

# Towards Boosting Video Popularity via Tag Selection

Elizeu Santos-Neto  
University of British Columbia  
Vancouver, BC, Canada  
elizeus@ece.ubc.ca

Jussara Almeida  
Univ. Federal de Minas Gerais  
Belo Horizonte, MG, Brazil  
jussara@dcc.ufmg.br

Tatiana Pontes  
Univ. Federal de Minas Gerais  
Belo Horizonte, MG, Brazil  
tpontes@dcc.ufmg.br

Matei Ripeanu  
University of British Columbia  
Vancouver, BC, Canada  
matei@ece.ubc.ca

## Abstract

Video content abounds on the Web. Although viewers may reach items via referrals, a large portion of the audience comes from keyword-based search. Consequently, the textual features of multimedia content (e.g., title, description, tags) will directly impact the view count of a particular item, and ultimately the advertisement-generated revenue.

This study makes progress on the problem of automating tag selection for online videos with the goal of increasing viewership. It brings two major insights: first, it describes a methodology to construct a ground truth to evaluate methods that aim to improve social content popularity; second, it provides evidence that the tags on existing YouTube videos can be improved by an automated tag recommendation process even for a sample of well curated videos; finally, it suggests a roadmap to explore low-cost techniques either based on crowdsourcing or on tag recommendation algorithms to improve the quality of tags for online video content.

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: S. Papadopoulos, P. Cesar, D. A. Shamma, A. Kelliher, R. Jain (eds.): Proceedings of the SoMuS ICMR 2014 Workshop, Glasgow, Scotland, 01-04-2014, published at <http://ceur-ws.org>

## 1 Introduction

Given the sheer volume content owners generate (e.g., YouTube receives 100 hours of video every minute<sup>1</sup>), it is common they offload online publication and monetization tasks to specialized content management companies. Content managers publish, monitor, and promote the owner's content, and revenues are generally shared with the owner. As revenue is directly related to the number of ad prints each piece of content receives, this incentivizes managers to boost content popularity.

Although viewers may reach a content item starting from many *leads* (e.g., an e-mail from a friend or a promotion campaign in an online social network), a large portion of viewers relies on keyword-based search and/or tag-based navigation to find videos. An argument supporting this assertion is the fact that, as of 2/Dec/2013, 14% of the unique visitors on YouTube come from Google.com searches<sup>2</sup>. The integration of Google and YouTube search will likely increase the volume of search traffic that leads to views on YouTube. Moreover, YouTube itself is the third most popular site on the web.

Consequently, the textual features of a video (e.g., title, description, tags, and comments) have a major impact on the view count of each particular item and, ultimately, on the advertisement-generated revenues [5, 11]. Similarly, in other contexts, it has been shown that even simple textual features produce positive results: for example, title suggestions on eBay have benefitted both sellers, who increased revenue, and buyers, who found relevant products faster [5].

---

<sup>1</sup>See: <http://www.youtube.com/yt/press/statistics.html>

<sup>2</sup><http://www.alexa.com/siteinfo/youtube.com#keywords>

Experts can produce the textual features associated with video content via manual inspection (and our industry contacts confirm this is a still current practice<sup>3</sup>). This solution, however, is manpower intensive and limits the scale at which content managers can operate. Therefore, mechanisms to support this process (e.g., automating tag suggestion) are desirable.

This study starts from the observation that, with the ever-increasing volume of user-generated textual content available on the Web, there is a plethora of sources from which an automated mechanism that suggests textual features, in general, and tags, in particular, could extract candidate terms that can improve multimedia content popularity. For example, Wikipedia (a peer-produced encyclopedia), MovieLens and Rotten Tomatoes (social networks where movie enthusiasts collaboratively catalog, rate, and annotate movies), New York Times movie review section (which includes over 28,000 movies) or even YouTube comments are potential sources of candidate keywords to annotate multimedia content such as videos. It is important to note that techniques to suggest textual features to improve multimedia content popularity are not restricted to movies. In fact, other types of content such as superbowl ads could benefit from a combination of information sources such as humans from a crowdsourcing service (e.g., Amazon Mechanical Turk).

