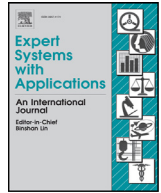




ELSEVIER

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Flash mobs, Arab Spring and protest movements: Can we analyse group identities in online conversations?

Natalia Criado^{a,*}, Awais Rashid^b, Larissa Leite^b^a King's College London, London, WC2R 2LS, United Kingdom^b Security Lancaster, Lancaster University, Lancaster LA1 4YW, UK

ARTICLE INFO

Article history:

Received 10 February 2016

Revised 11 April 2016

Accepted 12 June 2016

Available online 16 June 2016

Keywords:

Social identities

Online social media

Natural language processing

ABSTRACT

The Internet has provided people with new ways of expressing not only their individuality but also their collectivity i.e., their group affiliations. These group identities are the shared sense of belonging to a group. Online contact with others who share the same group identity can lead to cooperation and, even, coordination of social action initiatives both online and offline. Such social actions may be for the purposes of positive change, e.g., the Arab Spring in 2010, or disruptive, e.g., the England Riots in 2011. Stylometry and authorship attribution research has shown that it is possible to distinguish individuals based on their online language. In contrast, this work proposes and evaluates a model to analyse group identities online based on textual conversations amongst groups. We argue that textual features make it possible to automatically distinguish between different group identities and detect whether group identities are salient (i.e., most prominent) in the context of a particular conversation. We show that the salience of group identities can be detected with 95% accuracy and group identities can be distinguished from others with 84% accuracy. We also identify the most relevant features that may enable mal-actors to manipulate the actions of online groups. This has major implications for tools and techniques to drive positive social actions online or safeguard society from disruptive initiatives. At the same time, it poses privacy challenges given the potential ability to persuade or dissuade large groups online to move from rhetoric to action.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Global and national events over recent years have shown that online social media can be a force for good (e.g., Arab Spring in 2010) and harm (e.g., the England Riots in 2011). In both of these examples, social media played a key role in group formation and organisation, and in the coordination of the group's subsequent collective actions (i.e., the move from rhetoric to action) (Halliday, 2011; Tufekci & Wilson, 2012). Such coordinated actions are possible because individuals identify themselves with a particular social group or with an ideal (Taylor, Whittier, & Morris, 1992). Online identity in such contexts is, therefore, not so much about the categorisation of the self as a singular "I". Instead it is the conception and expression of group affiliations as a more inclusive "we".

This paper focuses on these online group identities. Offline group identities are usually referred to as *social identities* by social identity theory (Deaux, 1996; Stryker & Burke, 2000; Tajfel, 2010),

a social psychological theory that sets out to explain group processes, intergroup relationships and the social self. Social identity is the individual's explicit or implicit expression of belonging to certain social group, together with some emotional and value significance to him/her of the group membership (Tajfel, 2010). Thus, a person has not one "personal self" but rather multiple social identities that are culturally contingent and contextual (Hankin, 2013). The *salient identity* is the identity that comes into play and is invoked in a specific situation or context (Stryker & Burke, 2000). Thus, a social identity is salient when it is invoked across a group of persons who perceive themselves as members of a social group. Which identity becomes salient in a given situation depends on factors such as the level of commitment of a person to a particular identity. One component of commitment is the number of others with whom one is connected by possessing a particular identity. Thus, when a person shares a certain identity with a greater number of people, his/her commitment to that identity tends to be higher and this identity is likely to be more salient (Stryker, 1980).

Given the importance of online social media in orchestrating and coordinating large-scale group mobilisations—from democracy

* Corresponding author.

E-mail addresses: natalia.criado@kcl.ac.uk (N. Criado), a.rashid@lancaster.ac.uk (A. Rashid), larissa.leite@gmail.com (L. Leite).

and protest movements to hacktivist groups through to riots and extreme right wing marches— group identities are of key interest to a variety of stakeholders. They can be: mobilised as a resource for positive social change; studied to understand and counteract organised online actions that may compromise the safety and security of citizens; and even potentially be harnessed to build resilience in individuals and groups to limit the harmful effects of government or extremist efforts to disrupt online group formation and subsequent mobilisation.

