Utilizing Online Social Network and Location-Based Data to Recommend Products and Categories in Online Marketplaces

Emanuel Lacic[1]([✉]), Dominik Kowald[1], Lukas Eberhard[2], Christoph Trattner[3], Denis Parra[4], and Leandro Balby Marinho[5]

[1] Know-Center, Graz University of Technology, Graz, Austria
{elacic,dkowald}@know-center.at
[2] IICM, Graz University of Technology, Graz, Austria
lukas.eberhard@tugraz.at
[3] Norwegian University of Science and Technology, Trondheim, Norway
chritrat@idi.ntnu.no
[4] Pontificia Universidad Catlica de Chile, Santiago, Chile
dparra@ing.puc.cl
[5] UFCG, Campina Grande, Brazil
lbmarinho@dsc.ufcg.edu.br

Abstract. Recent research has unveiled the importance of online social networks for improving the quality of recommender systems and encouraged the research community to investigate better ways of exploiting the social information for recommendations. To contribute to this sparse field of research, in this paper we exploit users' interactions along three data sources (marketplace, social network and location-based) to assess their performance in a barely studied domain: recommending products and domains of interests (i.e., product categories) to people in an online marketplace environment. To that end we defined sets of content- and network-based user similarity features for each data source and studied them isolated using an user-based Collaborative Filtering (CF) approach and in combination via a hybrid recommender algorithm, to assess which one provides the best recommendation performance. Interestingly, in our experiments conducted on a rich dataset collected from SecondLife, a popular online virtual world, we found that recommenders relying on user similarity features obtained from the social network data clearly yielded the best results in terms of accuracy in case of predicting products, whereas the features obtained from the marketplace and location-based data sources also obtained very good results in case of predicting categories. This finding indicates that all three types of data sources are important and should be taken into account depending on the level of specialization of the recommendation task.

Keywords: Recommender systems · Online marketplaces · SNA · Social data · Location-based data · SecondLife · Collaborative filtering · Item recommendations · Product recommendations · Category prediction

1   Introduction

Research on recommender systems has gained tremendous popularity in recent years. Especially since the hype of the social Web and the rise of social media and networking platforms such as Twitter or Facebook, recommender systems are acknowledged as an essential feature helping users to, for instance, discover new connections between people or resources. Especially in the field of e-commerce sites, i.e., online marketplaces, current research is dealing with the improvement of the prediction task in order to recommend products that are more likely to match peoples preferences.

Typically, these online systems calculate personalized recommendations using only one data source, namely marketplace data (e.g., implicit user feedback such as previously viewed or purchased products - see also e.g., [1–3]). Although this approach has been well established and performs reasonably well, nowadays, online marketplaces often also have the opportunity to leverage additional information about the users coming from social and location-based data sources (e.g., via Facebook-connect). Even though previous research has shown that this kind of data can be useful in the wide field of recommender systems (see Sect. 2), it remains an open problem how to fully exploit these additional data sources (social network and location-based data) to improve the recommendation task in online marketplaces.

Moreover, it is often not the most important thing in online marketplaces to predict the exactly right products to the users but to suggest domains of interests (i.e., product categories) the users could like and could use for further browsing (e.g., [1,4]). Thus, it is not only important to investigate to which extent social and location-based data sources can be used to improve product recommendations but also to which extent this data can also be used for the recommendation of product categories.

To contribute to this sparse field of research, in this paper we present a first take on this problem in form of a research project that aims at understanding how different sources of interaction data can help in recommending products and categories to people in an online marketplace. In this respect, we are particularly interested in studying the efficiency of different user similarity features derived from various dimensions, not only from the marketplace but also from online social networks and location-based data to recommend products and categories to people via a user-based Collaborative Filtering (CF) approach (we have chosen a user-based CF approach since user-based CF is not only a well-established recommender algorithm but also allows us to incorporate various user-based similarity features coming from different data sources, which has been shown to play an important role in making more accurate predictions [4,5]). Specifically, we raise the following two research questions:

- *RQ1*: To which extent can user similarity features derived from marketplace, social network and location-based data sources be utilized for the recommendation of products and categories in online marketplaces?

– *RQ2*: Can the different marketplace, social network and location-based user similarity features and data sources be combined in order to create a hybrid recommender that provides more robust recommendations in terms of prediction accuracy, diversity and user coverage?

In order to address these research questions, we examined content-based and network-based user similarity feature sets for user-based CF over three data sources (marketplace, social and location-based data) as well as their combinations using a hybrid recommender algorithm and assessed the results via a more comprehensive set of recommender evaluation metrics than previous works. The study was conducted using a large-scale dataset crawled from the virtual world of SecondLife. In this way, we could study the utility of each user similarity feature separately as well as combine them in the form of hybrid approaches to show which combinations, per data source and globally, provide the best recommendations in terms of recommendation accuracy, diversity and user coverage. Summing this up, the contributions of this work are the following:

– Contrary to previous work in this area [1,2], we not only employ one source of data (marketplace) for the problem of predicting product purchases but show how data coming from three different sources (marketplace, social and location-based data) can be exploited in this context.
– In contrast to related work in the field, we provide also an extensive evaluation of various content-based and network-based user similarity features via user-based Collaborative Filtering as well as their combinations via a hybrid recommender approach.
– Finally, we also provide evidence to what extent top-level and sub-level purchase categories can be predicted which is in contrast to previous work (e.g., [1]) where the authors only focused on the problem recommending top-level categories to the users.

To the best of our knowledge, this is the first study that offers such a comprehensive user similarity feature selection and evaluation for product and category recommendation in online marketplaces.

Overall, the paper is structured as follows: we begin by discussing related work in Sect. 2. Then we present the datasets (Sect. 3) and the feature description (Sect. 4) used in our extensive evaluation. After that, we present our experimental setup in Sect. 5 and show the results of our experiments in Sect. 6. Finally, on Sect. 7 we conclude the paper and discuss the outlook.

