A Semantic Hybrid Approach for Sound Recommendation

Vito Claudio Ostuni[1], Sergio Oramas[2], Tommaso Di Noia[1], Xavier Serra[2], Eugenio Di Sciascio[1]
[1] Polytechnic University of Bari – Bari, Italy,
[2] Music Technology Group, Universitat Pompeu Fabra – Barcelona, Spain
{vitoclaudio.ostuni, tommaso.dinoia, eugenio.disciascio}@poliba.it
{sergio.oramas,xavier.serra}@upf.edu

ABSTRACT

In this work we describe a hybrid recommendation approach for recommending sounds to users by exploiting and semantically enriching textual information such as tags and sounds descriptions. As a case study we used Freesound, a popular site for sharing sound samples which counts more than 4 million registered users. Tags and textual sound descriptions are exploited to extract and link entities to external ontologies such as WordNet and DBpedia. The enriched data are eventually merged with a domain specific tagging ontology to form a knowledge graph. Based on this latter, recommendations are then computed using a semantic version of the feature combination hybrid approach. An evaluation on historical data shows improvements with respect to state of the art collaborative algorithms.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval

Keywords

Hybrid Recommender System; Ontologies; Tagging Systems; Entity Linking; Music; Aggregate Diversity; Novelty

1. A KNOWLEDGE GRAPH FOR SOUNDS DESCRIPTION

Freesound[1] is an online collaborative sound database where people with diverse interests share recorded sounds. When creators upload sounds they can tag and add textual descriptions while other users can comment and download those sounds. We used this information as a basis to feed a recommendation engine. A tagging ontology developed by [2] was initially used to classify the tags associated to sounds and build an initial knowledge graph which was further enriched by semanticing text-based data. In particular, we exploited tags and descriptions, to link more entities to the ones already present in the tagging ontology, thus enriching our knowledge graph. The first step in this direction was to link and disambiguate tags and keywords to Linked Open Data[2] resources. For this purpose we adopted Babelfy, a state of the art tool for Entity Linking and Word Sense Disambiguation [3]. Given a text, Babelfy returns the related WordNet *synset*, and/or the related DBpedia[3] URI associated with a set of related categories. We enriched our initial knowledge graph by (1) further expanding the Wordnet synsets considering their hypernymy relations; (2) retrieving the direct broader categories of the DBpedia entity. A snapshot of the final knowledge graph is available at `http://sisinflab.poliba.it/bag/www2015/#fs-wordnet-dbpedia-kg`.

This work is partly supported by the European Research Council FP7-2007-2013 ERC grant agreement 267583 (CompMusic)

[1] `https://www.freesound.org/`

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742775.

2. RECOMMENDATION APPROACH

In our scenario we adopt a hybrid recommendation approach to leverage both collaborative information coming from the user's sounds download data and content information coming from sound textual descriptions, tags and ontological knowledge. Following the taxonomy of hybrid recommender systems presented in [1] we develop a hybrid feature combination recommender system. Following this scheme we build item content feature vectors by considering both item graph-based descriptions represented in the knowledge graph and textual data and we then enrich such feature vectors with collaborative features. Subsequently, we use the obtained data to feed a content-based recommender system. Since we deal primarily with graph-based item representations we adopt the framework based on graph kernels presented in [5] as core content-based recommender which we briefly describe hereafter.

Let us formally define the knowledge graph as a multi-relational graph $G = \{t \mid t \in E \times R \times E\}$, where E denotes the set of entities and R indicates the set of properties or relations, namely the edge labels. With E_i^h we denote the set of entities reachable in *at most* h hops from i according to the shortest path in G. For a generic item i we then define its h-hop neighborhood graph $G_i^h = \{t = (e_i, r_j, e_k) \mid t \in E_i^h \times R \times E_i^h\}$ that is the subgraph of G induced by the set of triples involving entities in E_i^h. Starting from those item graph-based representations an explicit feature mapping based on entities is defined as follows.

[2] `http://linkeddata.org/`
[3] `http://dbpedia.org`

Entity-based item neighborhood mapping. Each feature refers to an entity in E and the corresponding score represents the weight associated to that entity in G_i^h. The resulting feature vector $\phi_E(G_i^h)$ is:

$$\phi_E(G_i^h) = (w_{i,e_1}, w_{i,e_2}, ...w_{i,e_m}, ..., w_{i,e_t})$$

where the weight associated to the generic entity e_m is computed as follows:

$$w_{i,e_m} = \sum_{l=1}^{h} \frac{|\{(e_n, p, e_m) \mid e_n \in \widehat{E}_i^l \vee e_m \in \widehat{E}_i^l\}|}{1 + \log(l)} \quad (1)$$

In Equation (1), $\widehat{E}_i^l = E_i^l \setminus E_i^{l-1}$ is the set of entities *exactly* l hops far from i. In particular, the numerator of the formula corresponds to the number of triples involving e_m, that is the *occurrence* of the entity e_m in the item neighborhood at distance l. The denominator can be seen as a decay factor depending on the distance l from the item i, whose aim is to incrementally penalize farther entities from the item.