To make progress on understanding whether textual information, tags in particular, associated with video content can be improved through an automated process, and on understanding what information sources provide the most valuable textual features (i.e., terms that can potentially improve videos popularity), this work focuses on the following research questions:

Q1 *What are the challenges in building a ground truth to evaluate popularity boosting of videos on social media via textual features optimization such as tagging with terms that can potentially improve the discoverability of the video via search? How can one leverage crowdsourcing channels such as Amazon Mechanical Turk for such purpose?*

Q2 *To what extent the tags currently associated with existing video content on social video-sharing websites, such as YouTube, are optimized to attract search traffic? Is there room for improvement possibly using automated tag recommendation solutions?*

It is worth highlighting that, to tackle these questions, this work uses tag recommender algorithms in a

---

<sup>3</sup>This work is motivated by our collaboration with a company specialized in promoting video content. An NDA prevents the disclosure of details.

different context than most previous studies: our goal is *not* to design novel and more efficient tag recommendation algorithms, but to study whether textual features of social content can be further optimized to improve the value of alternative data sources in providing tags to boost video popularity. In this sense, the mainstream recommender algorithms we use here provide a lower bound on the achievable quality. More complex algorithms, e.g., as proposed in [1, 3, 6, 8], can be tested and tuned using the methodology we propose here to further improve tag quality.

In particular, this study concentrates on the challenges related to constructing a ground truth to enable the evaluation information sources. Therefore, we adopt the following two-part methodology:

- Construct a ground truth by recruiting *turkers* from Amazon Mechanical Turk, asking them to watch YouTube videos, and to provide the keywords they would use to search for each of them, as opposed to simply describe the video (see Section 3);
- Prototype an automated tag recommendation pipeline, incorporate various recommender algorithms, and couple it with different input data sources (see Sections 2 and 4.1);
- Evaluate the tag quality of existing YouTube videos by comparing them with the ground truth (see Section 5);

In summary, the contributions of this work are:

- The production of a ground truth released to the community.
- Evidence that the tags associated with a sample of trailers of popular movies currently available on YouTube can be further optimized by an automated low-cost process. This process can either incorporate human computing engines (e.g., recruited through Amazon Mechanical Turk) at a much lower cost than using dedicated *channel managers* (the current industry practice), or, at an even lower cost, can use tag recommender algorithms to harness textual information from a multitude of data sources that are related to the video content.

## 2 Context of Our Assessment

This section describes the context for our investigation. Our main assumption is that annotating a video with the tags that match the terms users would use to search for it increases the chance that users view

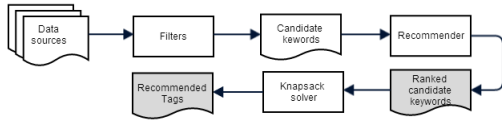


Figure 1: The recommendation pipeline.

the video. This view is supported by previous studies [11] and by the observation that a large portion of the traffic landing on a video comes from textual searches. As a result, textual sources that are related to the video and whose content can be automatically retrieved (e.g., movie reviews, comments, wiki pages, news items, blogs) can be used as inputs for recommenders to suggest tags for these video content items.

A recommendation pipeline that implements this idea is schematically presented in Figure 1: data sources feed the pipeline with textual input data. Next, the textual data is pre-processed by filters to both clean and augment it (e.g., remove stopwords, detect named entities). This first processing step provides candidate keywords for the recommenders. The recommendation step uses the candidate keywords (and their related statistics, such as frequency and co-occurrence) to produce a ranked list according to a scoring function implemented by a given recommender algorithm. Finally, as the space available for tags provided by video sharing websites, such as YouTube or Vimeo, is limited, the selection of most valuable candidate keywords is constrained by a budget, often defined by the number of words or characters. Therefore, the final step consists of solving an instance of the 0-1-knapsack problem [2] that selects a set of recommended tags from the ranked-list produced by the recommender.

In summary, the recommendation pipeline is composed of four main elements: data sources, filters, recommenders, and knapsack solver. The next paragraphs discuss each of these elements.

*Data sources.* This component provides the input textual data used by the tag recommenders. In particular, we are interested in peer-produced data sources such as Wikipedia and social tagging systems like MovieLens, as well as expert-produced data sources such as NYTimes movie reviews. We discuss in detail each of the data sources used in Section 4.1.

*Filters.* The raw textual data extracted from a data source is filtered to both clean and augment the input data, minimizing noise. We consider simple filters: stopword and punctuation removal, lowercasing, and named entity detection<sup>4</sup>.

