Enhancing Recommender Systems Using Linked Open Data-Based Semantic Analysis of Items

Rouzbeh Meymandpour and Joseph G. Davis
School of Information Technologies
The University of Sydney
NSW 2006, Australia
{rouzbeh.meymandpour, joseph.davis}@sydney.edu.au

Abstract

The Linked Open Data (LOD) project is a community effort that aims to publish structured data using open and liberal licences. The LOD cloud provides free access to datasets in diverse areas such as media, geography, publications and life sciences. These datasets are publicly available for machine and human consumption using Semantic Web standards and SPARQL endpoints. In addition to facilitating interoperability and integrity across diverse platforms, this movement not only opens up unique opportunities for developing novel and innovative applications but also makes the application development more efficient and cost-effective. This paper demonstrates how LOD can be a reliable and rich source of content information that supports recommender systems to overcome problems such as the item cold-start problem and limited content analysis that restrict many of the existing systems. By building on a robust measurement of the similarities between items using LOD, we present a hybrid recommender system that combines the semantic analysis of items with collaborative filtering approaches. The experimental evaluations of our proposed method using standard benchmark data and established measures show comparable overall accuracy and significant improvement in item cold-start situations.

Keywords: Linked Open Data, recommender systems, Semantic Web, Linked Data, recommendation, cold-start, similarity measures, Web of Data, content-based filtering, collaborative filtering, partitioned information content (PIC).

1 Introduction

Recommender systems provide users with information that helps them find products or items they are looking for. Their primary aim is to suggest a list of items in which the users may be interested. Recommender systems can be classified into two main categories: *content-based filtering (CBF)* and *collaborative filtering (CF)*. CBF approaches consider user- or item-specific information to identify and recommend items (e.g. books, movies, etc.) in which the user might be interested. In contrast, CF techniques analyse patterns in user ratings to identify groups of users with similar taste. Despite the extensive amount of research since the mid-1990s and during the Netflix Prize[1] competition from 2006 to 2009, several challenges still exist in the design and evaluation of recommender systems.

Although CF techniques are generally more accurate, they often suffer from the *item cold-start problem*; the lack of ratings provided by users for newly added items prevents them from being considered in the recommendations. On the other hand, CBF approaches successfully overcome the item cold-start problem. However, they struggle to deal with issues such as the user cold-start problem, overspecialisation and limited content analysis.

This paper demonstrates how structured content available through the community-driven effort of Linked Open Data (LOD) can be leveraged to support recommender systems in overcoming problems such as the item cold-start problem and limited capabilities for analysing item information. It presents a hybrid approach that combines the semantic analysis of items using LOD with collaborative filtering approaches. Experimental evaluation of this approach using well-established performance measures and benchmark datasets showed comparable overall accuracy and major improvement in item cold-start situations.

2 Using Linked Open Data to Enhance Recommender Systems

Semantic Web technologies such as RDF (resource description framework) allow publishing structured data in a standard manner that can be readily consumed by machines and shared across multiple applications. This transforms the conventional *Web of Documents*, associated with Web 1.0, into the *Web of Data*. The basic idea of the Web of Data, also known as *Linked Data*, is to publish structured data in a standard way on the Web and interlink them to other Linked Data in various domains. Linked Data principles ensure that the published data is represented using the standard RDF format and accessible for exploration via query languages such as SPARQL [Berners-Lee 2006].

Supported by Semantic Web standards and technologies, the *Linked Open Data (LOD)* project is a community effort that aims to publish Linked Data using open and liberal licences [Bizer et al. 2009]. LOD-based datasets

[1] In 2006, the online DVD rental service Netflix announced a one million dollar prize for the best CF method that improves the Netflix's own rating prediction algorithm by 10 per cent. The prize was awarded to 'BellKor's Pragmatic Chaos' team in 2009, who achieved 10.06% better predictions. (see http://www.netflixprize.com/)

are available for public use in a wide range of domains, from encyclopaedic knowledge bases to scientific data sources. The LOD cloud[2] provides free access to datasets in areas such as entertainment, geographic information, publications, life sciences and government data (see Figure 1). It contains structured information on diverse entities (i.e. resources or nodes in Linked Data) such as movies, actors, musicians, books, universities, cities, etc.

