. These systems are able to capture a significant portion of Web users: for instance as of January 2011, Facebook counts more than 500 millions active users and about 50% of active users log on to Facebook in any given day<sup>1</sup>.

Facebook users are allowed to publish online profiles describing both *demographic data* (e.g., place and date of birth) as well as *interests*. In addition, users may be involved in a large number of social activities like getting in touch with other people and creating friendship relationships with them, create groups with the goal of raising public awareness on political or social themes, sponsoring an event or declaring to participate to it and so on.

A central problem in this scenario is to determine whether two users can be considered *similar*. A tool capable of correctly identifying similar users is advantageous for many purposes. We can, in fact, identify people who share the same political and social ideas and suggest them to form groups in such a way to better promote and plead their causes. We can suggest new possible friendships to users in some way connected by common interests, activities, etc. We can find out in a social crowd, people who can possibly form groups representing a threat for the society because sharing extremist views in particular contexts, such as terrorism, criminal behaviours, etc. We could predict the connections and the interactions, which are likely to occur in the near future among similar users (Liben-Nowell & Kleinberg, 2007). From a commercial standpoint, the identification of groups of users tied by shared interests would be beneficial to promote and diffuse new technologies as well as to advertise commercial products (Kleinberg, 2008).

The problem of identifying the similarity among users has received a strong attention in

many fields of Computer Science (thinks of Recommender Systems (Resnick & Varian, 1997) or User Modelling (Kobsa, 2001)) but it is still largely unexplored in the context of very large social networks like Facebook.

We can put into evidence two research lines devoted to detect similarities between pairs of users. The first research line is based on social relationships (especially friendship relationships) existing among users in order to determine whether they are similar (Geyer, Dugan, Millen, Muller & Freyne, 2008, Spertus, Saham & Buyukkokten, 2005). Similarity derives from two different and competing factors (Crandall, Cosley, Huttenlocher, Kleinberg & Suri, 2008): social influence (Friedkin, 1998), according to which individuals adopt behaviors exhibited by those individuals they interact with, and homophily (Lazarsfeld & Merton, 1954, Mcpherson, Lovin & Cook, 2001), i.e., the tendency of individuals to create relationships with other individuals who are similar to them. Similarity can express along a broad range of dimensions like age, ethnicity, gender, religion and job. Extensive empirical research shows strong evidence of homophily in real contexts (Currarini, Jackson & Pin, 2009); for instance, a study on 12,067 people carried out between 1971 and 2003 indicated that a person has a high chance of being obese if her friends are obese too (Christakis & Fowler, 2007).

In online social networks like Facebook, friendship relationships are still a reliable indicator of similarity between two users but they are not enough. In fact, since the number of users of an online social network is typically huge, if we would select at random a pair of users, there would be a high chance they do not know each other. Selected users would be automatically recognized as not similar. Such a conclusion may be wrong because the two users may share, for instance, the same religious or political convictions and, then, a form of similarity between them could be envisaged.

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/finding-similar-users-facebook/61680](http://www.igi-global.com/chapter/finding-similar-users-facebook/61680)

## Related Content

---

### Blogger Mothers as a Transmediatic Narration: An Examination on Transmediatic Narration Used by Blogger Mothers

Ercan Aktan (2022). *Research Anthology on Usage, Identity, and Impact of Social Media on Society and Culture* (pp. 339-358).

[www.irma-international.org/chapter/blogger-mothers-as-a-transmediatic-narration/308610](http://www.irma-international.org/chapter/blogger-mothers-as-a-transmediatic-narration/308610)

### Still in Fashion?: A Study on Facebook Usage

Alberto Marcuzzo and Thanos Papadimitriou (2013). *International Journal of Virtual Communities and Social Networking* (pp. 42-61).

[www.irma-international.org/article/still-in-fashion/85356](http://www.irma-international.org/article/still-in-fashion/85356)

### Privilege-Seeking Activities in Organizational Politics and Its Effect on More Productive Employees

Gil S. Epstein and Bruce C. Herniter (2012). *International Journal of E-Politics* (pp. 16-30).

[www.irma-international.org/article/privilege-seeking-activities-organizational-politics/65550](http://www.irma-international.org/article/privilege-seeking-activities-organizational-politics/65550)

### Deleveraging Creative Capital: A Decade of YouTube Campaigning

Robert John Klotz (2019). *International Journal of E-Politics* (pp. 1-11).

[www.irma-international.org/article/deleveraging-creative-capital/241303](http://www.irma-international.org/article/deleveraging-creative-capital/241303)

### Flow as a Framework to Engage Youth in Participatory Politics on Social Media Platforms

Sohail Dahdal (2017). *International Journal of E-Politics* (pp. 34-48).

[www.irma-international.org/article/flow-as-a-framework-to-engage-youth-in-participatory-politics-on-social-media-platforms/193211](http://www.irma-international.org/article/flow-as-a-framework-to-engage-youth-in-participatory-politics-on-social-media-platforms/193211)