

Adaptive Faceted Ranking for Social Media Comments

Elaheh Momeni¹, Simon Braendle¹, and Eytan Adar²

Faculty of Computer Science, University of Vienna¹

Dept. of Information Science, University of Michigan²

momeni@cs.univie.ac.at, braendle@cs.univie.ac.at, eadar@umich.edu

Abstract. Online social media systems (such as YouTube or Reddit) provide commenting features to support augmentation of social objects (e.g. video clips or news articles). Unfortunately, many comments are not useful due to the varying intentions of the authors of comments as well as the perspectives of the readers. In this paper, we present, a framework and Web-based system for adaptive faceted ranking of social media comments, which enables users to explore different facets (e.g., subjectivity or topics) and select combinations of facets in order to extract and rank comments that match their interests and are useful for them. Based on an evaluation of the framework, we find that adaptive faceted ranking shows significant improvements over prevalent ranking methods, utilized by many platforms, with respect to the users’ preferences.

Demo: <http://amowa.cs.univie.ac.at:8080/Frontend/>

1 Introduction

User-generated comments are a vital part of the social media ecosystem. Comments provide a way for participants to “evolve” social media objects — ranging from YouTube videos, SoundCloud audio to more classic news articles. Unfortunately, most comment presentation systems are simple temporal streams that contain a diversity of focus, usefulness, and quality (with many comments being abusive or off-topic). Worse, due to the substantial number of comments on media objects with popular topics, identifying useful comments is often time-consuming and challenging. Without a mechanism for end-users to disentangle comment streams and identify those likely to be of interest, it is easy to imagine most end-users being overwhelmed and disappointed by their experience.

Automatic ranking of comments by “usefulness” is generally complex, mainly due to the subjective nature of usefulness [1]. The simplest method to provide ranking is wisdom-of-the-crowd approach (crowd-based ranking), which allows all users to vote on or rate comments. However, this strategy avoids an explicit definition of usefulness and voting is influenced by a number of factors (such as the “rich get richer” phenomenon) that may distort accuracy. Alternative relevant approaches propose topic-based browsing for micro-blogging platforms [2]. However, as comments have multiple explicit dimensions (such as language tone, physiological aspects, etc), grouping them exclusively based on topic, results in

a single imperfect faceted ranking which does not allow users to rank comments with regard to other potentially useful facets.

This work proposes a framework for enabling faceted ranking on comments attached to social media objects (such as comments on an online video or a news article). Our goal is to help users explore the comment space by offering facilities to extract a set of semantic facets dynamically and to adapt the ranking of comments on the fly for finding useful comments according to the users' preferences.

2 Proposed Framework

The proposed framework consists of four main components. First, the *Semantic Enrichment Component* enriches each comment along various semantic facets when an end-user requests an adaptive faceted ranking of comments on a media object. The system provides three types of semantic facets: (1) **Topic-related facets**, topics discussed within comments on a media object. The proposed framework extracts named entities to identify topics and as extracted named entities can be ambiguous, additional fine tuning is required. (2) **Subjective facets** such as comments with subjective tone, highly affective language, offensive oriented, sad oriented (by utilizing LIWC [3]). (3) **Objective facets** such as informative, video timestamp, religion referenced. Second, the *Facet Extraction and Selection Component* operates on semantically enriched comments and clusters comments along multiple explicit semantic facets. For clustering purposes, we utilize the centroid clustering method on enriched comments. It then extracts a set of facets, selects a list of proposed facets dynamically (using the Greedy Count algorithm). Third, the *Ranking Component* enables the user to explore and select a combination of facets, and ranks comments accordingly. Finally, the *Feedback Collector and Optimization Component* collects implicit (browsing behavior of the end-user) and explicit (explicit voting by end-users on comments) feedback from the end-user. This feedback facilitates the evaluation of the proposed framework and furthermore, the personalized selection of facets.

User-Experience of Faceted Ranking in Commenting Systems The interface (see Figure 1) of the proposed framework¹ consists of two parts, one for displaying the facets (on the left), and one for displaying the ranked list of comments (on the right). A user of the system can perform the following actions: (1) She can enter a media object ID, this triggers the system to crawl all comments related to the media object, semantically enrich each comment, cluster the comments into different semantic facets, and finally present a list of facets and topics on the left side of the interface. (2) She can select combinations of proposed facets based on her preferences, this triggers the system to present a ranked list of comments based on the chosen facets. (3) She can browse ranked comments and vote whether the comments match her interests and are relevant to her

¹ The development of the backend of the interface uses the REST style, permitting the interface to be easily integrated in any social media platform <http://amowa.cs.univie.ac.at:8080/Website/website.html>

selections of facets. When a list of comments is shown based on a combination of facets, the system also shows a short overview of all other possible facets for each comment.

