Time-aware online reputation analysis

Peetz, M.-H.

1 Introduction

Before social media emerged on the internet, most of the communication was personal and mass communication was left to journalists, politicians, companies, or public figures. Social media now enables everyone to communicate about anything to everyone [250]. Filtering this flood of information for relevancy is vital. Unlike traditional written media, social media is transient, similar to personal conversations. Do we really remember what a great-aunt said at the christmas dinner five years ago? Do we really care what our friends were posting on Facebook five months ago? As personal communication looses importance over time, so do mass-communicated conversations. But social media is dynamic as well: similar to everyday conversations it is prone to influences by events that affect us. Real-life conversations are real-time: the moment something happens, people talk about their thoughts, opinions, and events in their life. This is reflected in social media conversations as well. However, in everyday conversations, one person's word of mouth influences maximally 10 people, while using online social media, one may reach 10 million people [91].

The availability of this data is exciting for social media analysts and market researchers: the immediateness between publishing emotions, opinions, and feelings while using a product in daily life promises an understanding of customer satisfaction and a brand's reputation in real-time. Here, social media functions as a proxy to understand society's opinion on a certain brand. However, the amount of data is increasing and an all-covering manual inspection of relevant conversations ceased to be possible very quickly. In this thesis we develop tools to understand the online reputation of a company making use of the dynamic nature of social media.

Reputation

The reputation of a company is an integral part of its value [254]. In fact, a common hypothesis in business analytics is that the reputation of a company has a direct correlation with revenue [85, 234] and can act as a buffer from economic loss [123]. Why is that? A bad financial reputation (e.g., due to a bad financial report) shies away investors. A company with a bad reputation for the workspace environment shies away human capital. A bad reputation for customer relations takes the wrath of potential and standing customers upon them. Measuring reputation, however, is a difficult problem. In earlier days, media analysts followed public opinion based on mentions in newspapers and polls [254].

They formed an opinion on the reputation. Seminal work by Van Riel [253] and van Riel and Fombrun [254] adds a structural approach to the meaning of reputation, providing companies with a framework to measure reputation along different dimensions. Within this framework it is feasible to analyse newspaper and poll data manually. Based on this analysis, reputation used to be *manageable* with various means: dedicated advertisement campaigns like the responses towards health issues of smoking [263], or campaigns to gain back consumer trust [158].

With the arrival of social media into people's lives, media analysts have a greater pool of data to base their reputation analysis on. In the early days of social media when it mostly constituted of the blogosphere, this was still manually feasible [57]. *Microblogs* (posts on Twitter, Facebook, etc.) are often very close to the moment the customer has contact with the brand or product [105]: feedback is more immediate and emotionally attached than the thought out analysis often found in personal blogs (even though blogs have a higher conversion rate [70]). Additionally, more people publish microblogs than personal blogs,[1] allowing media analysts to be reached from a broader spectrum of people. However, this comes at a price: with 500 million tweets published per day[2] and 757 million daily active Facebook users [78], manual analysis loses its feasibility.

Online reputation analysis (ORA) is a discipline directly following from this problem: here, the use of computational tools allows for filtering relevant tweets and an approximate analysis of the reputation of an entity. Early ORA started with counting occurrences of a brand name in social media, therewith estimating the *knowledge/reach* of a brand. This early research in this area concerned the reputation of politicians [247] or company names [5]. Furthermore, ORA tries to estimate the *sentiment* in a sentence that mentions the brand [117] and aggregates this sentiment to measure the overall reputation. However, just using sentiment is an inaccurate proxy to measure the reputation a tweet has on the reputation of an entity [6]. In fact this has been one of the key insights at RepLab [6, 7], a workshop and challenge that focusses on classifying tweets for their influence on the reputation of an entity (in short: *reputation polarity* of a tweet) directly [6]. *Online reputation management* uses findings from ORA to manage the reputation, to directly interact with customers via Twitter (web care) or hype positive aspects of a brand [154].

