

Relevance, Novelty, Diversity and Personalization in Tag Recommendation

Fabiano Muniz Belém
fmuniz@dcc.ufmg.br
Dep. de Ciência da Computação
Universidade Federal de Minas Gerais

Jussara Marques Almeida
jussara@dcc.ufmg.br
Dep. de Ciência da Computação
Universidade Federal de Minas Gerais

Marcos André Gonçalves
mgoncalv@dcc.ufmg.br
Dep. de Ciência da Computação
Universidade Federal de Minas Gerais

ABSTRACT

The design and evaluation of tag recommendation methods have historically focused on maximizing the relevance of the suggested tags for a given object, such as a movie or a song. Tag relevance can be defined in two perspectives. In an *object-centered* perspective, a tag is relevant if it correctly describes the content of the target object, while in a *personalized* perspective, a relevant tag not only describes well the content of the target object, but also matches the interests of the target user. However, even enriched by a personalized perspective, relevance by itself may not be enough to guarantee recommendation usefulness. Promoting novelty and diversity in tag recommendation not only increases the chances that the user will select some of the recommended tags, but also promotes complementary information (i.e., tags), which helps cover multiple aspects or topics related to the target object. Yet, no prior work has tackled novelty and diversity in the specific context of tag recommendation. In this thesis, we aim at proposing novel solutions that effectively address multiple aspects related to the tag recommendation problem, notably, relevance, novelty, diversity and personalization of the suggested tags. We evaluate our strategies using real data from five Web 2.0 applications, namely, Bibsonomy, LastFM, MovieLens, YahooVideo and YouTube. Our experimental results demonstrate the effectiveness of our new methods over state-of-the-art approaches, and attest the viability to effectively increase novelty and diversity with only a slight impact (if any) on relevance. We also found that our proposed syntactic attributes are responsible for significant improvements (up to 17% in precision) over the best relevance-driven method in a cold start scenario. In addition, we assessed the benefits of personalization to provide better descriptions of the target object, with average gains of 15% in relevance over the best object-centered approach.

KEYWORDS

Tag Recommendation, Web 2.0, Relevance, Novelty, Diversity, Personalization

1 INTRODUCTION

This article presents a summary of the results obtained with the thesis *Relevance, Novelty, Diversity and Personalization in Tag Recommendation*, presented to the PPGCC (Programa de Pós-graduação em Ciência da Computação) at March, 6th, 2018. The author of the

thesis is the first author of this article, while his advisers are the two last authors of this article.

Web 2.0 applications offer a platform for end users to create and share large amounts of rich data content in the form of Web 2.0 objects. These objects usually comprise a main *media content* (e.g., a video, image, audio or text) along with other sources of data associated with it (e.g., links, meta-data), referred to as its *features*. The *textual features* of an object are well-defined blocks of text such as title, tags, description and user comments, used to describe the object's content, facilitating content organization and promoting findability [3]. Among all textual features, tags have gained special importance as they offer an effective data source for information retrieval (IR) services such as search [7], classification [5] and item recommendation [6], due to their capability of capturing personal or (sub-)community interests reasonably well [7]. In this context, there is a large interest in developing strategies to automatically recommend tags, providing *relevant* and *useful* tag suggestions for a target object, and indirectly improving the quality of the IR services that rely on them as a data source. Our thesis aims at proposing novel solutions that effectively address multiple aspects of tag recommendation, notably, relevance, novelty, diversity and personalization. We refer to our thesis for a formal definition of the problem.

1.1 Motivation

Tag recommenders have historically focused on maximizing the *relevance* of the recommended tags [3]. Tag relevance can be defined in two perspectives. In an *object-centered* perspective, a tag is relevant if it correctly describes the content of the target object. In a *personalized* perspective, a relevant tag not only describes well the content of the target object, but also matches the interests of the target user [10]. Personalization is motivated by the fact that users typically have different interests, levels of expertise or vocabulary biases, and may also have different purposes when choosing tags for a target object (e.g., content organization or content description). Thus, personalized tag recommendation aims at suggesting tags that not only are related to the object's content but also capture the user interests, profile and background. Furthermore, personalized tag recommendations may also provide, either in isolation or collectively (i.e., all personalized recommendations provided to all users who tagged an object) better and more complete descriptions of the object's content, compared to object-centered recommendations (which is also one of our findings in the thesis).