Of course, group identities are not the only variable that predicts behaviour, but they can provide a guide to likely behaviours—as stated by social identity theory the higher the salience of a social identity (i.e., the identification with a particular group), the greater the individual's willingness to contribute to the social action (Bagozzi & Dholakia, 2002; Stryker & Burke, 2000). Social identities have been shown to influence behaviour in many domains, including politics (Jackson & Smith, 1999), protest movements (Reicher, 1996) and fan behaviour (Platow et al., 1999). Knowing how salient is group identity can lead to predictions of how much the identity will influence the individuals' beliefs, emotions and actions. Since the activation of a social identity affects the way people think as well as their feelings and behaviours, our hypothesis is that such group identities also affect the way in which people communicate online. As such, our model characterises text-based online communications in terms of a set of textual features such as their language, their style and their interaction patterns (i.e., the way in which users interact). We then study the features that can best distinguish between different group identities online as well as those features that can indicate the salience, or lack thereof, group identities. We address research questions categorised as follows:

1. Detecting salience of group identities:
 - (a) Do group identities manifest in online conversations, i.e., is it possible to use textual features to automatically detect the presence of salient group identities?
 - (b) Is our analysis model generalizable to detect identity salience across different group identities and on different online social media?
 - (c) Which features are most suitable for detecting identity salience?
2. Distinguishing group identities:
 - (a) Is it possible to distinguish between different group identities on the basis of textual features automatically extracted from conversations?
 - (b) Is our analysis model generalizable to distinguish group identities over time and on different online social media?
 - (c) Which features enable a specific group identity to be accurately predicted?

Our evaluation shows that, by using a range of structural, grammatical, semantic, categorical and stylistic features, our model can detect the salience of group identities with 95% accuracy and distinguish between group identities with 84% accuracy. In general, our study reveals that there is much more valuable information available on social media than just personal data. We identify features of online conversations that can reveal important dynamics of online groups and, hence, potential drivers for mobilisation of such groups. Notwithstanding the importance of protecting personal data on online social media (Anthonysamy, Greenwood, & Rashid, 2013; Madejski, Johnson, & Bellovin, 2011), it is also important to study and understand how group identities are formed and could be exploited for positive or negative ends. While the former has the potential to adversely affect individuals, the latter has major implications for social action/inaction in our modern digital society.

The novel contributions of this paper are fourfold:

1. This is the first paper to propose a model to analyse online group identities based on social identity principles.
2. We use textual features to detect group identity and its salience. In contrast with other works that study the difficulties people encounter when interacting with heterogeneous groups in an online social network, e.g., (DiMicco & Millen, 2007), or how online identities are constructed and shaped, e.g. (Zhao, Grasmuck, & Martin, 2008), all the features analysed in our model are extracted fully automatically, i.e., no human intervention is required.
3. We demonstrate that group identities and their salience manifest themselves, with a high degree of accuracy, in text-based online communications through a range of structural, grammatical, semantic, categorical and stylistic features.
4. Our results open up key privacy challenges for the research community at large with regards to the potential exploitation of group identities to persuade or dissuade large groups online to move from rhetoric to action. We have implemented an online tool that enables study of features underpinning online group identities in order to investigate these challenges. We identify which features can put online groups at most risk of such manipulation by mal-actors so as to build resilience against such out-group influences.

The rest of the paper is organised as follows: Section 2 describes related work. Section 3 presents our model for analysing group identities including the features and the classifiers used in the analysis. Section 4 describes experiments that validate our model including the datasets used and the results obtained. We discuss the implications of our model and experiments in Section 5. Finally, Section 6 concludes the paper and identifies directions for future work.

2. Related work

Within the Artificial Intelligence field different computational models have been proposed to represent social identities. One of the most cited models is the ABIR (Agent-Based Identity Repertoire) model (Lustick, 2000), which seeks to refine, elaborate, and test theories of identity and identity shifts. This model has been used in agent-based simulations to analyse the emergence (Rousseau & Van Der Veen, 2005) and dynamics (Smaldino, Pickett, Sherman, & Schank, 2012) of social identities offline. To the best of our knowledge, ours is the first model for the automatic analysis of group identities invoked on different online social media.

There are empirical proposals, as ours, that draw conclusions about identity from information extracted from online social media. DiMicco and Millen (DiMicco & Millen, 2007) describe a study about the way in which people present themselves (i.e., the way in which people invoke their identities) on Facebook. Specifically, the authors analysed Facebook profiles and interviewed employees belonging to the same company with the aim of understanding how they managed their identity when interacting with different social groups (e.g., family, friends from school, workmates, etc.) on Facebook. The main contribution of their study was the identification of the difficulties that people encounter when interacting with heterogeneous groups using the same online social network; and the identification of the need for more sophisticated controls that help one to manage one's identities online. Similarly, Zhao et al. (Zhao et al., 2008) analysed Facebook profiles of students in a university to study how these students presented themselves on Facebook. They focused on how the online identities of these persons were "built" on Facebook. An interesting conclusion of their study is that identities are usually claimed implicitly on Facebook (e.g., people express that they belong to a group of friends by posting pictures with these friends instead of writing it in their self-description).

Download English Version:

<https://daneshyari.com/en/article/383552>

Download Persian Version:

<https://daneshyari.com/article/383552>

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