This thesis focuses on online reputation analysis. We show that filtering works very well with manual assistance and that the key to estimating reputation polarity is not only the overall sentiment, but also the textual, as well as metadata features that depend on the entities.

Time in Social Media

Reputation is dynamic and changes over time [253, 254]. But also in social media analysis, time proves to be an important aspect (see Chapter 2). The content is user-generated and therefore dynamic: Real life events impact events in social media [152, 213, 247, 262], and more and more events in social media impact events in real life [47, 251].

[1] 500 million tweets per day [248] 38.0 million new word press posts per month, 1.26 million per day [275]. Wordpress constitutes 43% of the market [41].

[2] August 2013.

The past is often reflected in personal blogs, where people talk about their experiences, they often function as a diary. Diary entries are influenced by current events: A recipe blog around super bowl time, will likely feature recipes related to superbowl snacks. The moods and feelings can be monitored [165] and they clearly change over time.

With the introduction of mobile devices in our life, social media becomes more immediate. In Twitter, users post about their present status [170] more than about anything else. This can be used to track and monitor current topics like the trending topics in Twitter.[3] However, it can also be used to monitor real-time events such as earthquakes [213], flu pandemics [137], or revolutions [152]. This knowledge has been used to predict the outcome of elections [247] or the stock market [32].

What does this mean for ORA? First of all, timeliness of responses to changes in reputation is of utter importance. This allows the analyst to manage and prevent the opinion to tip over. Secondly, new topics related to a brand emerge. Using algorithms that are blind to temporal changes, we cannot identify content relevant to emerging topics. Finally, the reputation of a brand is not static. Negative and positive aspects are being forgotten over time. In this thesis we find events in user-generated content, and make use of priors to filter social media according to their age.

1.1 Research Outline and Questions

This thesis addresses the following question: *How can we estimate the reputation of a brand automatically?* Let us summarize this thesis in a few sentences. Before we can proceed with the estimation of reputation polarity we need to understand what reputation polarity is (**RQ1.1–1.2**). We then develop an algorithm and a setting to estimate the reputation polarity (**RQ2.1–2.3**). Estimating reputation polarity only works if we can filter relevant material for training. We propose to use burst modeling (**RQ3.1–3.5**) to estimate correct query terms to find relevant documents in a collection, and use active learning to improve filtering of streaming data (**RQ4.1–4.2**).

In more detail, as a prerequisite to estimating reputation automatically, we need to understand reputation and the ingredients used by social media analysts. In Chapter 4 we want to understand the procedures and the features used to annotate the reputation polarity of tweets. In particular we ask

RQ1.1 What are the *procedures* of (social media) analysts in the analysis and annotation of reputation polarity?

RQ1.2 On a per tweet level, what are the indicators that (social media) analysts use to annotate the tweet's impact on the reputation of a company?

We find that analysts use indicators based on the (topical) authority of the author of the tweet to estimate a tweet's reputation polarity. This authority is based on online data and offline experience. They also look at the reach of a tweet. Based on this information in Chapter 5 we apply the extracted features and procedures on social media data (in particular Twitter). We identify three main settings in which a reputation polarity estimator

[3] `https://support.twitter.com/articles/101125-about-trending-topics`

can be trained: entity-independent, entity-dependent, and domain-dependent. Entity-independent training creates *one* model for all tweets, while entity-dependent training creates a models for every entity. For domain-dependent training, we create a model for each domain.

RQ2.1 For the task of estimating reputation polarity, can we improve the effectiveness of baseline sentiment classifiers by adding additional information based on the sender, message, and receiver communication model?

RQ2.2 For the task of estimating reputation polarity, how do different groups of features perform when trained on entity-(in)dependent or domain-dependent training sets?

RQ2.3 What is the added value of features in terms of effectiveness in the task of estimating reputation polarity?