*Recommenders.* A recommender scores the candidate keywords based on their relevant statistics (e.g.,

single word frequency, word co-occurrence frequency). Note that there are many ways of defining scoring functions; and, it is not our goal to advocate a specific one, as we focus on the value provided by various information sources. We discuss the recommenders used in this work in Section 4.2.

*Knapsack Solver.* Finally, after ranking candidate keywords, the final step is selecting the ones that best fit the budget. In this paper, the budget is expressed in terms of the number of characters, as done in video sharing systems such as YouTube, where one can use a limited number of characters (500) for tags. This step is formulated as a 0-1-knapsack problem.

Let  $v$  be a video and  $C = \langle k_i \rangle$ ,  $i = 1, \dots, n$ , be a list of candidate keywords provided by a data source when used as input to a tag recommendation algorithm. Additionally, let us denote the length of a keyword  $k_i$  as  $w_i$  in bytes. Therefore, the problem of selecting the best tags to improve viewership of  $v$  is equivalent to solving the following optimization [2]:

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n f(k_i, v) x_i \\ & \text{subject to} && \sum_{i=1}^n w_i x_i \leq B \end{aligned}$$

where  $B$  is the budget in terms of number of characters allowed in the tags field,  $x_i \in \{0, 1\}$  is an indicator variable, and  $f(k_i, v)$  is a scoring function provided by the recommender for the keyword  $k_i$  with respect to video  $v$ .

Considering that the cost<sup>5</sup> (i.e., the keyword length) and the scores are both nonnegative, we use a well-known dynamic programming algorithm [2] to solve this optimization problem.

**Our goal** is primarily to understand whether videos currently published on a popular social video sharing website have their tags optimized to attract search traffic. If tags can still be further optimized, one could evaluate how the choice of the data source used as input for a recommendation pipeline impacts the quality of the recommended tag-set. Next, we discuss how to build a ground truth that enables testing whether the tags assigned to a sample of videos available on a popular video sharing website are optimized. Additionally, such ground truth, in a future work, could enable the evaluation of potential data sources that provide candidate tags.

### 3 Building the Ground Truth

Our main assumption is that annotating a video with the tags that match the terms users would use to

<sup>4</sup>We use OpenCalais.com for entity detection.

<sup>5</sup>Our study can be easily extended to consider the budget as the number of tags (as in Vimeo).

search for it increases the chance that users watch the video. As a result, textual sources that are related to the video and whose content can be automatically retrieved (e.g., movie reviews, comments, wiki-pages, news items, blogs) can be used as inputs for recommenders to suggest tags for these video content items.

The ideal method to collect the ground truth would consist of experiments that vary the set of tags associated to videos, and capture their impact on the number of views attracted. However, collecting this requires the publishing rights for the videos and implies executing experiments over a considerable duration.

We decided for an alternative solution: we built a ground truth by setting up a survey using the Amazon Mechanical Turk (AMT)<sup>6</sup>. The survey asks participants to watch a video and answer the question: What query terms would you use to search for this video? The rationale is that these terms would, if used as tags to annotate the video, maximize its retrieval by the YouTube search engine, and indirectly maximize viewership.

*Content Selection.* Our study focuses on movie trailers<sup>7</sup>. The reason is that they are often short (about 5 minutes or less), and this makes the evaluation process more dynamic, encouraging turkers (i.e., the AMT workers who accept to participate in the survey) to watch more trailers and associate more keywords to them. The dataset consists of 382 videos selected to meet the following constraints: they must be publicly available on YouTube and have available content in the data sources used to extract candidate keywords (e.g., a page on Wikipedia, a NYTimes review page).

*Survey.* First, we conducted a pilot survey by recruiting participants via our internal mailing lists and online social networks. Their task was to watch the trailers and provide terms they would use to search for those videos. This pilot highlighted two major issues: (i) relying only on volunteerism to mobilize participants was insufficient; and, (ii) quality control of answers (e.g., typos in the keywords) is much more difficult - all videos in the survey are in English and there was no automatic way to recruit only participants that are fluent in English.