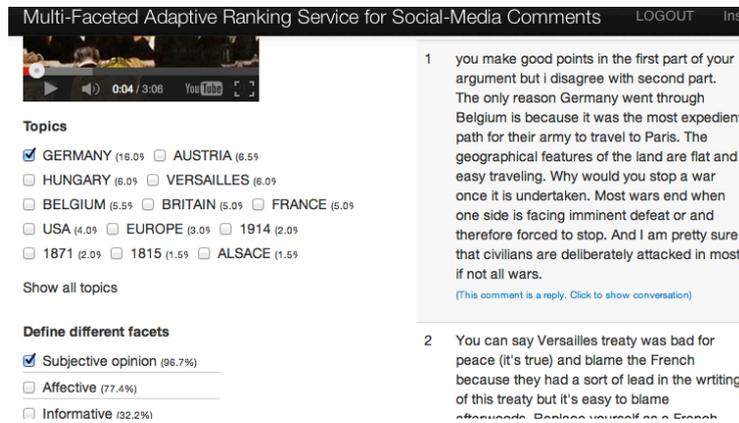


Fig. 1. User interface of the prototype implementation of proposed framework

3 Experimental Set up

In this section we compare the effectiveness of the proposed framework to the prevalent ranking methods: **Reverse-Chronological (RC)** and **Crowd-Based (CB)**. We recruited evaluation participants by distributing calls for participation through internal science mailing lists of two universities. From the respondents, we randomly selected 40 subjects. Participant ages ranged from 20 to 57 with a median of 29. Participants received via an online instruction page and received a gift voucher for their efforts in evaluating the system. After the training phase, they were asked to perform the following steps: (1) Use the prototype to select a title from a list of 30 videos (The primary set of YouTube videos and comments used for these studies is provided by [1]). We restrict these videos to ensure that each video is almost the same length and quality. The participants then choose and watch a video. (2) Utilize the prototype to retrieve a ranked list of the top 30 comments for a video based on reverse-chronological order and also based on crowd-based order (these were the top ranked comments highlighted by YouTube). (3) Use the prototype to retrieve a ranked list of the top 30 comments for the same video in accordance with their preferences by selecting combinations of facets and topics. (4) Vote on each comment and each ranking condition. In the facet-based ranking, each comment is voted along two dimensions: *interestingness* and *relevance*. Using these two scores, we believe, is more interpretable from the end-users perspective (as compared to “usefulness”). In the chronological and crowd-based ordering mode, only the *interestingness* is rated as *relevance*

is a very ambiguous concept without selecting a particular facet. This is because a comment is only considered relevant when it is relevant to the facet selection of a user. We restricted the size of the ranked list of comments to 30 in order to minimize judgment fatigue.

Results The results of our evaluation are shown in Table 1, which reflects performance of **adaptive faceted ranking (AF)**, with regard to three evaluation metrics (MAP–Mean Average Precision, P@10 and P@20). When considering the first default ranking method, the reverse-chronological ranking (RC), the measures indicate that this ranking is at least somewhat effective. Approximately half of the comments retrieved are *interesting* to the users. Furthermore, in consideration of the second ranking method, crowd-based ranking (CB), the effectiveness measures indicate that this ranking type is less effective compared to the reverse-chronological ranking. Approximately one third of the comments retrieved are determined to be *interesting* to the users. In contrast, in the ranking of comments retrieved with our adaptive faceted ranking strategy, approximately every two out of three results are deemed to be interesting.

Table 1. Effectiveness of faceted ranking

Ranking Method	#Ranking	Interesting			Relevant		
		MAP	P@10	P@20	MAP	P@10	P@20
RC	51	0.46	0.48	0.53	<i>Not applicable</i>		
CB	21	0.26	0.32	0.30	<i>Not applicable</i>		
AF	233	0.71	0.67	0.63	0.80	0.70	0.61

4 Conclusion and Future Work

Using the commonly employed ranking methods as two baselines, we have shown that the use of the faceted ranking significantly improves the ranking of comments with respect to relevance and interestingness. In the future, we will investigate the effectiveness of strategies for the selection of different types of facets and will explore the use of personalized ordering of facets and ranking strategies to further improve the interestingness and relevancy to individual users.

References

1. E. Momeni, C. Cardie, and M. Ott, “Properties, prediction, and prevalence of useful user-generated comments for descriptive annotation of social media objects,” in *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM2013)*. Boston, USA: AAAI, June 2013.
2. F. Abel, I. Celik, G.-J. Houben, and P. Siehndel, “Leveraging the semantics of tweets for adaptive faceted search on twitter,” in *Proceedings of the 10th International Conference on The Semantic Web*, ser. ISWC’11. Springer-Verlag, 2011, pp. 1–17.
3. Y. R. Tausczik and J. W. Pennebaker, “The psychological meaning of words: Liwc and computerized text analysis methods,” 2010.