In Chapter 4 we also identified that finding and filtering the right amount of information manually should be automated. Additionally, estimating the reputation polarity only works if we extract the right documents [230]. Therefore, in the second part of the thesis, we analyse how to retrieve and filter documents using temporal knowledge. In Chapter 6 we turn our attention to time and how to retrieve documents, in particular blogs, that have a temporal information need. For those queries, we ask:

RQ3.1 Are documents occurring within bursts more likely to be relevant than those outside of bursts?

RQ3.2 Can documents within bursts contribute more useful terms for query modeling than documents selected for relevance models?

RQ3.3 What is the impact on the retrieval effectiveness when we use a query model that rewards documents closer to the center of the bursts?

RQ3.4 Does the number of pseudo-relevant documents used for burst detection matter and how many documents should be considered for sampling terms? How many terms should each burst contribute?

RQ3.5 Is retrieval effectiveness influenced by query-independent factors, such as the quality of a document contained in the burst or size of a burst?

Watching social media analysts during their annotation, we learnt that, apart from temporal episodes (or bursts), recency of tweets is very important as well. Recency is defined by our own perception and directly related to memory. Psychologists model this recency with retention models. We can incorporate the retention models as recency priors and assess their impact on retrieval performance and other requirements. We ask:

RQ4.1 Does a prior based on exponential decay outperform other priors using cognitive retention functions with respect to effectiveness?

RQ4.2 In how far do the proposed recency priors meet requirements, such as efficiency, performance, and plausibility?

In Chapter 4, we found that social media analysts would like to be a part of the annotation and filtering process. This ensures a certain level of quality control. In Chapter 8 we then adjust the filtering task introduced at RepLab 2013 [7] to a streaming scenario and we propose an active learning approach. We introduce a baseline and ask:

RQ5.1 For the entity filtering task, does margin sampling improve effectiveness over random sampling, i.e., is it a strong baseline?

As entities and their description change over time, we also look at how we can incorporate temporal aspects into filtering algorithms. Based on the temporal recency priors from Chapter 7 we ask in Chapter 8:

RQ5.2 For the entity filtering task, does sampling based on recency priors and margin sampling together, outperform margin sampling with respect to F-score?

In Chapter 6 we found that burst detection works well on retrieving social media data. We propose two temporal reranking approaches, one based on bursts and one on the publication date. We would like to know:

RQ5.3 For the entity filtering task, does temporal reranking of margin sampled results based on bursts or recency, outperform margin sampling with respect to F-score?

We find that active learning is a feasible approach that needs very few documents to be annotated. The right kind of active learning depends on the entity.

Research questions **RQ1.1** and **RQ1.2** are addressed in Chapter 4. **RQ2.1**–**2.3** are answered in Chapter 5, and **RQ3.1–3.5** in Chapter 6. Finally, **RQ4.1–4.2** and **RQ5.1–5.3** are discussed in Chapter 7 and Chapter 8, respectively.

Having formulated our research questions, we list our main contributions below.

1.2 Main Contributions

This thesis contributes on different levels, we provide new *models*, new *analyses*, and use-case oriented new *task scenarios* We can summarise the two aspects as follows:

Models

- Effective algorithms to measure the impact of a tweet on the reputation polarity.
- Effective algorithms to model queries according to temporal bursts in the result set of that query.
- Bursts models using probability distributions that emphasize different time periods within the bursts.
- New and effective temporal priors based on cognitive retention models.
- Candidate selection models for active learning that sample tweets that need to be annotated according to recency or temporal bursts.

Task Scenarios

- An entity-dependent task for the estimation of reputation polarity using time based split evaluation.
- A temporal active learning scenario for entity filtering using passive training and testing data.

Analyses

- An analysis of the state of automatic online reputation analysis and anchoring this knowledge in the current economic and buisiness theories.
- A user study with social media analysts and finding the procedures and features used to determine the reputation of a company. Additionally, we provide a very new perspective on the concept of reputation polarity by understanding how experts annotate multimedia data.
- An analysis of the effectiveness and impact of different features in different scenarios to annotate reputation polarity. Based on this analysis we can give direct input into the creation of reputation polarity tasks that simulate the actual workflow of social media analysts.
- An analysis of the effectiveness of the query modelling approaches on news and social media data.
- A discussion of advantages and disadvantages of state of the art cognitive retention models.
- An analysis of the filtering task with respect to (temporal) active learning.

