

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/coseComputers
&
Security

Authorship disambiguation in a collaborative editing environment

Padma Polash Paul ^{a,*}, Madeena Sultana ^b, Sorin Adam Matei ^c,
Marina Gavrilova ^b

^a University of Oxford, Oxford, OX1 3PA, UK

^b Department of Computer Science, University of Calgary, 2500 University Dr NW, Calgary, AB, Canada

^c Department of Communication, University of Purdue, West Lafayette, USA

ARTICLE INFO

Article history:
Available online

Keywords:

Online communication
Authorship disambiguation
Score level fusion
Editing behavior
Cybersecurity
Social behavioral biometrics

ABSTRACT

“Wikipedia”, known as the world’s largest online free encyclopedia, is one of the remarkable examples of crowdsourcing, where millions of articles have been produced by volunteers from all over the world. Wikipedia allows anyone to edit articles without being pre-screened or authorized. A user can edit articles using either a valid ID or an IP address. The freedom of editing using ID or IP makes editorial identities ambiguous. This may affect the integrity of the research process. It also facilitates malicious users to vandalize Wikipedia content. Disambiguation of users’ identity in Wikipedia can assist in distinguishing between trusted and mischievous users and more important in defining authorship in a less ambiguous manner. The present paper introduces a new methodology to ascertain Wikipedia authorship and to reduce ambiguity of user IDs. Our methodology uses the editing activity of users as a distinguishing feature for identifying non-ambiguous profiles. Reducing ambiguity of authorship can facilitate understanding of human behavior in collaborative editing, predicting sock puppetry (duplicate accounts), detecting anomaly, identifying trustworthy as well as offensive authors, and in improving security procedures and research in online social media. Our experimental results indicate that it is possible to disambiguate with high degree of certainty the editorial activity by 75% at rank 1.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Social media, such as Wikipedia, Twitter, Facebook, Reddit, or Instagram, provide new and vast opportunities for communication and security research. Social media sites facilitate instant access to massive datasets storing information about online communities interacting in real time. The information that is stored and processed by social media covers a wide range of interests from entertainment, news, and leisure to business, commerce, and politics. Wikipedia as a user generated site is a prominent social medium offering a rich terrain for research

on human behavior, collaboration, knowledge production, organizational science or crowdsourcing especially in the context of computer security (Fan et al., 2015).

One important issue that can be explored on Wikipedia is anonymity. Aside from ethical or identity-driven concerns, anonymity may render data unreliable and authorship problematic. If the same individuals create multiple accounts, or if they interact behind a computer address (IP address), it is often hard to tell who we are observing and how to aggregate data over time or across units of observation. Deception is also a damaging issue (Avery et al., 2017) since it makes it hard to ascertain claims of vandalism, malicious editing, or attempts

* Corresponding author.

E-mail address: pppaul@live.ca (P.P. Paul).

<https://doi.org/10.1016/j.cose.2018.01.010>

0167-4048/© 2018 Elsevier Ltd. All rights reserved.

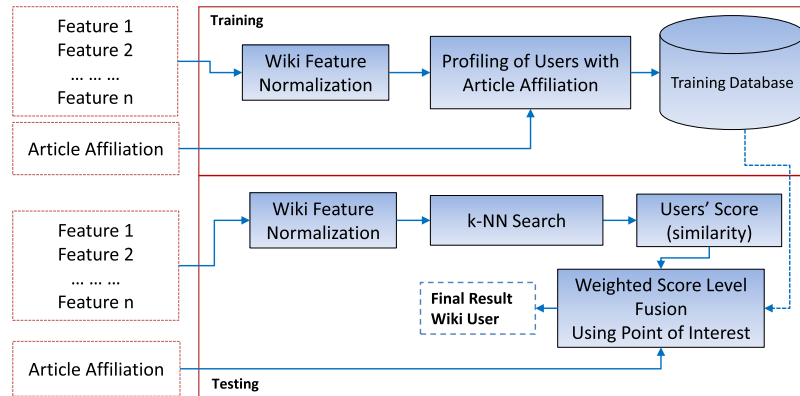


Fig. 1 – The basic building blocks of the proposed framework of authorship disambiguation.

