Chapter 4
Welcome to the Party:
Modeling the Topic Evolution of Political Parties in Microblogs Using Dynamic Topic Models

Kai Heinrich
TU-Dresden, Germany

ABSTRACT
Modeling topic distributions over documents has become a recent method for coping with the problematic of huge amounts of unstructured data. Especially in the context of Web communities, topic models can capture the zeitgeist as a snapshot of people’s communication. However, the problem that arises from that static snapshot is that it fails to capture the dynamics of a community. To cope with this problem, dynamic topic models were introduced. This chapter makes use of those topic models in order to capture dynamics in user behavior within microblog communities such as Twitter. However, only applying topic models yields no interpretable results, so a method is proposed that compares different political parties over time using regression models based on DTM output. For evaluation purposes, a Twitter data set divided into different political communities is analyzed and results and findings are presented.

INTRODUCTION
It is clear that every day, we are becoming increasingly enmeshed in the Web. Millions of individuals are communicating with each other using essentially a few clicks of a mouse. These discourses, supported by a growing number of online media–forums, Wikipedia, Twitter, Facebook and blogs–are engaging millions of individuals globally. Within the political and social realm, a recent study reports that close to a fifth of US Internet users have posted online or used a social networking site for civic or political engagement (Dugan & Smith, 2013). Another study found that 55% of the adult US population went online in 2008 in order to get involved in the political process or to seek information about the last US election (Ha Thuc, 2011). A key characteristic of social data is that it is generated by a large number of people from different cultures, locations, age groups, religions, income categories etc. This is in contrast to the case of traditional media such as

DOI: 10.4018/978-1-4666-6639-9.ch004
newswire generated by a small group of journalists. Because of its popularity and diversity, social data is an excellent source for understanding the perspectives of large groups of people.

The Analysis of social data from Facebook, Twitter, LinkedIn or other sources is crucial in order to gain information about the social behavior of people in online communities (Aarts, Schraagen, van Maanen, & Ouboter, 2012). The research that is done here, while broadly applicable, will be focused on the microblogging platform Twitter. There are various information that can be gained by analyzing Twitter. Examples include demographic analysis, social network analysis, as well as message analysis. Focusing on the message characteristics, it is important to capture user behavior and the dynamics of topics within online social networks.

Along with the exponential growth of text data on the Web, particularly of the user-generated content, comes an increasing need for hierarchically organizing documents, retrieving documents accurately, and discovering evolutionary trends of various popular topics from the data. However, all of these are challenging due to the diversity, heterogeneity, noisiness and time-sensitivity of Web 2.0 data.

While there is a variety of tasks that can be dealt with when facing social media analysis, this chapter focuses on the message characteristics to reflect on the behavior of people within a social network. Other topics in this area of Twitter analysis include topological and geographical properties (Hong & Davison, 2010), novelty and opinion leader detection (e.g. Bi, Tian, Sismanis, Balmin, & Cho, 2014), this work aims to provide a framework for evolution monitoring in social networks. Therefore we will examine related work in the fields of Twitter analytics and topic models in the following background section. As an output of the literature review concerning topic modeling in Twitter, we define our research goal, as well as certain sub problems, that we have to deal with in order to reach that goal.

BACKGROUND

Social Behavior in Online Networks

In the research field of social behavior in online networks, (Aarts et al., 2012) found that three specific research tasks can be identified: actor characteristics, message characteristics and network characteristics.

Actor characteristics determine to what extent individuals are part of the network and how much influence the individual actor has to spread the information further. Evidently, actors form the building blocks of a social network. It is between these actors that communication or information exchange may or may not occur.

Message characteristics form a part of the explanation to the extent to which information circulates through the social network. Message characteristics play a role in the distribution and propagation of information. An attractive or stimulating message is more likely to spread to various individuals within the network. Unappealing or uninformative messages are likely to die out quickly.

Network characteristics give us insight in the relational features between actors instead of features confined to one actor only. The extent to which information spreads is not solely determined by either the characteristics of the actor or the message characteristics. The social context or network (who is connected to whom, and in which way) to a large extent determines the flow and diffusion of information. This relational approach is a distinctive feature of a social network approach.

The research in the area focusing on message characteristics includes topic trend research (Lau, Collier, & Baldwin, 2012; Perkiö, Buntine, & Sami Perttu, 2004; Wang & McCallum, 2006), as well as retweet mechanics (Liu, Niculescu-Mizil, & Gryc, 2009). The pure motivational aspects of posting tweets and dealing with social communication in Twitter, e.g. the spread of rumors have also been dealt with (Java, Song, Finin, & Tseng, 2007).
Related Content

Hand Gesture Recognition Using Multivariate Fuzzy Decision Tree and User Adaptation
[www.igi-global.com/article/hand-gesture-recognition-using-multivariate/55994?camid=4v1a](www.igi-global.com/article/hand-gesture-recognition-using-multivariate/55994?camid=4v1a)

Evolution of e-Sales as A Form of e-Entrepreneurship in Poland: An Analysis of Opportunities and Threats

Ubiquitous Mediation and Critical Interventions: Reflections on the Function of Signs and the Purposes of Peirce’s Semeiotic
[www.igi-global.com/article/ubiquitous-mediation-critical-interventions/56444?camid=4v1a](www.igi-global.com/article/ubiquitous-mediation-critical-interventions/56444?camid=4v1a)

Hybrid Approach for Analyzing Acute Spots of Clinical Speech Data Using Fuzzy Inference System