1.3 Thesis Overview

This thesis is organised in 9 chapters.

Chapter 2

In this chapter we introduce related work for both temporal information retrieval in social media and current approaches to online reputation analysis. The latter is discussed from the computer science angle as well as from the business angle.

Chapter 3

In this chapter we introduce the benchmark datasets used in the remainder of the thesis.

Chapter 4

In this chapter we investigate how social media analysts determine reputation polarity of individual media expressions and which factors they use to assess their polarity.

To this end we analyse the annotation process of 15 social media experts annotating 331 media expressions to determine the reputation polarity of companies in 19 sectors. Three different types of data are analysed: (i) questionnaire data, about the actual steps in the process as well what indicators social media analysts take into account for the annotation; (ii) log data of the tool used by the analysts; and (iii) videos of analysts following the thinking out loud protocol during the annotation process. We find new features that can be used for automatic estimation of reputation polarity and we find which parts of the annotation process should be automated.

Chapter 5

Here we use features that proved important in Chapter 4 to automate the classification of reputation polarity. Additionally, we introduce a new feature that measures the impact of a tweet. We use alternative training and testing settings, allowing for entity-dependent, entity-independent, and type-dependent training on static data.

Chapter 6

Chapter 4 shows analysts desire the filtering of media to be automated. Query modeling can be used to filter and search for documents according to a specific query or entity [38]. In this chapter we present an approach to query modeling that leverages the temporal distribution of documents in an initially retrieved set of documents. In news-related document collections such distributions tend to exhibit bursts. Here, we define a burst to be a time period where unusually many documents are published. In our approach we detect bursts in result lists returned for a query. We then model the term distributions of the bursts using a reduced result list and select its most descriptive terms. Finally, we merge the sets of terms obtained in this manner so as to arrive at a reformulation of the original query.

Chapter 7

Chapter 6 introduces a retrieval approach that uses query modeling based on a time period, defined by a burst. A different feature of social media is the desire to prioritise recent data. To that end, temporal document priors are often used to adjust the score of a document based on its publication time. In this chapter, we consider a class of temporal document priors that is inspired by retention functions considered in cognitive psychology; such functions are used to model the decay of memory. We introduce a requirement framework consisting of efficiency, performance, and cognitive plausibility, the priors need to follow.

Chapter 8

This chapter combines several temporal ideas from the earlier chapters with an active learning paradigm for entity filtering. We introduce a scenario to use passive entity filtering data in a streaming setting. We show the feasibility of margin sampling as a strong baseline. We introduce various temporal extensions to this baseline: temporal recency priors from Chapter 7 and reranking based on bursts from Chapter 6 and recency.

Chapter 9

This chapter concludes the thesis. We revisit the research questions introduced earlier and answer them. We look forward and formulate open questions in automatic online reputation analysis and temporal information retrieval.

Chapter 5 and 8 depend on Chapter 4, but can be read independently from another. Chapter 8 depends on Chapter 7, but can be read independently. Chapter 6 and 7 can be read independently from the others in general.

1.4 Origins

The material of this thesis was previously published in various conference and journal publications:

- Chapter 4 is based on Peetz et al. [191].
- Chapter 5 is based on Peetz et al. [190].
- Chapter 6 is based on Peetz et al. [187, 188].
- Chapter 7 is based on Peetz and de Rijke [183].
- Chapter 8 is partially based on Peetz et al. [189].

Additionally, Chapter 5 and Chapter 8 were inspired by the RepLab submissions [185] and [189]. The knowledge and intuitions on temporal IR were developed while working on [38, 204], and while working on several demos [92, 174, 186]. General insights to academic work were gained in [22, 150, 184].