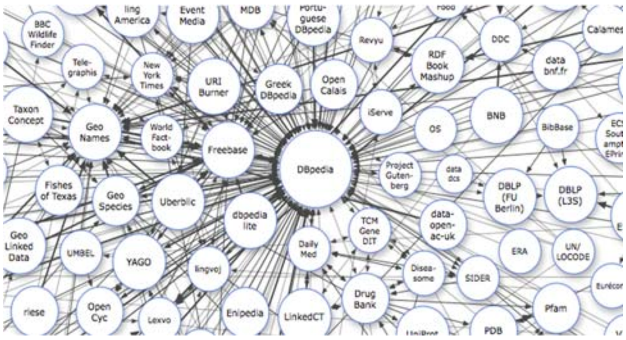

Figure 1. Part of the Linked Open Data (LOD) Diagram (source: Cyganiak and Jentzsch [2011])

Recommender systems, in general, and content-based filtering (CBF) approaches, in particular, are often required to analyse a large corpus of textual, unstructured information to gain knowledge about items. Due to the ambiguities associated with natural language processing, they suffer from limited content analysis. When item information is not available in a structured format, CBF systems have to deal with issues such as polysemy (same words can refer to different concepts), synonyms (multiple words can have similar meanings) and different forms of a word (such as singular forms, plural forms, and acronyms) [Lops et al. 2011]. In contrast, obtaining structured information about items from reliable sources such as Linked Open Data provides recommender systems with several potential benefits such as the low cost of acquiring structured information and the possibility of in-depth semantic analysis of items.

In order to be able to perform systematic semantic analysis of items using LOD, specific measures have to be developed to analyse and compare resources in Linked Open Data. The next section presents a systematic approach for measuring the semantic similarities of resources based on the concept of information content.

3 System Architecture and Approach

In this section, we present our hybrid recommender system as an *item-based collaborative filtering (IBCF)* method that uses Linked Open Data (LOD) for semantic analysis of items. IBCF is a *nearest neighbourhood*-based approach that considers items similar to those liked by the target user [Schafer et al. 2007; Ekstrand 2010]. In conventional IBCF recommender systems, item similarities are computed by analysing the way users have rated the items: "two items are similar if several users of the system have rated these items in a similar fashion" [Desrosiers and Karypis 2011, p112]. Cosine similarity and Pearson's correlation coefficient are among the similarity measures widely used in nearest neighbourhood-based recommender systems for computing item similarities [Schafer et al. 2007; Ekstrand 2010]. However, enough ratings have to be collected from the users before the system can provide any recommendation (the cold-start problem). In the following sections, we first summarise our LOD-based approach for semantic analysis of items. We proceed to present our hybrid recommendation approach.

3.1 Semantic Similarity Analysis Using Linked Open Data

In order to enable the Linked Open Data (LOD)-based comparison and analysis of items in recommender systems, a systematic approach that considers specific characteristics of LOD is required. *Semantic similarity* measures evaluate the degree of overlap between entities based on a set of pre-defined factors such as taxonomic relationships, particular characteristics of the entities or statistical information derived from the underlying knowledge base. They are fundamental to several fields including psychology, computational linguistics, artificial intelligence (AI) and natural language processing (NLP). However, previous approaches for semantic similarity measurement are largely restricted to the structural representation of their underlying domains. For example, feature-based methods such as Tversky [1977], Jaccard [1901] and Dice [1945] compute the similarity between entities or concepts based on the number of shared features between them. Therefore, they are only applicable in contexts where concepts or entities are represented using sets of features. Distance-based metrics such as the shortest path method proposed by Rada et al. [1989] are developed for taxonomies and only consider 'is-a' relations, while LOD is characterised by many different kinds of links, of which the 'is-a' relation (expressed by $rdf{:}type$[3] and $rdfs{:}subClassOf$[4] properties) is only one type. Another series of semantic similarity measures developed for taxonomies is known as the *information content (IC)*-based measures where the importance of the factors influencing the similarity value is derived based on their informativeness, that is, the amount of information conveyed by their presence. They showed a higher accuracy compared to the feature-based and distance-based measures (e.g. see Resnik [1995], Jiang and Conrath [1997] and Lin [1998]).

Information theory, as proposed by Claude E. Shannon [1948], describes the mathematical foundations of communication; transmitting information over communication channels by means of coding schemes (i.e. *source codes*). Based on earlier work by Hartley [1928], Shannon's key achievement was to define information as a measurable mathematical quantity, *information content (IC)*. Shannon [1948] describes information content as a measure of information conveyed by the occurrence of a probable event. Based on its information theory and probability theory foundations, information content-based measurement is

[2] http://lod-cloud.net/

[3] 'rdf:' is the prefix for the namespace http://www.w3.org/1999/02/22-rdf-syntax-ns#

[4] 'rdfs:' is the prefix for the namespace http://www.w3.org/2000/01/rdf-schema#

premised on the notion that common events are less informative than distinctive, infrequent events.