2   Related Work

Most of the literature that leverages social data for recommendations is focused on recommending users, (e.g., [2,6]), tags (e.g., [7]) or points-of-interest (e.g., [4]), although some works have exploited social information for item or product recommendation, being the most important ones model-based. Jamali et al. [5]

introduced SocialMF, a matrix factorization model that incorporates social relations into a rating prediction task, decreasing RMSE with respect to previous work. Similarly, Ma et al. [4] incorporated social information in two models of matrix factorization with social regularization, with improvements in both MAE and RMSE for rating prediction. Among their evaluations, they concluded that choosing the right similarity feature between users plays an important role in making a more accurate prediction.

On a more general approach, Karatzoglou et al. [8] use implicit feedback and social graph data to recommend places and items, evaluating with a ranking task and reporting significant improvements over past related methods. Compared to these state-of-the-art approaches, our focus on this paper is at providing a richer analysis of feature selection (similarity features) with a more comprehensive evaluation than previous works, and in a rarely investigates domain: product recommendation in a social online marketplace. For instance, in Guo et al. [2] or Trattner et al. [3] the authors leveraged social interactions between sellers and buyers in order to predict sellers to customers. Other relevant work in this context is the study of Zhang and Pennacchiotti [1] who showed how top-level categories can be better predicted in a cold-start setting on eBay by exploiting the user's "likes" from Facebook.

3 Datasets

In our study we rely on three datasets[1] obtained from the virtual world Second-Life (SL). The main reason for choosing SL over real world sources are manifold but mainly due to the fact that currently there are no other datasets available that comprise marketplace, social and location data of users at the same time. For our study we focused on users who are contained in all three sources of data, which are 7,029 users in total. To collect the data (see Table 1) we crawled the SL platform as described in our previous work [9, 10].

3.1 Marketplace Dataset

Similar to eBay, SecondLife provides an online trading platform where users can trade virtual goods with each other. Every seller in the SL marketplace[2] owns her own seller's store and publicly offers all of the store's products classified into different categories (a hierarchy with up to a maximum of four different categories per product). Furthermore, sellers can apply meta-data such as price, title or description to their products. Customers in turn, are able to provide reviews to products.

We extracted 29,802 complete store profiles, with corresponding 39,055 trading interactions, and 2,162,466 products, out of which 30,185 were purchased.

---

[1] **Note:** The datasets could be obtained by contacting the fourth author of this work.
[2] https://marketplace.secondlife.com/.

Table 1. Basic statistics of the SL datasets used in our study.

Marketplace dataset (Market)	
Number of products	$30,185$
Number of products with categories	$24,276$
Number of purchases	$39,055$
Number of purchases with categories	$31,164$
Mean number of purchases per user	5.56
Mean number of purchases per products	1.29
Mean number of categories per product	2.86
Number of top-level categories	23
Number of low-level categories	532
Mean number of top-level categories purchase	$1,354.96$
Mean number of low-level categories purchase	58.58
Number of sellers	$8,149$
Mean number of purchases per seller	3.70
Online social network dataset (Social)	
Number of interactions	$490,236$
Mean number of interactions	69.75
Number of groups	$39,180$
Mean number of groups per user	$9,419$
Number of interests	5.57
Mean number of interests per user	1.34
Location-based dataset (Location)	
Number of different favorite locations	$10,538$
Mean number of favorite locations per user	5.77
Number of different shared locations	$5,736$
Mean number of shared locations per user	1.94
Number of different monitored locations	$1,887$
Mean number of monitored locations per user	6.52

From the purchased products, 24,276 are described using categories which are differentiated in top-level categories and low-level categories (i.e., the assigned product categories on the lowest possible level of the category hierarchy). An example of a product in the marketplace of SL is shown in the first image of Fig. 1.

3.2   Online Social Network Dataset

The online social network MySecondLife[3] is similar to Facebook with regard to postings: users can interact with each other by sharing text messages and

(a) SecondLife store

(b) SecondLife social stream
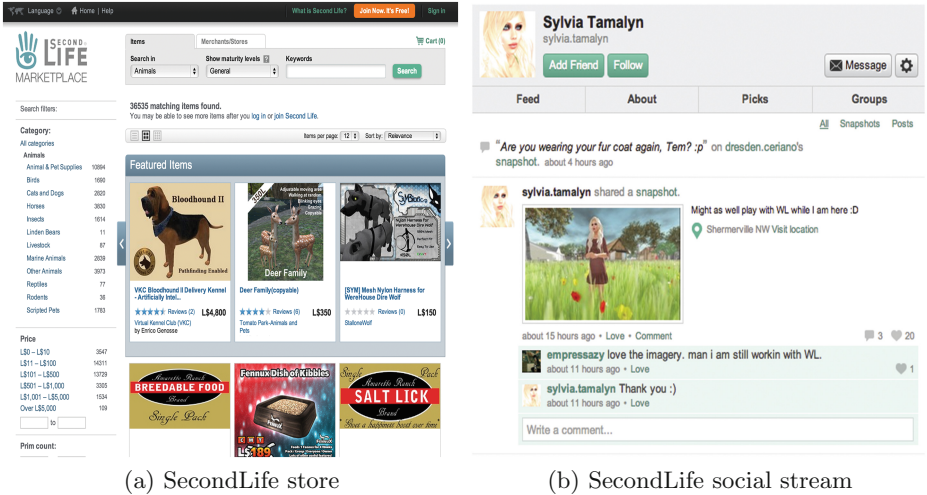
Fig. 1. Examples for a store in the marketplace and a social stream of an user in the online social network of the virtual world SecondLife.

commenting or loving (= liking) these messages. From the extracted 7,029 complete user profiles, we gathered 39,180 different groups users belong to, 9,419 different interests users defined for themselves and 490,236 interactions between them. The second image of Fig. 1 shows an example of an user profile in the online social network of SL.

3.3   Location-Based Dataset

The world of Second Life is contained within regions, i.e., locations which are independent from each other. Overall, we extracted three different sources of location-based data in our experiments:

(a) Favored Locations: Every user of SL can specify up to 10 so-called "Picks" in their profile representing her favorite locations that other users can view in the user's MySecondLife profile. We found that the extracted users picked 40,558 locations from 10,538 unique locations;

(b) Shared Locations: Users in SL can also share information about their current in-world position through in-world pictures called "snapshots", which also include in-world GPS information (similar as Foursquare). Overall, we identified 13,637 snapshots in 5,736 unique locations;

(c) Monitored Locations: As in real life, users in SL can create events in the virtual world and publicly announce them in a public event calendar. We collected these events, with an accurate location and start time, and extracted 157,765 user-location-time triples, with 1,887 unique locations.