Thus, we published an AMT task<sup>8</sup> requiring the turkers to watch trailers, and provide the query terms (3 to 10 keywords to each video, as queries are typically of that length [4]) they would use to search for the videos they had watched. Following AMT pay guidelines, each participant was paid \$0.30 per task assignment, which had an average completion time of

<sup>6</sup><http://www.mturk.com>

<sup>7</sup>As long as there are data sources to extract candidate tags from, other content types can benefit from our methodology.

<sup>8</sup>Links to data and code can be provided upon request.

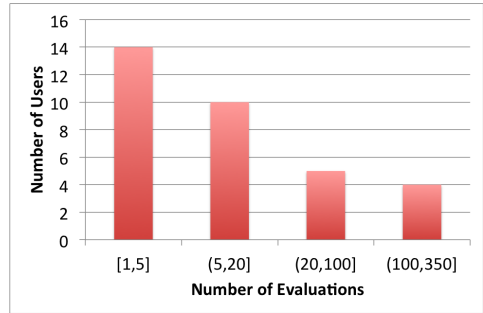


Figure 2: Histogram with number of evaluations performed by turkers

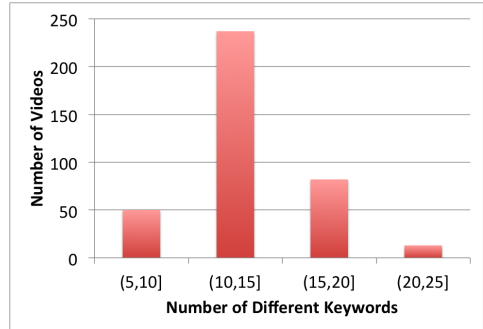


Figure 3: Number of distinct keywords assigned by turkers to each video.

6 minutes (total cost to conduct the survey: \$345). This leads to \$3 per hour, which is much cheaper than the wage paid to dedicated *channel managers*.

We also performed simple quality control by inspecting each answer to avoid accepting spam (which is expected to be rare, due to the reputation mechanism adopted by AMT). In fact, only one submission was rejected as unrelated URLs were assigned as answers instead of keywords.

*A brief characterization of the ground truth.* In total, 33 turkers submitted solutions. Figure 2 shows the number of videos evaluated per turker: as we can observe, 19 turkers (58%) evaluated more than 5 videos, with the maximum reaching 333 videos. Figure 3 shows a histogram presenting the number of different keywords each video received. Even though we asked the turkers to associate at least 3 keywords to each video, 82% of the evaluations provided more than the required minimum, which resulted in 96% of the videos with 10 or more different keywords.

Figure 4 presents the total number of characters in the set of unique keywords associated to each video. The length of the ground truth varies from 51 (min) to 264 (max) characters; in fact, 32% of the videos have tags summing up to 100 characters. These values guided the budget parameter in our experiments, as we explain in Section 4.3.

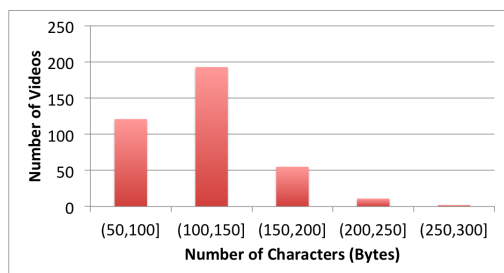


Figure 4: Histogram with the number of characters in bytes to each video.

To gain an understanding on what types of keywords would drive search traffic to these videos, we look at the set of most popular terms (overall) in the ground truth. Among the top 10% most frequent search terms provided by turkers, 68% of them are named entities (e.g., actor, director, and producer names). Another category of terms with strong presence is genre-describing terms. This suggests that a strategy that aim to boost popularity of videos by optimizing the tags associated with the content should use sources that provide named entities related to the video.

It is important to note that we found some evidence that this happens to videos other than movie trailer. In a smaller sample of Super Bowl ads videos, we observed that terms users provide as keywords they would use to search for the ads are also dominated by named entities like the brand names.

## 4 Experimental Setup

This section presents the instances of data sources and tag recommenders, as well as the success metrics used in our evaluation on whether the existing tags on YouTube videos are optimized.

### 4.1 Data Sources

To understand whether the tags assigned to existing online video content are optimized to attract search traffic, it is necessary to compare the current tags to a set of tags recommended by using other data sources as input.

