

A Package Recommendation Framework for Trip Planning Activities

Idir Benouaret
Sorbonne Universités
Université de technologie de Compiègne
CNRS UMR 7253 Heudiasyc
CS 60 319 - 60 203 Compiègne cedex
idir.benouaret@hds.utc.fr

Dominique Lenne
Sorbonne Universités
Université de technologie de Compiègne
CNRS UMR 7253 Heudiasyc
CS 60 319 - 60 203 Compiègne cedex
dominique.lenne@hds.utc.fr

ABSTRACT

Classical recommender systems provide users with ranked lists of recommendations, where each one consists of a single item. However, these ranked lists are not suitable for applications such as trip planning, which deal with heterogeneous items. In this paper, we focus on the problem of recommending a set of packages to the user, where each package is constituted with a set of different Points of Interest that may constitute a tour. Given a collection of POIs, our goal is to recommend the most interesting packages for the user, where each package satisfies the budget constraints. We formally define the problem and we present a novel composite recommendation system, inspired from composite retrieval. Experimental evaluation of our proposed system, using a real-world dataset demonstrates its quality and its ability to improve both diversity and relevance of recommendations.

Keywords

recommender system, package, top-k, diversity, trip planning

1. INTRODUCTION

Classical recommender systems provide the user with recommendations as ranked lists consisting in single items, e.g., movie, book. In trip planning, a user is interested in suggestions for points of interest (POI), that could be very heterogeneous, e.g., museum, park, restaurant, etc. A tourism recommender system can benefit from a system capable of recommending items organized in packages (bundles) rather than ranked lists, which constitute an improved exploratory experience for the visitor. Then, there is a need to recommend for the user the best packages that match his preferences. Furthermore, there may be a cost for visiting each point of interest (time, price, etc.) that the user may want to constraint with a budget. The budget can also simply be the number of items per package. Some so-called "third generation" travel planning web sites, such as Tripadvisor and YourTour, aim at assisting the user with suggestions of

places to visits integrating these kinds of constraints, but the suggestions are often based only on the most popular places and neglect the personalization aspect for the user. Thus, the use of these web sites is very limited. Our contribution in this paper is the design and implementation of a suggestion model promoting diversity and inspired from composite retrieval [1]. The approach we propose is to group suggestions in different packages, where each package is constituted with a set of diverse POIs. Each POI has a cost and a time of visit associated with it, and the user specifies a maximum total budget for any recommended package of items. POIs in each package are chosen using a scoring function that takes into account the personalization, the diversity and the popularity of the items in the package. The evaluation of our proposed system using a real dataset with data crawled from the website Tripadvisor, shows that our system is competitive, it can improve the diversity of recommendations without deteriorating their relevance. The road map of the paper is as follows. After discussing related work in more details (Section 2), we give a formalization of the problem (Section 3). We then describe the model and define the quality criteria of packages. We show how we calculate the score of a package (section 4). In section 5, we describe our algorithm for calculating the top-k packages. In section 6, we subject our system to thorough experimental analysis using a real dataset. We investigate the quality of the recommended packages. Finally, we discuss future work and conclude the paper in Section 7.

2. RELATED WORK

In [2], authors are interested in finding the top-k tuples of entities. Examples of entities could be cities, hotels and airlines, while packages are tuples, they query documents using keywords in order to determine entity scores. A package in their framework has a fixed size, e.g., one city, one hotel and one airline. Instead, in our work, we allow packages of variable size, subject constraints specified by the user. Other closely related work is [4] where a framework is proposed to recommend travel itineraries from online user-generated data, like picture uploads. They formulate the problem of recommending travel itineraries that might be interesting for users while the travel time is under a given time budget. However, in this work, the value (score) of each POI is only determined by the number of times it was mentioned by other users in the social network, whereas in our work, the importance of a POI is determined not only by the popularity of the POI but also with a personalized score depending on user's preferences and his ratings for other POIs. Finally,

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

RecSys '16, September 15-19, 2016, Boston, MA, USA

© 2016 ACM. ISBN 978-1-4503-4035-9/16/09...\$15.00

DOI: <http://dx.doi.org/10.1145/2959100.2959183>

the closest work to our's is [6], where the authors explore returning approximate solutions to composite recommendations. The focus of the work is on using a Fagin-style algorithm for variable size packages and proving its optimality. The same authors further develop the idea into a prototype of recommender system for travel planning (CompRec) [7]. However, the score of an item is just the predicted rating of a user, while we believe that using also the popularity of items improves the relevance of the recommended packages. Furthermore, none of these works accounts for the diversity in packages which leads to a better satisfaction of the user.