However, even enriched by personalization, relevance may not be enough, in isolation, to guarantee recommendation usefulness [2]. For example, a list of synonyms that well describe the object's content is arguably relevant, but also redundant and less useful than

In: I Concurso de Teses e Dissertações (CTD 2019), Rio de Janeiro, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2019.
ISSN 2596-1683

a more *diversified* list covering several aspects related to the object. Indeed, the utility of a recommended item (or tag, specifically) depends on the other items in the list of recommendations [2], due to the possible redundancy among them. Recommending tags that bring *novel* and *diverse* information with regards to previously ranked tags may promote complementary information, improving the coverage of the multiple aspects or topics related to the target object and, indirectly, improving results of tag-based information retrieval (IR) services.

Diversity is particularly important because multimedia objects on the Web 2.0 may be *multifaceted*, that is, they may be related to various aspects and topics. In fact, we observed that a large fraction of Web 2.0 objects present multiple categories. For example, 84% of the artists in our LastFM dataset and 63% of the movies in our MovieLens dataset are associated with two or more categories (style or genre). *Novelty*, on the other hand, can increase serendipity, coverage and recall of services that use more “specific” (yet relevant) recommended tags.

Thus, in our thesis, we define novelty as the capacity of recommending *long tail* [2] tags, that is, more rare tags. The idea is that a term used as tag many times tends to be a more “obvious” recommendation (if relevant at all), thus being of little use (if any) to improve the description of the target object provided by its tag set. We note that this concept is closely related to tag *specificity*, since rare words tend to be more specific (less general) [4]. Diversity, in turn, refers to the *exhaustivity* [4] of the set of recommended tags, which is defined as the coverage they provide for the topics of the target object.

To further illustrate the benefits of novelty and diversity in tag recommendation, Table 1 shows an example of recommendations produced for a MovieLens object (i.e., movie) by three recommenders: one focused on relevance only, a second one that directly incorporates diversity and a third one that, besides diversity, also considers novelty aspects¹. The relevance-driven recommender suggested the relevant tags “comics”, “superhero” and (though more vague) “based”, possibly referring to the fact that the movie is based on the Marvel’s comics X-Men. But it also suggested the general tag “dvd”. Notice also that, despite being driven by relevance, the recommender suggested an apparently irrelevant tag (as far as we can tell), “ummarti2006”. The second recommender, in turn, which incorporates diversity aspects, also suggested the tags “dvd” and “comics”, but together with “genetics” and “biology”, which may be seen as other important subjects of the movie plot (a fiction related to genetic evolution). Those two tags cover other topics related to the movie, increasing the diversity of the recommendations. The tag “mckellen”, also suggested by the second recommender, can also be considered relevant, as it refers to one of the main actors of the film, Ian McKellen. We also note that all recommended tags are, to some extent, relevant to the movie, which illustrates a “good side effect” of promoting diversity: ensuring that at least one relevant tag for each topic related to an object will be suggested may demote too general or noisy tags, improving the relevance of the recommendations. In fact, our experimental results corroborate this hypothesis. Finally, the third recommender, which fully exploits all

three aspects, brought one more novel and specific tag, “marvel”, which represents well the creators of the movie’s universe, not to mention that it is related to the “comics” topic in a more specific way. While this example illustrates that *diversity* and *novelty* are important aspects for tag recommendation, to our knowledge, no previous work has addressed aspects other than relevance and personalization in our specific context.

Another issue that has been mostly neglected in tag recommendation is the cold start problem, which refers to an insufficient amount of previous information about items or users (e.g., when new items or users are introduced in the system), making it difficult to provide effective recommendations [12]. In the specific tag recommendation domain, cold start has been defined as the absence of an initial set of tags associated with the target object [9].

Many state-of-the-art tag recommendation methods exploit co-occurrence patterns with the initial tag set, recommending to a target object o , associated with an initial set of tags I_o , tags that frequently co-occur with tags in I_o in a training collection [3]. Yet, as shown by [9], the effectiveness of these methods greatly suffers in a cold start scenario in which those initial tags are absent, due to the absence of such co-occurrence information.