Information content measurement has been successfully used in a wide range of applications (such as data compression and transmission) to estimate the importance or informativeness of an event, a term or a message chosen from a set of possible events, terms or messages. For example, in data compression, the more frequent terms in a corpus are considered to be less informative. Therefore, they can be stored using fewer bits.

In order to assess the semantic similarity of items using Linked Open Data, we adopt the partitioned information content (PIC) [Meymandpour and Davis 2013], a measure of information content designed for resources in Linked Data. It combines the simplicity of feature-based approaches with the accuracy of information content-based methods.

Features of a resource are defined as triples of kind $(relationType,\ targetResource,\ direction)$. The type of the relation, the target node (the node connected to the other end of the relation) and the direction of the relation (incoming [In]/outgoing [Out]) are considered in the definition of the features. As a simple illustration, the features of nodes a and b in Figure 2 are the sets F_a and F_b, respectively:
$F_a = \{(l_1, c, Out), (l_2, d, In), (l_3, e, Out), (l_4, f, Out)\}$
$F_b = \{(l_2, d, In), (l_4, e, Out), (l_4, f, Out), (l_5, g, Out)\}$.
Thus: $F_a \cap F_b = \{(l_2, d, In), (l_4, f, Out)\}$

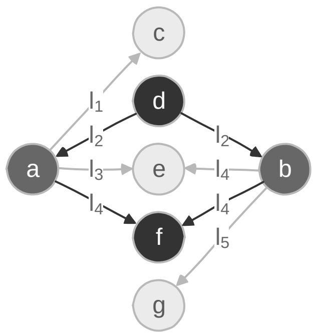

Figure 2. An example of resources and features in the Linked Data graph (a, b, … g are the resources and l_1, l_2, … l_5 are the links)

Based on this, PIC is defined as follows:

$$PIC(r) = \sum_{\forall f_i \in F_r} -log\left(\frac{\varphi(f_i)}{N}\right) \qquad (1)$$

The partitioned information content of the resource r, is the aggregate amount of information content conveyed by the features of the resource ($\forall f_i \in F_r$) ($PIC(r) \geq 0$). In this definition, $\varphi(f_i)$ is the frequency of the feature f_i in the underlying LOD dataset and N is the total number of resources in the underlying Linked Data.

The characteristics of PIC are derived from its information theory fundamentals. The logarithm is usually calculated to the base two; therefore, PIC is measured in units of information called *bits*. It is premised on the notion that highly probable features are general and less informative, while distinctive features, that is, features with a low number of occurrences, are more specific and convey more information. For example, based on the frequency of features in DBpedia (see Section 4.1), the fact that all actors are a 'Person' (represented using the feature $(rdf\!:\!type, foaf\!:\!Person, Out)$[5]) is substantially more popular than the fact that a particular actor starred in a movie (represented using the feature $(starring, movieURI, In)$). The former applies to millions of resources in DBpedia that describe a person, while the latter is only used when representing the actors of a movie (specified with $movieURI$). The frequency of the latter is equal to the number of actors who starred in the movie; therefore, it is more informative than the former (see Meymandpour and Davis [2013]).

Given the notion of the partitioned information content (PIC) of resources in LOD, our similarity measure can be derived by computing the PIC of shared and distinctive features of two resources:

$$PICSS(a,b) = \frac{PIC(F_a \cap F_b)}{PIC(F_a \cap F_b) + PIC(F_a - F_b) + PIC(F_b - F_a)} \qquad (2)$$

where the two resources that are being compared, that is, a and b, are represented as their sets of features F_a and F_b, respectively.

The similarity score computed by PICSS is normalised between 0 and 1, where the score of 0 represents no similarity between resources (perfectly dissimilar resources) and 1 represents a perfect similarity (identical resources).

The similarity value computed by PICSS is increased with more shared features and decreased with differences between resources. PICSS enables recommender systems to perform in-depth analysis of entities and to establish detailed comparison based on semantics acquired from Linked Open Data. In contrast to CF-based similarity measures, it is independent of the ratings provided by the users. Therefore, it is not biased by the popularity of items caused by the excessive attention of users or the lack of ratings on newly added items (the item cold-start problem).