4   Feature Description

As shown in our previous work (e.g., [10]), similarities between users can be derived in two different ways: either we calculate similarities between users on the content (= meta-data) provided directly by the user profiles or on the network structure of the user profiles interacting with each other. In the following sections we present more details about these ideas.

4.1   Content-Based User Similarity Features

We define our set of content-based user similarity features based on different types of entities or meta-data information that are directly associated with the user profiles in our data sources. In the case of the marketplace dataset these entities are purchased products, product categories and sellers of the products, in the case of the social network these are groups and interests the users have assigned, and in the case of the location-based data source these are favored locations, shared locations and monitored locations. Formally, we define the entities of a user u as $\Delta(u)$ in order to calculate the similarity between two users, u and v.

The first content-based user similarity feature we induce is based on the entities two users have in common. It is called *Common Entities* and is given by:

$$sim(u,v) = |\Delta(u) \cap \Delta(v)| \tag{1}$$

The second similarity feature, *Total Entities*, is defined as the union of two users' entities and is calculated by:

$$sim(u,v) = |\Delta(u) \cup \Delta(v)| \tag{2}$$

These two user similarity features are combined by *Jaccard's Coefficient for Entities* as the number of common entities divided by the total number of entities:

$$sim(u,v) = \frac{|\Delta(u) \cap \Delta(v)|}{|\Delta(u) \cup \Delta(v)|} \tag{3}$$

4.2   Network-Based User Similarity Features

In our experiments we consider all networks as an undirected graph $G\langle V, E\rangle$ with V representing the user profiles and $e = (u,v) \in E$ if user u performed an action on v (see also [10]). In the case of the social network, these actions are defined as social interactions, which are a combination of likes, comments and wallposts. In the case of the location-based dataset, actions between users are determined if they have met each other in the virtual world at the same time in the same location[4]. Furthermore, the weight of an edge $w_{action}(u,v)$ gives the frequency of

---

[4] **Note:** We derived the networks in our study from the location-based dataset only for the monitored locations, since the exact timestamps are not available for the favored nor the shared locations in the datasets.

a specific action between two users u and v. Finally, this network structure also let us determine the neighbors of users in order to calculate similarities based on this information. We define the set of neighbors of a node $v \in G$ as $\Gamma(v) = \{u \mid (u, v) \in E\}$.

The first network-based user similarity feature we introduce, uses the number of *Directed Interactions* between two users and is given by:

$$sim(u, v) = w_{action}(u, v) \tag{4}$$

In contrast to *Directed Interactions*, the following user similarity features are based on the neighborhood of two users: The first neighborhood similarity feature is called *Common Neighbors* and represents the number of neighbors two users have in common:

$$sim(u, v) = |\Gamma(u) \cap \Gamma(v)| \tag{5}$$

To also take into account the total number of neighbors of the users, we introduced *Jaccard's Coefficient for Common Neighbors*. It is defined as:

$$sim(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \tag{6}$$

A refinement of this feature was proposed as *Adamic Adar* [11], which adds weights to the links since not all neighbors in a network have the same tie strength:

$$sim(u, v) = \sum_{z \in \Gamma(u) \cap \Gamma(v)} \frac{1}{log(|\Gamma(z)|)} \tag{7}$$

Another related similarity feature introduced by Cranshaw et al. [12], called *Neighborhood Overlap*, measures the structural overlap of two users. Formally, this is written as:

$$sim(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u)| + |\Gamma(v)|} \tag{8}$$

The *Preferential Attachment Score*, first mentioned by Barabasi et al. [13], is another network-based similarity feature with the goal to prefer active users in the network. This score is the product of the number of neighbors of each user and is calculated by:

$$sim(u, v) = |\Gamma(u)| \cdot |\Gamma(v)| \tag{9}$$

5   Experimental Setup

In this section we provide a detailed description of our experimental setup. First, we describe the recommender approaches we have chosen in order to evaluate our three data sources (marketplace, social network and location-based data) as well as our derived user similarity features for the task of recommending products and categories. Afterwards, we describe the evaluation methodology and the metrics used in our study.

5.1   Recommender Approaches

In this subsection we describe the recommender approaches we have used to tackle our research questions described in the introductory section of this paper. All mentioned approaches have been implemented into our scalable big data social recommender framework *SocRec* [9,14], an open-source framework which can be obtained from our Github repository[5].

Baseline. As baseline for our study, we used a simple MostPopular approach recommending the most popular products or categories in terms of purchase frequency to the users.

Recommending Products. The main approach we adopt to evaluate our data sources and user similarity features for the task of recommending products is a *User-based Collaborative Filtering (CF)* approach. The basic idea of this approach is that users who are more similar to each other, e.g., have similar taste, will likely agree or rate on other resources in a similar manner [15]. Out of the different CF approaches, we used the non-probabilistic user-based nearest neighbor algorithm where we first find the k-nearest similar users and afterwards recommend the resources of those user as a ranked list of top-N products to the target user that are new to her (i.e., she has not purchased those products in the past). As outlined before, we have chosen this approach since user-based CF is not only a well-established recommender algorithm but also allows us to incorporate various user-based similarity features coming from different data sources, which has been shown to play an important role in making more accurate predictions [4,5].