Feature Combination. Since besides ontological knowledge there is also availability of different information sources such as usage download data and sound textual descriptions we use all of them to obtain the set of features best describing the items. Hence, in the feature vector we encode: graph-based features extracted from the knowledge graph, collaborative features from user implicit feedback, raw tags and keywords extracted from sound textual descriptions. We combine all those features by concatenation since we want to represent all different feature sources separately. Collaborative features are simply added by encoding in the feature vector those users who downloaded that item.

3. EXPERIMENTAL EVALUATION

We evaluated our approach on a historical data about sound downloads collected from February 2005 to October 2013 consisting in 20,000 users, 21,552 items and 12,132,548 downloads. We adopted the `All Unrated Items` methodology presented in [7] and used Precision and Recall as performance metrics for measuring recommendation accuracy. Following the hold-out strategy we used 65% of user downloads for building the training set, 15% for cross-validation and remaining 20% as test data. We used cross-validation for tuning the hyper-parameters of the Support Vector Regression algorithm used for learning the user model. For computing item feature vectors for each item we considered its 4-hops neighborhood graph for computing the entity-based mapping and added textual, tags and collaborative features as detailed before. For validating our approach we compared it with five state of the art recommendation algorithms. `MostPop` is a popularity-based baseline. `BPR-MF` [6] is a matrix factorization-based method optimized with Bayesian Personalized Ranking optimization criterion. `BPR Linear` is a hybrid version of BPR-MF. We used keywords and tags as item attribute data. `SLIM` [4] is based on a Sparse Linear method. `BPR-SLIM` is similar to `SLIM` but uses the BPR optimization criterion. The computation of the recommendations for all these comparative algorithms has been done with the publicly available software library *MyMediaLite*[4]. Figure 1 shows the precision-recall results we obtained in our experiments. We can see that our approach largely outperforms the others (all differences are statistically significant ($p<0.01$) according to the paired t-test). The only method which is close to the approach we propose is BPR-SLIM.

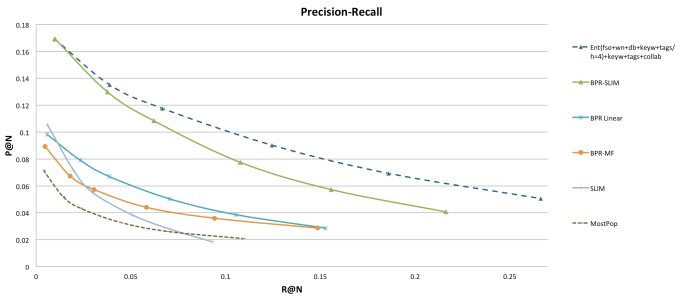

Figure 1: Precision-Recall results.

4. CONCLUSIONS

In this paper we presented an approach for computing semantic hybrid sound recommendations using Freesound as a case study. Preliminary experimental results showed that the proposed approach is able to improve recommendation accuracy with respect to several state of the art collaborative filtering algorithms. As future work we plan to better investigate the contribution of the different components of the proposed technique and measure other performances besides accuracy such as novelty and diversity.

5. REFERENCES

[1] R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, Nov. 2002.

[2] F. Font and S. Oramas. Extending Tagging Ontologies with Domain Specific Knowledge. *Interantional Semantic Web Conference (ISWC 2014)*, pages 1–4, 2014.

[3] A. Moro, A. Raganato, and R. Navigli. Entity Linking meets Word Sense Disambiguation : a Unified Approach. *Transactions of the Association for Computational Linguistics (TACL)*, 2014.

[4] X. Ning and G. Karypis. Sparse linear methods with side information for top-n recommendations. In *Proceedings of the sixth ACM conference on Recommender systems*, RecSys '12, pages 155–162, 2012.

[5] V. C. Ostuni, T. Di Noia, R. Mirizzi, and E. Di Sciascio. A linked data recommender system using a neighborhood-based graph kernel. In *The 15th International Conference on Electronic Commerce and Web Technologies*, 2014.

[6] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, UAI '09, pages 452–461, Arlington, Virginia, United States, 2009. AUAI Press.

[7] H. Steck. Evaluation of recommendations: Rating-prediction and ranking. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, 2013.

[4]http://www.mymedialite.net/