Therefore, we use a combination of data sources as inputs to recommender algorithms to produce a comparison basis for the existing tags on the videos in our sample. Next, we describe each of these data sources:

*MovieLens*<sup>9</sup> is a web system where users collaboratively build and maintain a catalog of movies and their ratings. Users can create and update movie entries, annotate movies with tags, review and rate them. Based on previous user activities, MovieLens suggests

<sup>9</sup><http://www.movielens.org>

movies a user may like to watch. For our evaluation, we use some of the data available in MovieLens: only the tags users produce while collaboratively annotating and bookmarking movies. This data is a publicly available trace of tag assignments<sup>10</sup>.

*Wikipedia* is a peer-produced encyclopedia where users collaboratively write articles about a multitude of topics. Users in Wikipedia also edit and maintain pages for specific movies<sup>11</sup>. We leverage these pages as the sources of candidate keywords for recommending tags for their respective movie trailers from our sample.

*NYTimes* reviews are written by movie critics considered experts on the subject. Similar to the data provided by Wikipedia, we leverage the review page of a movie as the source of candidate keywords for the tag recommendation task. The reviews are collected via the query interfaces<sup>12</sup> provided by the NYTimes API.

*Rotten Tomatoes* is a portal where users can rate and review movies. Moreover, users have access to critics' reviews and all credits information: actors and roles, directors, soundtrack, synopsis, etc. The portal links to critics' reviews as well. The information about the credits of a movie and the critics' reviews can be considered as produced by experts (likely the film credits are obtained from the movie producers, while the critics' reviews are similar to those from NYTimes). While users can review the movies as well (and this qualifies as peer-produced information), these reviews are available on the website, but not accessible via the API at the time of our investigation. The rest of the information about the movies together with links to the experts' reviews is available via the Rotten Tomatoes API<sup>13</sup>.

*YouTube* is a video-sharing website, here used to test whether the tags already assigned to videos can be further optimized. To this end, we collect the tags assigned to the YouTube videos in our sample from the HTML source of each video's page (API requests in this sense are only accessible by the video publisher). The reason for using page scraping rather than API requests is that videos' tags are accessible via the API only to the video publisher, even though these tags are still used by the search engine to match queries and are available in the HTML of the video page. YouTube data source figures in the expert-produced end of the spectrum since only the publisher can assign tags to the video (it is reasonable to assume that a publisher is an expert on the own video and aims to optimize its textual features to attract more views).

<sup>10</sup><http://www.grouplens.org/taxonomy/term/14>

<sup>11</sup>[http://en.wikipedia.org/wiki/Pulp\\_Fiction\\_\(Film\)](http://en.wikipedia.org/wiki/Pulp_Fiction_(Film))

<sup>12</sup><http://developers.nytimes.com>

<sup>13</sup><http://developer.rottentomatoes.com/>

## 4.2 Tag Recommenders

The experiments use two tag recommendation algorithms that process the input provided by the data sources: FREQUENCY and RANDOMWALK. We selected them primarily because they harness some fundamental aspects of the tag recommendation problem that more sophisticated methods (e.g., [1, 6, 9]) also use: tag frequency, and tag co-occurrence patterns. Moreover, our goal is to understand the relative influence of the data sources on the quality of the recommended tags. We note that the methodology we describe and the ground truth can be used to evaluate other, more sophisticated, recommender algorithms as well.

The FREQUENCY recommender scores the candidate keywords based on how often each keyword appears in the data source. Given the movie title, our pipeline finds the documents in the data source that match the title and extract a list of candidate keywords. For example, in Wikipedia, the candidate keywords for recommendation to a given movie are extracted from the Wikipedia page about the movie, and its frequency are the number of times each one appears in that page. Similarly, in MovieLens, the frequency is the number of times a tag is assigned to a movie.

The RANDOMWALK recommender harnesses both the frequency and the co-occurrence between keywords. The co-occurrence is detected differently depending on the data source. In MovieLens, two keywords co-occur if they are assigned to the movie by the same user, while in NYTimes, Rotten Tomatoes, and Wikipedia two keywords co-occur if they appear in the same page related to the movie (i.e., review, movie record, and movie page, respectively). The RANDOMWALK recommender builds a graph based on keyword co-occurrence, where each keyword is a node.