3.2 A Hybrid Recommendation Approach

Despite the superior accuracy of collaborative filtering (CF) techniques (e.g. see Pilászy and Tikk [2009]), they often suffer from the item cold-start problem. In contrast, content-based filtering (CBF) approaches successfully overcome the item cold-start problem. However, they often struggle dealing with issues such as the user cold-start problem, overspecialisation and limited analysis of items. Therefore, a hybrid approach that balances between various aspects of the quality of the recommendations seems to be the optimal solution [Schein et al. 2002].

We present our hybrid recommender system as an item-based collaborative filtering (IBCF) method that uses our Linked Open Data-based semantic similarity measure, PICSS (Equation (2)) to assess the similarity between items using their corresponding resources on

[5] 'foaf:' is the prefix for http://xmlns.com/foaf/0.1/

LOD. Once the most similar items to the target, unrated item are identified, the prediction is computed based on the weighted sum of similarities between the target item and the similar items rated by the user [Schafer et al. 2007; Ekstrand 2010]:

$$\tilde{r}_{ui} = \frac{\sum_{j \in I_u^k} PICSS(i,j)\, r_{uj}}{\sum_{j \in I_u^k} |PICSS(i,j)|} \quad (3)$$

where $\tilde{r}_{ui}$ is the predicted rating of the user u on the item i, I_u^k is the set of k most similar items to i rated by the user u, r_{ui} is the rating score given to i by the user u and $Sim(i,j)$ is the similarity between items i and j.

In order to incorporate individual user- and item-specific effects (e.g. some users tend to give higher ratings than others and popular items often receive higher ratings than others), a first-order approximation of user and item biases can be added to the model [Koren et al. 2009; Koren 2010a]:

$$\tilde{r}_{ui} = b_{ui} + \frac{\sum_{j \in I_u^k} PICSS(i,j)\,(r_{uj} - b_{ui})}{\sum_{j \in I_u^k} |PICSS(i,j)|} \quad (4)$$

such that

$$b_{ui} = \mu + b_u + b_i \quad (5)$$

The term μ denotes the global average; the overall average of ratings. The bias terms b_u and b_i are the deviations of the ratings given by the user u and the ratings given to the item i, respectively, from the global average. The user and item biases, b_u and b_i, can be learned using a regularised model (see Koren [2010b]).

In the next section, we conduct experimental evaluations to assess the performance of the presented recommendation method.

4 Experimental Context and Platform

In this section, we report our experimental context and datasets used for the experimental evaluation of the presented recommendation approach.

4.1 Experimental Datasets

The primary dataset used for our experiments was DBpedia,[6] one of the most successful initiatives developed based on the Linked Open Data (LOD) principles [Auer et al. 2007]. We used the English version of DBpedia 3.8, released on August 2012.

In addition to DBpedia, in some parts of the experiments, we used several datasets from the LOD cloud, namely, Freebase,[7] LinkedMDB[8] and YAGO.[9]

4.2 Evaluation Datasets and Metrics

A well-established benchmark dataset widely used in the recommender systems community is MovieLens[10] [Herlocker et al. 2004]. We employed the MovieLens100K and MovieLens1M datasets. MovieLens100K contains 100,000 integer-rating scores (in 1-5 range) by 943 users on 1,682 movies. MovieLens1M provides around one million ratings for 3,883 movies given by 6040 users.

In order to evaluate our LOD-based recommender system using these datasets, we had to link items in the MovieLens datasets to their corresponding resources in DBpedia (see Table 1). We had to match the title of movies in the MovieLens datasets with the rdfs:label property of resources in DBpedia and (if provided) their release year. For example, after manual double-checking the results, we found the exact match for 1,569 items (93.3%) in the MovieLens100K dataset.[11]

Table 1. A sample rating in DBpedia-MovieLens dataset

User Id	17
Item Id	858
Item Name	Godfather, The (1972)
Respective Resource URI	dbr:The_Godfather[a]
Rating	5

[a] 'dbr:' is the prefix for http://dbpedia.org/resource/

We performed five-fold cross-validations on five randomly-split test (20%) and training (80%) sets. The average results are presented.

In addition to the evaluation of the rating prediction accuracy based on all available ratings, we assessed methods in an extreme item cold-start situation: the methods were evaluated for providing recommendations for items without any ratings given by users (known as the strict cold-start). In this experiment, items in the test set without any ratings in the training set were used to simulate the item cold-start situation.

The rating prediction accuracy of the evaluated methods was assessed using RMSE (root mean square error) that puts more weight on larger errors. RMSE is calculated by measuring the average square error between predictions ($\tilde{r}_{ui}$) and the actual ratings (r_{ui}) for all items in the test set (T):

$$RMSE = \sqrt{\frac{1}{|T|} \sum_{(u,i) \in T} (\tilde{r}_{ui} - r_{ui})^2} \quad (6)$$

All experiments were performed using an external evaluation platform developed using the MyMediaLite library[12] [Gantner et al. 2011].