In order to address this issue, previous work has exploited other textual features (e.g., title, description), extracting candidate tags directly from the text associated with the target object [8, 11], or from neighbors (similar/related objects) [9]. However, these previous efforts focus only on statistical properties of the occurrence of words, such as term frequency (TF) and inverse document frequency (IDF) [1]. These properties in isolation may fail to identify the most relevant candidate tags, specially from the typically small and possibly low quality texts associated with Web 2.0 objects [5]. Thus, it is necessary to propose alternative tag quality attributes to distinguish relevant from non-relevant candidate tags, as well as alternative sources to generate candidate tags, which is one of the topics we tackle in our thesis, as we will discuss in Section 2.

1.2 Objectives

In sum, our thesis hypothesis is that we can simultaneously address various aspects of the quality of recommended tags, some of which are conflicting objectives (e.g., focusing on novelty can be detrimental to relevance), by proposing and combining different tag quality attributes to address scenarios with and without cold start. More specifically, by automatically combining various tag quality attributes (some of them proposed in our thesis), using learning-to-rank techniques, we can improve the relevance of the recommended tags. Novelty and diversity aspects can also be captured by tag quality attributes, and further improved by re-ranking strategies. Finally, by designing suitable attributes to deal with the cold start and personalization issues, it is also possible to improve the effectiveness of the recommendations.

1.3 Publications and Awards

The results of our thesis were published in various journals and conferences of high impact, including:

- (1) Exploiting syntactic and neighbourhood attributes to address cold start in tag recommendation. Fabiano M. Belém,

¹These are real recommendations produced by some of our proposed methods and baselines.

Table 1: Example of tag recommendations for a MovieLens object.

<i>Title</i>	<i>Relevance Only Recommender</i>	<i>Relevance + Diversity Recommender</i>	<i>Relevance + Diversity + Novelty Recommender</i>
X-Men: The Last Stand	dvd, comics, ummarti2006, super-hero, based	dvd, genetics, biology, comics, mckellen	genetics, dvd, biology, mckellen, marvel

André G. Heringer Jussara M. Almeida Marcos A. Gonçalves. Information Processing Management 56(3), 2019

- (2) F. Belém, C. Santos, R. Santos, J. Almeida, M. Goncalves. A Survey on Tag Recommendation Techniques. Journal of the Association for Information Science and Technology (JASIST), 2017
- (3) F. Belém, C. Santos, R. Santos, J. Almeida, M. Gonçalves. Beyond Relevance: Explicitly Promoting Novelty and Diversity in Tag Recommendation. ACM Transactions on Intelligent Systems and Technology. 7(3), 2015
- (4) F. Belém, E. Martins, J. Almeida, M. Gonçalves. Personalized and object-centered tag recommendation methods for Web 2.0 applications. Information Processing Management, 50(4), 2014
- (5) S. Canuto, F. Belém, J. Almeida, M. Gonçalves. A comparative study of learning-to-rank techniques for tag recommendation. Journal of Information and Data Management (JIDM), 4:453–468, 2014
- (6) F. Belém, R. Santos, J. Almeida, M. Gonçalves. Topic Diversity in Tag Recommendation. In Proceedings of the 7th ACM Conference on Recommender Systems (RecSys), 2013
- (7) F. Belém, E. Martins, J. Almeida, M. Gonçalves. Exploiting Relevance, Novelty and Diversity in Tag Recommendation. In European Conference on Information Retrieval (ECIR), 2013
- (8) F. Belém, E. Martins, T. Pontes, J. Almeida, M. Gonçalves. Associative tag recommendation exploiting multiple textual features. In ACM SIGIR Conference on Information Retrieval, 2011

It is also worth mentioning that our work has been recognized and has been awarded by:

- 2015 **IBM Ph.D. Fellowship Award**
- 2016-2017 **Google Research Awards for Latin America** (two consecutive years)
- **Prêmios CAPES e UFMG de Teses 2019**: our work has been selected by PPGCC (the DCC/UFMG post graduation program) among the best thesis to compete in the national award.

2 CONTRIBUTIONS AND RESULTS

2.1 Proposal and Combination of Tag Quality Attributes

We combined a number of tag quality attributes by means of heuristics and L2R-based techniques. Some of these attributes and techniques have already been proposed and evaluated in our previous work [3]. Other attributes, namely, the topic-related attributes, and syntactic attributes, briefly discussed in Sections 2.2 and 2.4, respectively, are novel contributions of our thesis.