The similarity values of the user pairs $sim(u, v)$ are calculated based on the user similarity features proposed in Sect. 4 (i.e., constructing the neighborhood). Based on these similarity values, each item i of the k most similar users for the target user u is ranked using the following formula [15]:

$$pred(u, i) = \sum_{v \in neighbors(u)} sim(u, v) \qquad (10)$$

Recommending Categories. For the task of recommending categories and in contrast to previous work (e.g., [1]), we are not only focusing here on the prediction of top-level categories but also on the prediction of low-level categories. The prediction of categories was implemented as an extension of product predictions. Thus, for each product in the list of recommended products (i.e., the products obtained from the k-nearest neighbors of user u based on a user similarity feature), we extracted the assigned category on the highest level in the case of predicting top-level categories and the assigned category on the lowest level in the case of predicting low-level categories. Afterwards, we assigned a score to

---

[5] https://github.com/learning-layers/SocRec.

each extracted category e_i in the set of all extracted categories E_u for the target user u based on a similar method as proposed in [1]:

$$pred(u, e_i) = \frac{purc(E_u, e_i)}{\sum\limits_{e \in E_u} purc(E_u, e)} \tag{11}$$

where $purc(E_u, e_i)$ gives the number of times the category e_i occurs in the set of all extracted categories E_u for user u.

Combining User Similarity Features and Data Sources. To further explore how to combine our data sources and features for recommendation, we investigated different hybridization methods (see also [16]). The hybrid approach chosen in the end is known as *Weighted Sum*. The score of each recommended item in the *Weighted Sum* algorithm is calculated as the weighted sum of the scores for all recommender approaches. It is given by:

$$W_{rec_i} = \sum_{s_j \in S} (W_{rec_i, s_j} \cdot W_{s_j}) \tag{12}$$

where the combined weighting of the recommended item i, W_{rec_i}, is given by the sum of all single weightings for each recommended item in an approach W_{rec_i, s_j} multiplied by the weightings of the recommender approaches W_{s_j}. We weighted each recommender approach W_{s_j} based on the nDCG@10 value obtained from the individual approaches (calculated based on an additional evaluation set where we withheld 20 purchased products - see Sect. 5.2).

We also experimented with other hybrid approaches, known as *Cross-source* and *Mixed Hybrid* [16]. However, these approaches have not yielded better results than the *Weighted Sum* algorithm.

5.2   Evaluation Method and Metrics

To evaluate the performance of each approach in a recommender setting, we performed a number of off-line experiments. Therefore, we split the SL dataset in two different sets (training and test set) using a method similar to the one described in [9], i.e., for each user we withheld 10 purchased products from the complete dataset and added them to the test set to be predicted. Since we did not use a *p*-core pruning technique to prevent a biased evaluation, there are also users with less than 10 relevant products. We did not include these users into our evaluation since they did not have enough purchase data available that could be used to produce reasonable recommendations based on the marketplace data, although this data is worthwhile for our user-based CF approach to find suitable neighbors. Thus, we used a post-filtering method, where all the recommendations were still calculated on the whole datasets but accuracy estimates were calculated only based on these filtered user profiles (= 959 users).

To finally quantify the performance of each of our recommender approaches, we used a diverse set of well-established metrics in recommender systems [17,18]. These metrics are as follows:

Recall (R@k) is calculated as the number of correctly recommended products divided by the number of relevant products, where r_u^k denotes the top k recommended products and R_u the list of relevant products of a user u in the set of all users U. Recall is given by [19]:

$$R@k = \frac{1}{|U|} \sum_{u \in U} \left(\frac{|r_u^k \cap R_u|}{|R_u|} \right) \tag{13}$$

Precision (P@k) is calculated as the number of correctly recommended products divided by the number of recommended products k. Precision is defined as [19]:

$$P@k = \frac{1}{|U|} \sum_{u \in U} \left(\frac{|r_u^k \cap R_u|}{k} \right) \tag{14}$$

Normalized Discounted Cumulative Gain (nDCG@k) is a ranking-dependent metric that not only measures how many products can be correctly predicted but also takes the position of the products in the recommended list with length k into account. The nDCG metric is based on the *Discounted Cummulative Gain (DCG@k)* which is given by [20]:

$$DCG@k = \sum_{k=1}^{|r_u^k|} \left(\frac{2^{B(k)} - 1}{log_2(1 + k)} \right) \tag{15}$$

where $B(k)$ is a function that returns 1 if the recommended product at position i in the recommended list is relevant. nDCG@k is calculated as DCG@k divided by the ideal DCG value iDCG@k which is the highest possible DCG value that can be achieved if all the relevant products would be recommended in the correct order. Taken together, it is given by the following formula [20]:

$$nDCG@k = \frac{1}{|U|} \sum_{u \in U} \left(\frac{DCG@k}{iDCG@k} \right) \tag{16}$$

Diversity (D@k), as defined in [17], can be calculated as the average dissimilarity of all pairs of resources in the list of recommended products r_u^k. Given a distance function $d(i, j)$ that is the distance, or the dissimilarity between two products i and j, D is given as the average dissimilarity of all pairs of products in the list of recommended products [17]:

$$D@k = \frac{1}{|U|} \sum_{u \in U} \left(\frac{1}{k \cdot (k - 1)} \sum_{i \in R} \sum_{j \in r_u^k, j \neq i} d(i, j) \right) \tag{17}$$

User Coverage (UC) is defined as the number of users for whom at least one product recommendation could have been calculated ($|U_r|$) divided by total number of users $|U|$ [21]:

$$UC = \frac{|U_r|}{|U|} \tag{18}$$

All mentioned performance metrics are calculated and reported based on the top-10 recommended products.

6   Results

In this section we highlight the results of our experiments for predicting products, low-level and top-level categories in terms of algorithmic performance in order to tackle our two research questions presented in Sect. 1. Our evaluation has been conducted in two steps: in the first step we compared the different recommender approaches with the corresponding user similarity features isolated (*RQ1*), see Table 2) and in the second step we combined these approaches in the form of hybrid recommendations (*RQ2*, see Table 3). All results are presented by recommender accuracy, given by nDCG@10, P@10 (Precision) and R@10 (Recall), D@10 (Diversity) and UC (User Coverage).