to manipulate the system. Simple tasks, such as determining the degree to which certain individuals contribute to online communication spaces, are undermined by the chronic uncertainty about the basic units of analysis that are observed. Political or cultural polarization research becomes equally difficult (Song, 2015). When actors are anonymous or when they use pseudonyms, it is hard to tell how many individuals are in each camp or if the declared affiliations are the real ones. Finally, research on online identity is equally impacted by the anonymous nature of online interactions. When we are uncertain about the reliability of basic actor data, it is hard to determine who created certain identities and how identity building works.

Anonymity is a crucial research issue because although some social media, such as Facebook, demand at least in principle if not always that users use their real names, most social media operate under the veil of anonymity. Wikipedia, Twitter, Reddit, or Instagram handles (i.e. user names) can be completely arbitrary and the same individuals may use multiple accounts. The same account can also be used by multiple people, especially in an institutional or PR context. Some social media activities do not require any kind of authentication. For example, most (although not all) Wikipedia pages can be safely edited without logging in. Research of the entire Wikipedia editorial history between 2001 and 2010 suggests that over 85% of the unique editors are in fact anonymous users who never log in and can only be identified by their IP (computer) addresses (Matei and Bruno, 2015).

There is a pressing need to develop better research strategies for increasing data reliability by identifying patterns of activity that may be associated with clusters of registered identities. Simply put, we need to examine the registered identities of social media users and cluster them into potential “real” identities. This is especially important in researching sites that are used as reference information sources, such as Wikipedia. Being one of the top ten most visited sites, Wikipedia is one of the most influential sources of information in the world (Wikipedia). Knowing who and to what degree changes specific articles and to what degree the project is influenced by what group of individuals is of crucial importance. Limitations imposed by anonymity on research may go beyond Wikipedia. Editorial activity on sites like Reddit, which has become one of the most popular news sources, is also anonymous. A method of

identifying association between Twitter accounts based on their users’ social interactions has recently presented in Sultana et al. (2017a, 2017b). User clustering based on behavioral patterns in online social media can also assist in situation awareness (Kandias et al., 2017) and risk assessment (Farahmand and Spafford, 2010) as well as in differentiating between legitimate users and spammers (Fu et al., 2018).

Our research addresses the ambiguous authorship problem in a collaborative editing environment, Wikipedia. We use an editorial-behavioral and security perspective. We propose an editing behavior “fingerprinting” method, which relies on editing style and editing quality originating from the same author (contributor, user ID). An overview of the proposed method is presented in Fig. 1. We use a search approach widely used in information sciences, namely, k-nearest neighbor (k-NN) (Thomas and Hart, 1967). The reason for using k-NN search is that given a test editing profile we wanted to rank editing profiles from most to least similar, which facilitates combining some of them on the basis of a similarity threshold. In addition, the fundamental k-NN search will be used to show that the selected feature set is good enough for authorship disambiguation without applying any advanced classification techniques. The final goal of choosing this method is to create valid user profiles across editorial interventions of unknown users. Our method utilizes similarity of user activity to disambiguate identity. We use both testing and training profile datasets based on certain editing features. The test profile includes edits of registered users, whose identity is not given to the system. The training profile dataset includes all edits that were made by registered users along with their identities. Pattern recognition is operated on both groups, the aim being to show that the identification method used for the unknown users is just as good as that for editorial activity committed by known (registered) users. A final ranked list of similarity between test and training datasets is obtained by weighting user profiles based on article affiliations. The data thus classified will be used to test the effectiveness of the disambiguation method and to demonstrate the impact that the process of identification may have on future security research.

The methodology we propose contributes in several ways to security research. First, we utilize editing behavior for the first time to disambiguate Wikipedia users identity. Second, we identify “article affiliation” as a disambiguation feature and

Download English Version:

<https://daneshyari.com/en/article/11002558>

Download Persian Version:

<https://daneshyari.com/article/11002558>

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