The best analyzed L2R-based strategy outperforms the state-of-the-art heuristic (LATRE+wTS, or Lazy Associative Tag Recommender + weighted Term Spread [3]), producing gains of up to 29% in average NDCG. Among the L2R based strategies, there is a clear winner group of methods: Random Forests (RF), MART and λ -MART, which produces gains ranging from 4% to 12% in average NDCG over the best of the remaining L2R-based methods (i.e., the previously proposed methods GP, RankSVM and Rankboost). Furthermore, we found that the L2R approach presents a very low additional recommendation time (under 3%) when compared with LATRE+wTS. Besides the promising results, the flexibility of the L2R framework in terms of the incorporation of new attributes and ability to maximize different target measures (as we did in the thesis, when adding personalization, novelty and diversity aspects, as well as addressing cold start) makes it a very suitable solution.

It is also worth mentioning that, besides combining tag quality attributes, we also tested combinations of L2R methods by means of a stacking technique, and some straightforward strategies such as summing up the scores given by each method. However, for the same set of attributes, these strategies did not outperform the best L2R based method (RF) in isolation. In general, we note that the highest gains are achieved when we combine different (and complementary) sources of candidate tags and tag quality attributes.

2.2 Addressing Cold Start with Syntactic Attributes and Neighborhood Expansion

In this front of work, we proposed syntactic related attributes and nearest neighbor techniques to extend and improve tag recommendation methods in a cold start scenario. We note that these techniques provide much higher gains in this particular scenario than in a scenario in which there are some tags available in the target object, since tag co-occurrences with these initial tags are strong evidence of the quality of a candidate tag, reducing the need for additional attributes.

First, we investigated syntactic patterns of the text associated with Web 2.0 objects that can be exploited to identify and recommend tags. For example, the probability of a word being relevant as a tag given its part-of-speech (PoS) label (e.g., noun, verb) or syntactic function (e.g., subject, adverbial adjunct). We also proposed new tag quality attributes based on these syntactic patterns, exploiting them to further improve our proposed L2R-based tag recommenders in the given scenario. Our experiments showed that our proposed syntactic attributes are responsible for significant improvements (up to 17% in precision over the best relevance-driven method). A feature importance analysis confirmed that our new attributes are among the most discriminative for the problem in hand, in particular the sequence of syntactic dependencies between the candidate tag and the root of the sentence, the token connected to the candidate tag in the syntactic tree, and the root of the sentence.

Moreover, we also analyzed to which extent we can further improve tag recommendations by exploiting the neighborhood of the target object (i.e., similar objects). We used the L2R-based tag recommender with syntactic attributes to compute a new, complementary neighborhood. Recommendations based on this new neighborhood outperformed those generated from traditional nearest neighbors approaches, which exploit only TFIDF as weights for terms in an object. Finally, $KNN_{synt} + RF_{synt}$, our combination of both neighborhood and L2R-based tag recommenders, consistently produced the best results, with gains of up to 21% over RF .

2.3 Personalization of Tag Recommendation

We proposed four heuristics and evaluated three new L2R-based methods (RF , $RankSVM$ and GP) to address the personalized tag recommendation problem. Furthermore, we have provided a quantitative assessment of the benefits of personalized tag recommendation to provide better descriptions of the target object. We found that our heuristics produced gains of up to 157% in average NDCG over a state-of-the-art personalized tag recommendation method ($PITF$). Our best L2R method, RF , provided average relevance gains of 9% over our best heuristic and gains of 5% over our previous L2R-based strategies (i.e., $RankSVM$ and GP). Comparing our best personalized and object-centered tag recommendation methods, both based on the RF technique, we found that the former outperforms the latter, with average gains of 15% in relevance. Thus, we found that personalization brings benefits when applied to provide better descriptions of the target object.