6.1   Recommendations Based on Single User Similarity Features

The results for the recommendation of products, low-level categories and top-level categories, using content-based and network-based user similarity features derived from our three data sources (marketplace, social and location-based data), are shown in Table 2 in order to address our first research question (*RQ1*). The results also include the *Most Popular (MP)* approach as a baseline. Additionally, the performance of the different data sources is also shown in Fig. 2 in form of Recall/Precision plots.
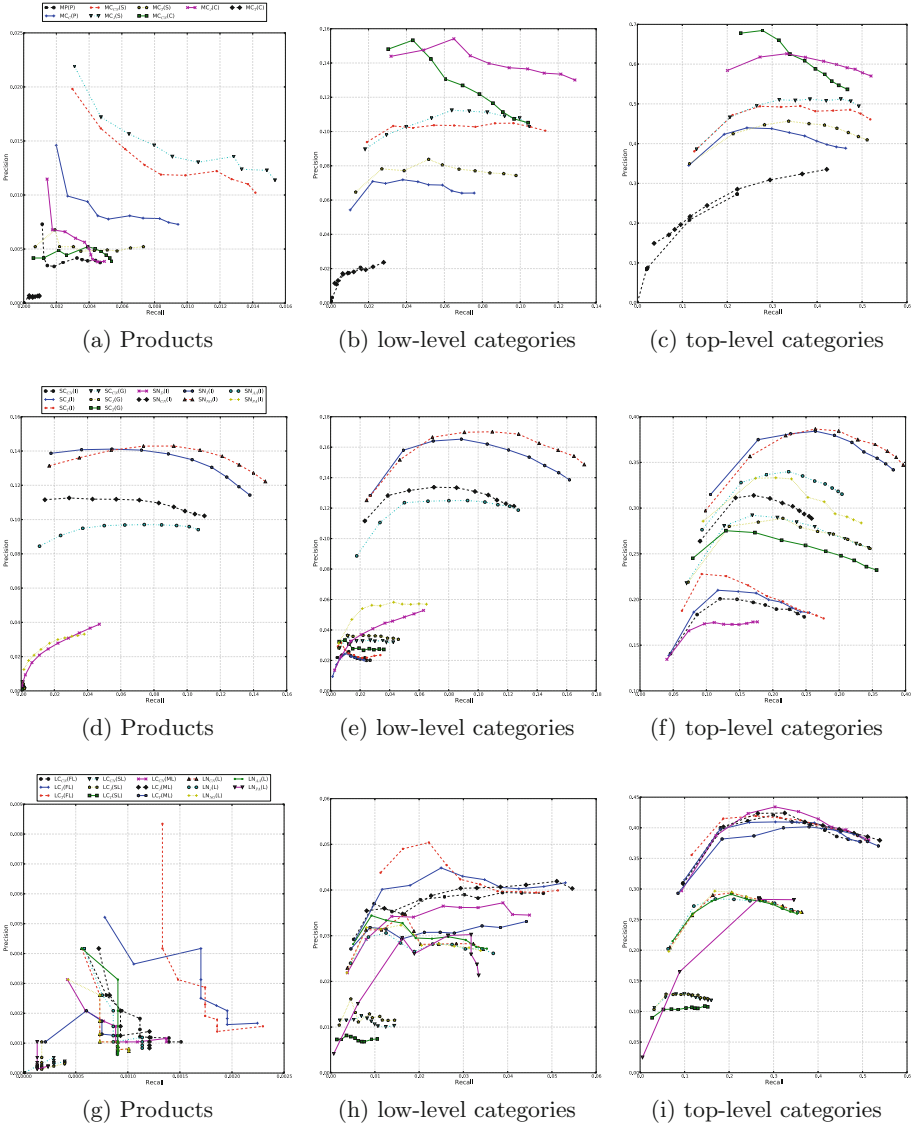
Recommending Products. Regarding the task of predicting product purchases (first column in Table 2), the best results in terms of recommender accuracy are reached by the network-based features based on interactions (e.g., loves, comments, wallposts) between the users in the social network. Surprisingly these approaches clearly outperform the user-based CF approaches relying on marketplace data, which implies that social interactions of the users are a better predictor to recommend products to people than marketplace data.

Another interesting finding is that the neighborhood-based features (*Common Neighbors*, *Jaccard*, *Neighborhood Overlap* and *Adamic/Adar*) also seem to be better indicators to determine the similarity between users than the direct interactions between these pairs. Only the *Preferential Attachment Score* based recommender approach does not perform well in this context, although it still performs better than the features derived from the marketplace data source. This is to some extent expected and reveals that the individual's taste is more driven by the user's peers rather than by the popular users in the SecondLife social network.

In terms of the marketplace and location-based user similarity features, the results reveal that they do not provide high estimates of accuracy. This is

Table 2. Results for the user-based CF approaches based on various user similarity features showing their performance for the tasks of predicting products, low-level categories and top-level categories, respectively (*RQ1*). *Note:* Bold numbers indicate the highest accuracy values per feature set and "*" indicate the overall highest accuracy estimates.