[6] http://dbpedia.org/

[7] http://www.freebase.com/

[8] http://linkedmdb.org/

[9] http://www.yago-knowledge.org/

[10] http://www.grouplens.org/

[11] The DBpedia-MovieLens evaluation datasets and the mapping of MovieLens ItemIDs to DBpedia URIs as well as the mappings between DBpedia and Freebase, LinkedMDB and YAGO datasets are available for download at the following address: http://sydney.edu.au/engineering/it/~rouzbeh/files/DBpedia-MovieLens.zip

[12] http://mymedialite.net/

4.3 Evaluated Methods

PICSS was evaluated against a number of well-known methods. Several user-based and item-based nearest neighbourhood collaborative filtering methods were evaluated. The similarity measures used for the neighbourhood selection include cosine, Jaccard coefficient and Pearson's correlation coefficient [Karypis 2001; Sarwar et al. 2001; Candillier et al. 2007; Schafer et al. 2007; Su and Khoshgoftaar 2009]. These methods were implemented using the same approach to that presented in Equation (4). For comparison purposes, a baseline predictor considering only the baseline estimates (Equation (5)) is also evaluated.

In order to conduct a fair evaluation, the optimisation problem was optimised for the baseline predictor regardless of the choice of the similarity measure. It was optimised to maximise the item cold-start performance.

In order to study the proposed semantic similarity measures, two variations of our PICSS-based recommender system were evaluated: 1) $PICSS^{DBpedia}$ which only uses DBpedia as the source of items information and 2) $PICSS^{LOD}$ that employs several LOD datasets (as explained in Section 4.1) as the source of item information. By traversing the $owl{:}\,sameAs$[13] links, the features of resources are extracted from various datasets in the LOD cloud.

We also compared our hybrid recommender system against two state-of-the-art matrix factorization methods, namely, SVD++ and the integrated model [Koren 2008].[14] The latter method achieved the lowest RMSE among the others based on the Netflix Prize dataset.

5 Results

A detailed comparison of our hybrid recommender system against conventional and state-of-the-art recommender systems is reported in Table 2. In both benchmark datasets, that is, DBpedia-MovieLens100K (100K) and DBpedia-MovieLens1M (1M), $PICSS^{LOD}$ outperformed other methods in the new item cold-start situations. It also showed a comparable overall performance in our experimental evaluations.

The higher accuracy of $PICSS^{LOD}$ in the new item cold-start situations is observable. By achieving the new items RMSE of 0.993 and 0.994 on the 100K and 1M datasets, respectively, its new item cold-start performance was respectively 15.8% and 5.3% better than that of the baseline predictor (lower RMSE). Compared to the matrix factorization techniques, it performed significantly better for overcoming the new item cold-start problem. Based on the 100K benchmark dataset, it showed 14.6% lower RMSE compared to SVD++. The improvement is also noticeable based on the larger 1M dataset (8.8% lower RMSE). Compared to SVD++, the integrated model showed slightly better performance on the new item cold-start conditions; however, $PICSS^{LOD}$ performed 11.4% and 8.5% better than the integrated model on the 100K and 1M datasets, respectively.

[13] 'owl:' is the prefix for the namespace http://www.w3.org/2002/07/owl#

[14] Note that the performance of matrix factorization methods in various situations can be influenced by the choice of the parameter values (e.g. learning rate, number of factors, etc.). Therefore, the presented results are for comparison purposes only and are not necessarily an indication of the actual performance of the evaluated methods. In our experiments, we used the values suggested by the MyMediaLite 3.9 library.

Table 2. The results of five-fold cross-validations on DBpedia-MovieLens100K and DBpedia-MovieLens1M datasets; the minimum amount of RMSE in each column is shown in bold.