2.4 Improving Novelty and Diversity of Tag Recommendation

We have proposed four new tag recommendation methods aiming at exploiting novelty and diversity, in different levels. Our first method, called GP_{rnd} , extends the relevance-driven method GP , which already incorporates some novelty aspects at the attribute level, to include novelty and diversity metrics at both attribute and objective function levels. The second method, called RF_t , extends the relevance-driven approach based on RF to include new tag attributes that capture the extent to which a candidate tag is related to the topics (e.g., categories) of the target object. This solution indirectly captures topic diversity while trying to maximize relevance in its objective function. Unlike RF_t , our third method, Explicit Tag Recommendation Diversifier ($xTReD$), directly exploits topic diversity, by *re-ranking* the recommendations provided by any tag recommender. Finally, our fourth proposal, called Explicit Tag Recommendation Diversifier with Novelty Promotion ($xTReND$), generalizes $xTReD$, to fully exploit relevance, novelty and topic diversity.

Our evaluation showed that GP_{rnd} provides gains in novelty without significantly harming relevance when compared to GP , but the gains in diversity (both implicit and explicit) are modest (at best). We also found that the use of our new topic related metrics at the attribute level (as performed by RF_t) does contribute to produce better tag recommendations, particularly if predefined categories are used as topics, allowing substantial gains in diversity (up to 35%) as well as some (modest) gains in relevance (up to 5% in average NDCG), when compared to RF . Overall, our new method,

$xTReND$, is the best out of the four new methods, considering the trade-offs among relevance, novelty and diversity. Though more modest, the improvements of our new methods over the baselines are still significant if LDA (Latent Dirichlet Allocation) topics are used, implying that such unsupervised topic inference strategy can be used to extend the applicability of our solutions to applications where predefined categories are not available. Finally, although relevance, novelty and diversity of recommendations may seem conflicting objectives, it is possible to effectively increase novelty and diversity with only a slight impact on relevance.

3 FINAL CONSIDERATIONS

Besides the research contributions summarized above and our various publications in conferences and journals of high impact, it is worth mentioning that we are currently working with new research and industrial/commercial applications of our thesis's results. In a cooperation with the Industry, we are adapting our methods to the e-commerce context, along with an exploitation of new objectives (revenue, consumers' satisfaction), new targets (collections of objects) and integrating product search and tag recommendation in a virtuous cycle of reinforcement. This shows the wide range of applicability of our work.

REFERENCES

- [1] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. 1999.
- [2] F. Belém, C. Batista, R. Santos, J. Almeida, and M. Gonçalves. Beyond relevance: exploiting novelty and diversity in tag recommendation. *ACM Transactions on Intelligent Systems Technology*, 7:26:1–26:34, 2016.
- [3] F. Belém, E. Martins, T. Pontes, J. Almeida, and M. Gonçalves. Associative Tag Recommendation Exploiting Multiple Textual Features. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1033–1042, 2011.
- [4] Y. Choi. A complete assessment of tagging quality: A consolidated methodology. *Journal of the Association for Information Science and Technology*, 66:798–817, 2015.
- [5] F. Figueiredo, F. Belém, H. Pinto, J. Almeida, and M. Gonçalves. Assessing the quality of textual features in social media. *Information Processing & Management*, 49:222–247, 2012.
- [6] N. Ifada and R. Nayak. How Relevant is the Irrelevant Data: Leveraging the Tagging Data for a Learning-to-Rank Model. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 23–32, 2016.
- [7] X. Li, L. Guo, and Y. E. Zhao. Tag-based social interest discovery. In *Proceedings of the 17th International Conference on World Wide Web*, pages 675–684, 2008.
- [8] M. Lipczak and E. Milios. Efficient tag recommendation for real-life data. *ACM Transactions on Intelligent Systems Technology*, 3(1):2:1–2:21, 2011.
- [9] E. F. Martins, F. M. Belém, J. M. Almeida, and M. A. Gonçalves. On cold start for associative tag recommendation. *Journal of the Association for Information Science and Technology*, 67(1):83–105, 2016.
- [10] S. Rendle and L. Schmidt-Thie. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pages 81–90, 2010.
- [11] I. S. Ribeiro, R. L. Santos, M. A. Gonçalves, and A. H. Laender. On tag recommendation for expertise profiling: A case study in the scientific domain. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 189–198, 2015.
- [12] M. Saveski and A. Mantrach. Item cold-start recommendations: Learning local collective embeddings. In *Proceedings of the ACM Conference on Recommender Systems*, pages 89–96, 2014.