## 4.3 Budget Adjustment

To make the comparison fair, for each movie trailer, we adjust the budget to the size of the tag set of that video in the ground truth. The knapsack solver uses this budget to select the recommended tags for a particular video. The reason for setting a budget per video is that a number of recommended tags greater than the ground truth size penalizes some evaluation metrics, such as the F3-measure (see definition below).

## 4.4 Success Metrics

The final step in the experiment is to estimate, for each video and for various input data sources and recommender algorithms, the quality of the recommended tag-set against the ground truth. To this end, we use F3-measure. Let  $T_v$  and  $S_v$  be the set of distinct keywords in the ground truth and in the recommended

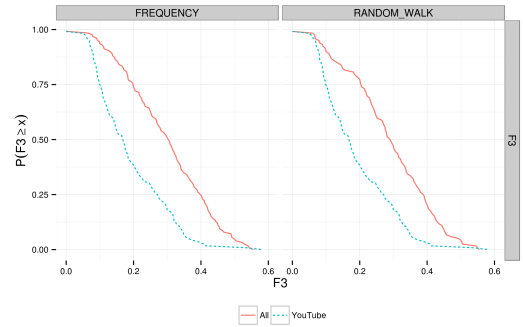


Figure 5: CCDF of F3-measure for YouTube tags and recommended tags.

tag set, respectively, for video  $v$ . The metric is defined as follows:

**F3-metric.**  $F_3(v) = \frac{10 \cdot P(v)R(v)}{9 \cdot P(v) + R(v)}$ , where  $P(v) = \frac{|T_v \cap S_v|}{S_v}$  and  $R(v) = \frac{T_v \cap S_v}{T_v}$  are the precision and recall of tag recommendation for video  $v$ , respectively.

## 5 Experimental Results

This section presents our experimental results to address the following research questions:

*To what extent the tags currently associated with existing YouTube content are optimized to attract search traffic? Is there room for improvement using automated tag recommendation solutions?*

To address these questions, we perform an experiment to assess the quality of tags already assigned to existing YouTube videos and whether there is room for improvement. By improvement we mean extending/modifying the tags to better match the ground truth. To this end, we compare the tags to the ground truth for each video and observe a wide gap. The dotted (blue) line on the left in Figure 5 presents the Complementary Cumulative Distribution Function (CCDF) for the F3-measure. A point in the curve indicates the percentage of videos (on y-axis) for which the F3-measure is larger than the corresponding value on x-axis, thus, the closer the line is to the top-right corner, the better.

To explore whether the gap than can be covered, at least partially, by automated tag recommendations, we explore the performance of the tag recommendations using as inputs all data sources combined (MovieLens, Rotten Tomatoes, Wikipedia, and NY Times). The results are presented in Figure 5 as the solid (red) line.

The *Kolmogorov-Smirnov test* of significance indicates that the performance of using All data sources is significantly higher than that achieved by the YouTube tags (FREQUENCY:  $D^- = 0.44$ ,  $p$ -value =  $3.9 \times 10^{16}$ ; RANDOMWALK:  $D^- = 0.43$ ,  $p$ -value =  $5.5 \times 10^{-15}$ ) implying that the tags recommended by both meth-

ods are better than those currently assigned to the videos on YouTube. *Therefore, the tags currently assigned to the YouTube videos can still be improved by automated methods to attract more search traffic, and, hence, boost video popularity.*

## 6 Related Work

The quest to improve visibility of one’s content (e.g., a website, a video) is not new – the whole Search Engine Optimization segment has seen uninterrupted attention. Multiple avenues are available, ranging from some that are viewed as abusive (e.g., link-farms) to perfectly legitimate ones (e.g., better content organization, good summaries in the titlebar of web pages). Our exploration falls into this latter category.

The related literature falls into two broad categories: automated content annotation and tag value assessment. The majority of related work on automated content annotation (or tag recommendation) focuses on suggesting tags to annotate content items such that they maximize the relevance of the tag given the content [1, 5, 7, 9], with a few exceptions where authors propose to leverage other aspects such as diversity [1].

Although finding relevant tags to a given content item is an important component of improving the tags assigned to this item, previous studies fail to account for the potential improvement on the view count of the annotated content – an aspect which is valuable to content managers and publishers, as they monetize based on the audience that is able to find their content.