Set	User Sim. Feature	Products			low-level categories			top-level categories			D@10 UC
		nDCG@10	P@10	R@10	nDCG@10	P@10	R@10	nDCG@10	P@10	R@10	
	Most Popular	.0048	.0037	.0047	.0185	.0207	.0157	.2380	.2730	.2221	.6392 100.0%
Market / Content	Common Purchases	.0097	.0073	.0094	.0724	.0641	.0757	.4557	.3884	.4636	.5892 90.51%
	Common Sellers	.0146	.0102	.0142	.1119	.1005	.1132	.5251	.4610	.5183	.6372 99.06%
	Jaccard Sellers	**.0158**	**.0114**	**.0154**	.1092	.1029	.1047	.5061	.4940	.4927	.6054 99.06%
	Total Sellers	.0065	.0052	.0073	.0929	.0743	.0977	.5079	.4094	.5113	.6566 99.06%
	Common Categories	.0050	.0039	.0054	.1090	.1051	.1041	.5073	.5366	.4674	.6123 99.48%
	Jaccard Categories	.0058	.0039	.0049	**.1361**	**.1301**	**.1288**	**.5456***	**.5701***	**.5200***	.6364 99.48%
	Total Categories	.0007	.0006	.0009	.0225	.0236	.0280	.3317	.3353	.4215	**.6575** 99.48%
Social / Content	Common Groups	.0022	.0010	.0014	.0402	.0320	.0425	.3233	**.2567**	.3439	.4307 64.13%
	Jaccard Groups	**.0027**	**.0016**	**.0021**	**.0433**	**.0339**	**.0459**	**.3272**	.2557	**.3464**	.4332 64.13%
	Total Groups	.0006	.0005	.0007	.0324	.0272	.0361	.3214	.2323	.3563	**.4466** 64.13%
	Common Interests	.0005	.0002	.0002	.0235	.0201	.0267	.2285	.1810	.2474	.3185 46.51%
	Jaccard Interests	.0003	.0002	.0003	.0219	.0201	.0227	.2424	.1857	.2551	.3161 46.51%
	Total Interests	.0004	.0002	.0003	.0285	.0235	.0337	.2639	.1793	.2765	.3319 46.51%
Social / Network	Directed Interactions	.0345	.0389	.0471	.0582	.0528	.0630	.1743	.1756	.1769	.2169 38.79%
	Common Neighbors	.1107	.1021	.1104	.1216	.1212	.1243	.2633	.2887	.2584	.3300 62.88%
	Jaccard Common Neighbors	.1381	.1143	.1378	.1523	.1386	.1618	.3726	.3419	.3814	.4455 71.53%
	Neighborhood Overlap	**.1434***	**.1222***	**.1471***	**.1620***	**.1486***	**.1719***	**.3855**	**.3471**	**.3965**	**.4514** 71.53%
	Adamic/Adar	.1013	.0941	.1067	.1241	.1187	.1272	.3028	.3153	.3042	.3762 69.34%
	Pref. Attach. Score	.0317	.0331	.0380	.0630	.0569	.0650	.3202	.2838	.3332	.4420 70.70%
Location / Content	Common Favored Locations	.0019	.0010	.0015	.0427	.0393	.0481	.4674	.3773	.4946	.6437 96.35%
	Jaccard Favored Locations	.0028	**.0017**	.0022	.0472	**.0416**	.0531	.4636	.3777	.4919	.6490 96.35%
	Total Favored Locations	**.0031**	.0016	**.0023**	.0459	.0400	.0513	.4794	.3802	.5061	.6635 96.35%
	Common Shared Locations	.0003	.0003	.0004	.0130	.0103	.0144	.1449	.1180	.1599	.2067 30.45%
	Jaccard Shared Locations	.0005	.0003	.0004	.0134	.0115	.0145	.1420	.1208	.1522	.2042 30.45%
	Total Shared Locations	.0000	.0000	.0000	.0092	.0074	.0106	.1340	.1076	.1520	.2031 30.45%
	Common Monitored Locations	.0016	.0010	.0014	.0408	.0345	.0449	.4825	**.3804**	.5139	.6734 98.23%
	Jaccard Monitored Locations	.0017	.0008	.0012	**.0473**	.0403	**.0546**	**.4987**	.3795	**.5387**	**.6760** 98.23%
	Total Monitored Locations	.0011	.0006	.0009	.0366	.0331	.0442	.4770	.3703	.5354	.6757 98.23%
Location / Network	Common Neighbors	.0015	.0007	.0010	.0298	.0271	.0345	.3377	.2623	**.3632**	.4609 67.05%
	Jaccard Common Neighbors	**.0016**	**.0008**	**.0011**	**.0322**	.0261	**.0367**	.3306	.2623	.3545	.4579 67.05%
	Neighborhood Overlap	.0014	.0007	.0010	.0295	.0267	.0339	.3359	.2614	.3608	**.4615** 67.05%
	Adamic/Adar	.0015	.0006	.0009	.0320	**.0272**	.0353	.3332	.2598	.3530	.4595 67.05%
	Pref. Attach. Score	.0003	.0002	.0002	.0270	.0213	.0335	**.3634**	**.2825**	.3458	.4583 70.59%

(a) Products   (b) low-level categories   (c) top-level categories

(d) Products   (e) low-level categories   (f) top-level categories

(g) Products   (h) low-level categories   (i) top-level categories

Fig. 2. Recall/Precision plots for the single user similarity features derived from the marketplace (a, b, c), social (d, e, f) and location-based (g, h, i) data sources, showing the performance of each feature for $k = 1$–10 recommended items, low-level categories or top-level categories, respectively (*RQ1*). *Note:* Each feature name in the legends is derived in the following way: the first two letters describe the data source, the subscript denotes the user similarity feature and the value in brackets defines the used data field (e.g., $SN_{NO}(I)$ stands for the Social Network data source, the Neighborhood Overlap similarity feature and Interactions data field).

interesting since our previous work [10,22] showed that these features perform extremely well in predicting tie strength or social interactions between users. However, the features derived from the marketplace and the location-based data sources provide the best results with respect to Diversity (D) and User Coverage (UC).

Recommending Categories. Regarding the tasks of predicting low-level and top-level categories, the second and third column of Table 2 report the accuracy estimates for the different user similarity features based on the extracted categories. As expected, all user similarity features end up with a much higher accuracy than for predicting products, especially in the case of top-level categories, because of the lower level of specialization of these recommendation tasks. In the case of the low-level category predictions, the approaches based on social interaction features still perform better than the approaches based on features of the marketplace or location-based data sources. Interestingly, the content-based user similarity features derived from the social network as well as the location-based features, which performed the worst at product predictions, perform much better for low-level categories, now also outperforming the *MP* baseline.

In the case of the top-level category recommendations, it can be seen that the user similarity features of all three data sources provide quite similar results in terms of recommender accuracy. The approach based on the Jaccard's coefficient for categories performs best in terms of nDCG@10, P@10 and R@10. This result is very interesting since this feature is based on the marketplace data source that provided quite bad results in the case of product predictions. Summed up, we see that user similarity features derived from all data sources are very useful indicators for recommendations, although they depend on the level of specialization of the recommendation task (*RQ1*).