			DBpedia-MovieLens 100K		DBpedia-MovieLens 1M	
			All	*New Items*	*All*	*New Items*
Conventional		Baseline Predictor	0.947	1.180	0.914	1.050
	Item-Based CF	Cosine	0.922	1.180	0.878	1.050
		Jaccard	0.915	1.180	0.872	1.050
		Pearson Correlation	0.936	1.180	0.883	1.050
	User-Based CF	Cosine	0.935	1.180	0.898	1.050
		Jaccard	0.935	1.180	0.897	1.050
		Pearson Correlation	0.934	1.180	0.895	1.050
LOD-based		$PICSS^{DBpedia}$	0.916	1.074	0.882	1.011
		$PICSS^{LOD}$	**0.910**	**0.993**	0.875	**0.994**
MF		SVD++	0.947	1.164	0.892	1.090
		Integrated Model	0.931	1.121	**0.863**	1.086

$PICSS^{LOD}$ also showed a comparable overall accuracy. It achieved the lowest amount of overall RMSE compared to the evaluated methods on the 100K benchmark data. The overall RMSE of $PICSS^{LOD}$ was 3.9% lower than that of the baseline predictor. It also outperformed matrix factorization techniques on the 100K dataset. In comparison with the integrated model, $PICSS^{LOD}$ obtained 2.2% lower overall RMSE. Despite a slightly lower overall prediction accuracy (1.4% higher RMSE) on the 1M dataset compared to the integrated model, $PICSS^{LOD}$ showed promising performance by achieving 0.875 overall RMSE.

Significant enhancement of the LOD cloud version of PICSS over its DBpedia variant is also noticeable. $PICSS^{LOD}$ achieved higher accuracy both for overall and new items performance on both benchmark datasets. These differences were all statistically significant with at least 95% confidence ($p < 0.05$) based on the paired t-test (two-tailed) on the average of per-user prediction errors using five-fold cross-validations (see Shani and Gunawardana [2011]).

6 Discussion

We compared our approach against well-established, conventional user-based and item-based collaborative filtering (CF) techniques as well as advanced matrix factorization (MF) methods.

As collaborative filtering techniques rely solely on the ratings provided by users, they are unable to provide accurate recommendations in cold-start situations when no ratings are provided by the users. In these situations, one approach is to use a baseline predictor to deal with the cold-start problem. The baseline predictor used in our experiments (Equation (5)) estimates user ratings on unrated items based on the deviations of the ratings given by the target user and the ratings given to the target item from the global average. As shown in the experimental results, due to the fact that the evaluated CF methods use the same baseline estimates in cold-start situations, their predictions have the same amount of accuracy as the baseline predictor. In contrast, PICSS is not dependent on the user ratings for computing the item similarities. By incorporating the semantic analysis of items in the recommendation procedure, our hybrid approach showed significant enhancement in the item cold-start conditions over the conventional CF methods.

Matrix factorization is an effective way of extracting latent semantic factors from user ratings, allowing the recommender system to identify various aspects of users' preferences and their interactions with items. However, they are mainly dependent on an extensive amount of historical user rating data with access to 1M, 10M, or in the case of the Netflix Prize competition, more than 100M ratings. As our experiments showed, the evaluated MF approaches, namely, SVD++ and the integrated model, achieved low overall accuracy in the DBpedia-MovieLens100K dataset. By incorporating a neighbourhood model that considers the similarity between items, the integrated model performed better than SVD++ in cold-start situations. As expected, access to more user rating data, that is, using the larger DBpedia-MovieLens1M benchmark dataset, led them to provide predictions that are more accurate. Nevertheless, the performance of our approach in the cold-start conditions was significantly better on both datasets.

The results of the experimental evaluation showed meaningful differences between the variations of PICSS. Obtaining features of entities from multiple datasets in the LOD cloud in addition to DBpedia significantly increased the overall and new items performance of PICSS ($PICSS^{LOD}$). The differences between $PICSS^{DBpedia}$ and $PICSS^{LOD}$ confirm that PICSS not only succeeded in managing the availability of a large collection of features acquired from various LOD datasets, but also, its performance was significantly improved. The RMSE of $PICSS^{LOD}$ in the new item cold-start situations was lower by 7.5% on the 100K dataset and 1.7% on the 1M dataset when compared to $PICSS^{DBpedia}$. A slight (less than one per cent lower RMSE), but statistically significant improvement in the overall performance of $PICSS^{LOD}$ was also noticeable. Nevertheless, the accuracy of $PICSS^{DBpedia}$ that only uses DBpedia as the source of item information is also promising.

These experiments support that PICSS can effectively take advantage of the large amount of semantic content on the LOD cloud and provide robust predictions in all scenarios including strict cold-start situations.