The study presented by Zhou et al. [10] is, to the best of our knowledge, the closest to our work. However, contrary to our study that focus on testing whether tags can be further optimized to attract traffic, Zhou et al. propose to boost video popularity by suggesting ways to connect a video to other influential videos (e.g., making title and description similar to those of influential videos) as a way to leverage the related video recommendations.

Our study is different from these previous efforts, as it focuses on testing the hypothesis that textual features of social content, such as online videos, can be further optimized to potentially attract search traffic. This motivates our future work on evaluating the impact of data source choice to produce recommendations.

## 7 Summary and Future Work

A large portion of traffic received by video content on the web originates from keyword-based search and/or tag-based navigation. Consequently, the textual features of this content will directly impact the popularity

of a particular content item, and ultimately the advertisement generated revenue. Therefore, understanding the performance of automatic tag recommenders is important to optimize the view count of content items.

First, we discuss the challenges on building a ground truth to evaluate data sources and techniques that aim to boost the popularity of multimedia content on the web. Next, this study provides evidence that tags currently assigned to a sample of YouTube videos can be further improved to attract more search traffic. To this end, we show an experiment that compares how close the tags currently assigned to the videos in the sample and tags harnessed from a combination of data sources are to the ground truth. The results show that using simple recommenders and a combination of data sources can improve the tags.

These preliminary results suggest a few directions of future research. Initially, one may perform comparisons between data sources individually and/or grouped by type (peer- and expert-produced, structured vs. unstructured) with the goal of understanding their relative value as inputs for tag recommenders. For example, are combinations of peer-produced data sources relatively more valuable than expert-based ones in the context of boosting multimedia content popularity?

Additionally, more experiments could provide deeper explanations on the performance of peer-produced data sources. For instance, does the value of tags extracted from peer-produced sources (for boosting content popularity), such as Wikipedia or MovieLens, increase with the number of contributors? All these questions are part of our future efforts.

## References

- [1] F. Belém, E. Martins, J. Almeida, and M. Gonçalves. Exploiting Novelty and Diversity in Tag Recommendation. In P. Serdyukov, P. Braslavski, S. Kuznetsov, J. Kamps, S. Rüger, E. Agichtein, I. Segalovich, and E. Yilmaz, editors, *Advances in Information Retrieval SE - 32*, volume 7814 of *Lecture Notes in Computer Science*, pages 380–391. Springer Berlin Heidelberg, 2013.
- [2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. The MIT Press, third edit edition, July 2009.
- [3] Z. Guan, J. Bu, Q. Mei, C. Chen, and C. Wang. Personalized tag recommendation using graph-based ranking on multi-type interrelated objects. SIGIR ’09, pages 540–547, Boston, MA, USA, 2009. ACM.

- [4] B. He and I. Ounis. Query performance prediction. *Information Systems*, 31(7):585–594, Nov. 2006.
- [5] S. Huang, X. Wu, and A. Bolivar. The effect of title term suggestion on e-commerce sites. WIDM '08, pages 31–38, Napa Valley, California, USA, 2008. ACM.
- [6] R. Krestel and P. Fankhauser. Language Models and Topic Models for Personalizing Tag Recommendation. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 82–89, Toronto, AB, Canada, Aug. 2010. IEEE.
- [7] D. Liu, X. S. Hua, L. Yang, M. Wang, and H. J. Zhang. Tag ranking. WWW '09, pages 351–360, Madrid, Spain, 2009. ACM.
- [8] S. Rendle and L. Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the third ACM international conference on Web search and data mining - WSDM '10*, page 81, New York, New York, USA, 2010. ACM Press.
- [9] C. Wang, F. Jing, L. Zhang, and H. J. Zhang. Image annotation refinement using random walk with restarts. MULTIMEDIA '06, pages 647–650, Santa Barbara, CA, USA, 2006. ACM.
- [10] D. Zhou, J. Bian, S. Zheng, H. Zha, and C. L. Giles. Exploring social annotations for information retrieval. In *17th International World Wide Web Conference*, pages 715–724, Beijing, China, 2008. ACM.
- [11] R. Zhou, S. Khemmarat, L. Gao, and H. Wang. Boosting video popularity through recommendation systems. In *Databases and Social Networks on - DBSocial '11*, pages 13–18, New York, New York, USA, June 2011. ACM Press.