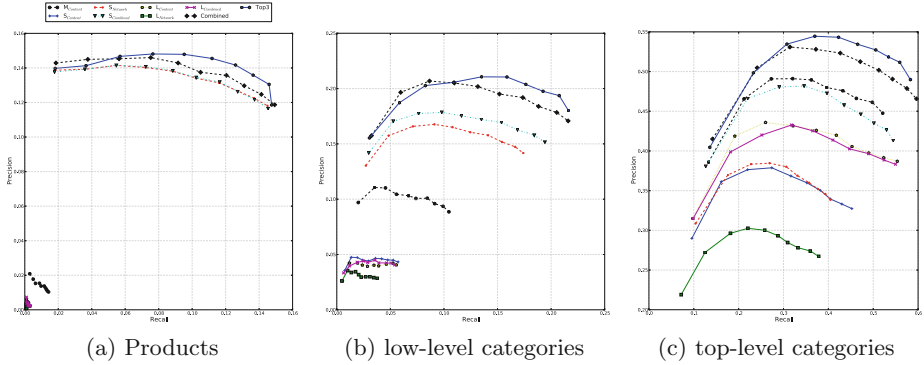## 6.2   Recommendations Based on Combined Data Sources

The findings of the last subsection suggests that a combination of features from all three data sources (marketplace, social network and location-based data) should provide more robust recommendations in case of both tasks, predicting products and categories (*RQ2*). Thus, Table 3 shows the results of the hybrid approaches based on theses data sources in order to tackle our second research question. As before, the first column indicates the results for the product prediction, the second column for the low-level category prediction and the third column for the top-level category prediction. Additionally, Fig. 3 shows the performance of the hybrid approaches based on the three data sources for $k = 1$–10 recommended products, low-level categories or top-level categories, respectively, in form of Recall/Precision plots.

Recommending Products. Regarding the product prediction task, we see again that the recommender approaches based on the social network data source clearly outperform the ones based on the marketplace and location-based data sources as well as the *MP* baseline. Furthermore, when combining all three data sources, not only the overall recommendation accuracy is increased with respect

Table 3. Results for the hybrid approaches based on our three data sources for the tasks of predicting products, low-level categories and top-level categories. The results show that all three data sources (marketplace, social network and location-based data) are important indicators for calculating recommendations, since a hybrid combination of all data sources provided the best results in case of predicting products, low-level categories and top-level categories (*RQ2*). *Note:* Bold numbers indicate the highest accuracy values across the data sources and "*" indicate the overall highest accuracy estimates.

	Sets	Products			low-level categories			top-level Categories				
		nDCG@10	P@10	R@10	nDCG@10	P@10	R@10	nDCG@10	P@10	R@10	D@10	UC
	Most Popular	.0082	.0021	.0122	.0185	.0207	.0157	.2380	.2730	.2221	.5945	100.00%
Market	Content	.0151	.0103	.0142	.1051	.0886	.1046	.5232	**.4476**	.5210	.6421	99.79%
Social	Content	.0030	.0019	.0025	.0702	.0432	.0571	.5555	.3274	.4523	.5603	81.65%
	Network	.1416	**.1180**	.1450	**.2454***	.1416	.1740	.5708	.3385	.4047	.4591	71.53%
	Combined	**.1418**	.1165	**.1454**	.2104	**.1515**	**.1942**	**.5901**	.4131	.5446	.6245	92.70%
Location	Content	.0036	.0021	.0030	.0556	.0406	.0556	.5535	.3870	**.5535**	.6923	100.00%
	Network	.0015	.0007	.0010	.0535	.0284	.0378	.5359	.2672	.3783	.4864	70.59%
	Combined	.0036	.0022	.0031	.0540	.0406	.0540	.5497	.3832	.5497	.6914	100.00%
Weighted Sum	Combined	**.1460***	**.1187***	**.1493***	.2163	.1708	.2159	**.5978***	.4656	**.5965***	.6642	100.00%
	Combined Top 3	.1459	.1186	.1475	.2161	**.1801***	**.2161***	.5829	**.4898***	.5829	.6530	100.00%

(a) Products   (b) low-level categories   (c) top-level categories

Fig. 3. Recall/Precision plots for the hybrid approaches showing the performance of each data source for $k = 1$–10 recommended products, low-level categories or top-level categories, respectively (*RQ2*).

to nDCG@10, P@10 and R@10, but also the User Coverage (UC) is increased to the maximum of 100 %.

This means that the hybrid approach combines the strengths of the user similarity features of all three data sources in order to be capable of providing accurate recommendations for all users in the datasets. Another hybrid approach shown in Table 3 combines only the best user similarity features from each data source (referred to as *Top 3*) and reaches higher accuracy estimates, but lower Diversity (D).

Recommending Categories. In contrast to the results of the product predictions, that showed that the recommender based on the social network data source clearly outperforms the recommenders based on the marketplace and location-based data sources, the results of the category predictions (second and third column of Table 3) do not show that big differences between the three data sources. With respect to the low-level category predictions, we again observe that the recommender based on the social network data source still provides the highest accuracy estimates.

Interestingly, in this case also the recommenders based on the other two data sources provide reasonable results, which has not been the case of predicting products, where the location-based recommender even was outperformed by the *MP* baseline. Based on these results we would assume that marketplace and location-based data sources are suitable of providing accurate predictions in more general recommendation tasks. The results for the top-level category predictions prove this assumption since in this case the recommenders based on marketplace and location-based data sources even provide better results in terms of recommender accuracy, Diversity and User Coverage than the one based on social network data in case of the content-based features. As before, the combination of all three data sources provide again the best results. Summed up, this results prove our assumption derived from *RQ1*, that all three data sources

are important for calculating recommendations, since a combination of all data sources provided the best results in case of predicting products, low-level categories and top-level categories (*RQ2*).

7   Conclusions and Future Work

In this work we presented first results of a recently started project that tries to utilize various user similarity features derived from three data sources (marketplace, social network and location-based data) to recommend products and points of interests (i.e., low-level and top-level categories) to people in an online marketplace setting. This section concludes the paper with respect to our two research questions and gives an outlook into the future.

The first research question of this work (*RQ1*) dealt with the question as to which extent user similarity features derived from marketplace, social network and location-based data sources can be utilized for the recommendation of products and categories in online marketplaces. To tackle this question we implemented various user-based Collaborative Filtering (CF) approaches based on the user similarity features from the data sources and tested them isolated. As the results have shown, the user-based CF approaches that utilize features of online social network data to calculate the similarities between users performed best in case of predicting products, significantly outperforming the other approaches relying on both – marketplace and location-based user data. However, this behavior changed in the case of predicting low-level and top-level categories where the differences between the three data sources got substantially smaller. Surprisingly, with respect to the top-level category predictions, the marketplace and location-based features even reached the highest results in terms of accuracy, Diversity (D) and User Coverage (UC).