7 Related Work

Several methods aimed to exploit Linked Open Data (LOD) for recommendation provision. However, they are often restricted to the semantic content of DBpedia. Moreover, none of the existing approaches was evaluated properly against well-established collaborative filtering (CF) methods and the state-of-the-art recommender systems. Passant [2010a, b] presented a distance measure based on the number of direct and indirect paths between resources in DBpedia. The distance measure was used for providing recommendations in the music domain. A small collection of link types was considered for computing paths between resources. In addition, as this approach is based only on the number of paths between the resources and all kinds of relations have the same importance in the distance function, the semantics of relations are not fully considered in the method. Furthermore, it is only applied on a manually cleaned dataset of DBpedia. In contrast, our proposed hybrid recommender system and semantic similarity measure are applicable in a wide range of domains. As showed in our experiments, PICSS can retrieve semantic content from various datasets on the LOD cloud without any need for extensive manual pre-processing.

A series of studies aimed to develop a content-based movie recommender system using LOD [Noia, Mirizzi, et al. 2012; Noia, Ostuni, et al. 2012]. The method applied a cosine similarity metric on a TF-IDF (term frequency-inverse document frequency)-based three-dimensional vector space model (VSM) that consists of movies, movie properties and values of properties extracted from DBpedia and LinkedMDB. The proposed recommendation approach showed improved performance compared to the method presented by Passant [2010a]. This method also demonstrated a higher accuracy in comparison with a user-based collaborative filtering technique (Pearson's correlation coefficient). As shown in our experimental evaluations (see Table 2) and reported by other authors [Karypis 2001; Sarwar et al. 2001], user-based collaborative techniques (UBCF) provide recommendations that are less accurate than those provided by item-based collaborative filtering techniques (IBCF). In contrast, we conducted extensive experimental evaluations that compare our approach against a wide range of established recommender systems including IBCF methods and recent matrix factorization techniques based on two standard benchmark datasets.

8 Conclusion

This paper demonstrated the applicability of Linked Open Data (LOD) for providing semantic analysis of items. The experiments showed that the accurate measurement of item similarities using LOD has the potential to improve the performance of recommender systems, especially, in situations where an insufficient amount of user ratings is available. The combination of semantic analysis of items with collaborative filtering-based recommendation in the proposed hybrid recommender system presented comparable overall accuracy, in addition to significant improvement in resolving the item cold-start problem.

The core of the presented approach is PICSS: the partitioned information content (PIC)-based semantic similarity measure, which is based solely on semantics retrieved from various datasets on the LOD cloud. As a pure content based similarity measure, it is not biased by the popularity of an item caused by the excessive attention of users or the lack of ratings on the newly added items.

The hybrid recommender system proposed in this paper was developed as an extension to current approaches, which makes it usable in combination with other methods. In terms of feasibility and scalability, once the similarity scores between items have been computed, for a newly added item, the system only needs to assess the similarity between the new item and others. This incremental update is advantageous compared to model-based CF techniques that require expensive computations to update the recommendation model and user/item profiles.

9 References

Auer, S., C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. 2007. "DBpedia: A Nucleus for a Web of Open Data." In *The Semantic Web*, edited by K. Aberer, et al., 722-735. Springer Berlin / Heidelberg.

Berners-Lee, T. *Linked Data - Design Issues*. World Wide Web Consortium (W3C) 2006 [cited 2014-04-14]. Available from http://www.w3.org/DesignIssues/LinkedData.html.

Bizer, C., T. Heath, and T. Berners-Lee. 2009. "Linked Data - The Story So Far." *International Journal on Semantic Web and Information Systems* no. 5 (3):1-22.

Candillier, L., F. Meyer, and M. Boulle. 2007. Comparing State-of-the-Art Collaborative Filtering Systems. In *the 5th International Conference on Machine Learning and Data Mining in Pattern Recognition*. Leipzig, Germany: Springer-Verlag.

Cyganiak, R., and A. Jentzsch. *Linking Open Data Cloud Diagram* 2011 [cited 2014-04-23]. Available from http://lod-cloud.net/.

Desrosiers, C., and G. Karypis. 2011. "A Comprehensive Survey of Neighborhood-based Recommendation Methods." In *Recommender Systems Handbook*, edited by F. Ricci, et al., 107-144. Springer.

Dice, L. R. 1945. "Measures of the Amount of Ecologic Association Between Species." *Ecology* no. 26 (3):297-302.

Ekstrand, M. D. 2010. "Collaborative Filtering Recommender Systems." *Foundations and Trends® in Human–Computer Interaction* no. 4 (2):81-173.

Gantner, Z., S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. 2011. MyMediaLite: a free recommender system library. In *the 5th ACM Conference on Recommender Systems*. Chicago, Illinois, USA: ACM.