These results showed that user similarity features of all three data sources are important indicators for recommendations and suggests that combining them should result into more robust recommendations, especially in cases of multiple recommendation tasks on different levels of specialization (topics and categories). Thus, our second research question (*RQ2*) tried to tackle the question if the different marketplace, social network and location-based user similarity features and data sources can be combined in order to create a hybrid recommender that provides more robust recommendations in terms of prediction accuracy, diversity and user coverage. In order to address this question we implemented and evaluated hybrid recommenders that combined the features of the data sources. The results proved our assumption and showed that hybrid recommender that combined user similarity features of all three data sources provided the best results across all accuracy metrics (nDCG@10, P@10, R@10) and all settings (product, low-level category and top-level category recommendations). Moreover, this hybrid recommender also provided a User Coverage of 100 % and thus, is able to provide these accurate recommendation to all users in the datasets.

Although the results of this study are based on a dataset obtained from the virtual world SecondLife, we believe that it bears great potential to create a

sequence of interesting studies that may have implications for the "real" world (see e.g., [23]). For instance, one of the potential interesting issues we are currently exploring is predicting products and categories to users in a cold-start setting (i.e., for users that only have purchases a few or even no products in the past) by a diversity of features. Other important work we plan is the use of state-of-the-art model-based approaches in order to assess whether the signals extracted from similarity features in the current analysis can be replicated (in the case of social data) or improved (in the case of location data) for different recommendation tasks.

We have also shown that using the interaction information between users improves not only the task of product recommendation, but also the recommendation of low-level and top-level categories. Thus, we are also interested in studying the extent to which recommendations can be improved by utilizing content-based similarity features derived from the users' social streams.

References

1. Zhang, Y., Pennacchiotti, M.: Predicting purchase behaviors from social media. In: Proceedings of WWW '13, pp. 1521–1532 (2013)
2. Guo, S., Wang, M., Leskovec, J.: The role of social networks in online shopping: Information passing, price of trust, and consumer choice. In: Proceedings of EC '11, pp. 157–166. ACM (2011)
3. Trattner, C., Parra, D., Eberhard, L., Wen, X.: Who will trade with whom? Predicting buyer-seller interactions in online trading platforms through social networks. In: Proceedings of WWW '14, pp. 387–388. ACM (2014)
4. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I.: Recommender systems with social regularization. In: Proceedings of WSDM '11, pp. 287–296. ACM (2011)
5. Jamali, M., Ester, M.: A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of RecSys '10, pp. 135–142. ACM, New York (2010)
6. Bischoff, K.: We love rock'n'roll: analyzing and predicting friendship links in Last.fm. In: Proceedings of WebSci '12, pp. 47–56. ACM (2012)
7. Feng, W., Wang, J.: Incorporating heterogeneous information for personalized tag recommendation in social tagging systems. In: Proceedings of KDD '12, pp. 1276–1284. ACM (2012)

8. Delporte, J., Karatzoglou, A., Matuszczyk, T., Canu, S.: Socially enabled prefer-ence learning from implicit feedback data. In: Blockeel, H., Kersting, K., Nijssen, S., Železný, F. (eds.) ECML PKDD 2013, Part II. LNCS, vol. 8189, pp. 145–160. Springer, Heidelberg (2013)
9. Lacic, E., Kowald, D., Parra, D., Kahr, M., Trattner, C.: Towards a scalable social recommender engine for online marketplaces: The case of apache solr. In: Proceed-ings of WWW '14, pp. 817–822. ACM (2014)
10. Steurer, M., Trattner, C.: Acquaintance or partner? Predicting partnership in online and location-based social networks. In: Proceedings of ASONAM'13. IEEE/ACM (2013)
11. Adamic, L., Adar, E.: Friends and neighbors on the web. Soci. Netw. **25**, 211–230 (2003)
12. Cranshaw, J., Toch, E., Hong, J., Kittur, A., Sadeh, N.: Bridging the gap between physical location and online social networks. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, pp. 119–128. ACM (2010)
13. Barabási, A., Albert, R.: Emergence of scaling in random networks. Science **286**, 509–512 (1999)
14. Lacic, E., Kowald, D., Trattner, C.: Socrecm: A scalable social recommender engine for online marketplaces. In: Proceedings of HT '14, pp. 308–310 (2014)
15. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recom-mender systems. In: The adaptive web. Springer (2007) 291–324.
16. Bostandjiev, S., O'Donovan, J., Höllerer, T.: Tasteweights: a visual interactive hybrid recommender system. In: Proceedings of RecSys '12, pp. 35–42. ACM (2012)
17. Smyth, B., McClave, P.: Similarity vs. Diversity. In: Aha, D.W., Watson, I. (eds.) ICCBR 2001. LNCS (LNAI), vol. 2080, pp. 347–361. Springer, Heidelberg (2001)
18. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. (TOIS) **22**, 5–53 (2004)
19. Van Rijsbergen, C.J.: Foundation of evaluation. J. Doc. **30**, 365–373 (1974)
20. Parra, D., Sahebi, S.: Recommender systems: sources of knowledge and evaluation metrics. In: Velásquez, J.D., Palade, V., Jain, L.C. (eds.) Advanced Techniques in Web Intelligence-2. SCI, vol. 452, pp. 149–176. Springer, Heidelberg (2013)
21. Ge, M., Delgado-Battenfeld, C., Jannach, D.: Beyond accuracy: evaluating recom-mender systems by coverage and serendipity. In: Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 257–260. ACM (2010)
22. Steurer, M., Trattner, C.: Predicting interactions in online social networks: an experiment in second life. In: Proceedings of the 4th International Workshop on Modeling Social Media, p. 5. ACM (2013)
23. Szell, M., Sinatra, R., Petri, G., Thurner, S., Latora, V.: Understanding mobility in a social petri dish. Scientific Reports 2 (2012)