Hartley, R. V. 1928. "Transmission of Information." *Bell System Technical Journal* no. 7 (3):535-563.

Herlocker, J. L., J. A. Konstan, L. G. Terveen, and J. T. Riedl. 2004. "Evaluating collaborative filtering recommender systems." *ACM Transactions on Information Systems* no. 22 (1):5-53.

Jaccard, P. 1901. "Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines." *Bulletin de la Société Vaudoise des Sciences Naturelles* no. 37:241-272.

Jiang, J. J., and D. W. Conrath. 1997. Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In *International Conference Research on Computational Linguistics (ROCLING X)*. Taiwan.

Karypis, G. 2001. Evaluation of Item-Based Top-N Recommendation Algorithms. In *the 10th International Conference on Information and Knowledge Management*. Atlanta, Georgia, USA: ACM.

Koren, Y. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Las Vegas, Nevada, USA: ACM.

Koren, Y. 2010a. "Collaborative filtering with temporal dynamics." *Communications of the ACM* no. 53 (4):89-97.

Koren, Y. 2010b. "Factor in the neighbors: Scalable and accurate collaborative filtering." *ACM Transactions on Knowledge Discovery from Data* no. 4 (1):1-24.

Koren, Y., R. Bell, and C. Volinsky. 2009. "Matrix Factorization Techniques for Recommender Systems." *Computer* no. 42 (8):30-37.

Lin, D. 1998. An information-theoretic definition of similarity. In *the 15th International Conference on Machine Learning*: Morgan Kaufmann, San Francisco, CA.

Lops, P., M. d. Gemmis, and G. Semeraro. 2011. "Content-based recommender systems: State of the art and trends." In *Recommender Systems Handbook*, edited by F. Ricci, et al., 73-105. New York: Springer.

Meymandpour, R., and J. G. Davis. 2013. "Linked Data Informativeness." In *Web Technologies and Applications*, edited by Y. Ishikawa, et al., 629-637. Springer Berlin Heidelberg.

Noia, T. D., R. Mirizzi, V. C. Ostuni, and D. Romito. 2012. Exploiting the web of data in model-based recommender systems. In *the 6th ACM Conference on Recommender Systems*. Dublin, Ireland: ACM.

Noia, T. D., V. C. Ostuni, D. Romito, R. Mirizzi, and M. Zanker. 2012. Linked Open Data to support Content-based Recommender Systems. In *the 8th International Conference on Semantic Systems (I-Semantics 2012)*. Graz, Austria.

Passant, A. 2010a. "dbrec — Music Recommendations Using DBpedia." In *The Semantic Web – ISWC 2010*, edited by P. Patel-Schneider, et al., 209-224. Springer Berlin / Heidelberg.

Passant, A. 2010b. Measuring Semantic Distance on Linking Data and Using it for Resources Recommendations. In *AAAI Spring Symposium Linked Data Meets Artificial Intelligence*.

Pilászy, I., and D. Tikk. 2009. Recommending new movies: even a few ratings are more valuable than metadata. In *the 3rd ACM Conference on Recommender Systems*. New York, New York, USA: ACM.

Rada, R., H. Mili, E. Bicknell, and M. Blettner. 1989. "Development and application of a metric on semantic nets." *IEEE Transactions on Systems, Man and Cybernetics* no. 19 (1):17-30.

Resnik, P. 1995. Using information content to evaluate semantic similarity in a taxonomy. In *the 14th International Joint Conference on Artificial Intelligence*.

Sarwar, B., G. Karypis, J. Konstan, and J. Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *the 10th International Conference on World Wide Web*. Hong Kong, Hong Kong: ACM.

Schafer, J. B., D. Frankowski, J. Herlocker, and S. Sen. 2007. "Collaborative Filtering Recommender Systems." In *The Adaptive Web*, edited by P. Brusilovsky, et al., 291-324. Springer Berlin / Heidelberg.

Schein, A. I., A. Popescul, L. H. Ungar, and D. M. Pennock. 2002. Methods and metrics for cold-start recommendations. In *the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Tampere, Finland: ACM.

Shani, G., and A. Gunawardana. 2011. "Evaluating Recommendation Systems." In *Recommender Systems Handbook*, edited by F. Ricci, et al., 257-297. New York: Springer.

Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* no. 27:379–423, 623–656.

Su, X., and T. M. Khoshgoftaar. 2009. "A Survey of Collaborative Filtering Techniques." *Advances in Artificial Intelligence* no. 2009:1-19.

Tversky, A. 1977. "Features of similarity." *Psychological Review* no. 84 (4):327-352.