3. PROBLEM STATEMENT

Given a set I of POIs, a set U of users, an active user $u \in U$ and a POI $i \in I$. We denote by $c(i)$ the cost of POI i and by $t(i)$ the average time needed for visiting POI i . Given a set of POIs $P \subset I$, we define $Score(P)$ the score of a package P , which estimates the quality of a package (see more details in section 4.2.4), $c(P) = \sum_{i \in P} c(i)$ the cost for a package P and $t(P) = \sum_{i \in P} t(i)$ the time to visit POIs in package P . Given a cost budget B_c and a time budget B_t , a package P is said *Valid* iff $c(P) \leq B_c$ and $t(P) \leq B_t$.

PROBLEM 1. *Top-k Composite recommendations*

Given a set I of POIs, an active user u with his preferences background, a cost budget B_c , a time budget B_t and an integer k , a top-k composite recommender system has to determine the top-k packages P_1, P_2, \dots, P_k such that each P_i has $c(P_i) \leq B_c$, $t(P_i) \leq B_t$, and among all Valid packages, P_1, P_2, \dots, P_k have the k highest scores, i.e. $Score(P) \leq Score(P_i)$ for all Valid packages $P \notin \{P_1, P_2, \dots, P_k\}$

4. MODEL

4.1 Distance and Similarity between POIs

Our distance between POIs is based on a taxonomy of hierarchical topic categories organized in a tree structure. Formally, we used a domain ontology developed by [3] to represent these categories. Let be I the set of all possible POIs. Each POI in I is associated to one category in the taxonomy, e.g. museum, park, building, etc. We define the topical distance $dist_t : I \times I \rightarrow \mathbb{N}$ between two POIs i and j as the length of the shortest path between the two categories of i and j in the taxonomy:

$$dist_t(i, j) = sp(c_i, c_j) \quad (1)$$

where c_i, c_j are the categories of POIs i, j and sp is the shortest path function. The topical similarity evaluates in which measure two POIs i and j deal with similar thematics or topics. The similarity depends on the topical distance between the two POIs i, j :

$$sim(i, j) = \frac{1}{1 + dist_t(i, j)} \quad (2)$$

4.2 Packages Quality Criteria

In order to suggest the top-k packages of POIs for a specific user, it is necessary to have a criteria estimating how "good" is a package P according to a user u . We denote it by the score of a package. To this end, we need to measure how a new POI not visited yet by the user would be interesting for him. The popularity is also an important factor in the general appreciation of a POI. In fact, the properties depending only on the popularity of a POI are often more important than the similarity with the user profile [5].

Moreover, we assume that the user is not only interested by visiting POIs he would like, but instead visiting POIs that best cover his interests. For example, a user who likes museums will not be satisfied with a recommended package with only museums. The diversity of POIs in the same package is thus an important criteria for the quality of the package.

4.2.1 Overall Popularity

The overall popularity measures the popularity of a POI $i \in I$:

$$opop(i) = \frac{pop(i)}{\max_{j \in I} pop(j)} \in [0, 1] \quad (3)$$

where j designates the POIs of I and $pop : I \rightarrow \mathbb{N}$ represents a popularity indicator of a POI. By extension, the overall popularity of a package P is:

$$opop(P) = \frac{\sum_{i \in P} opop(i)}{|P|} \in [0, 1] \quad (4)$$

4.2.2 Intra Package Diversity

Most of travel recommender systems focus on the modeling of user preferences in order to get a ranking of the most pertinent POIs. However, the diversity of suggestions has never been the focus point. Nevertheless, it has been suggested that the diversity has a large positive effect on the satisfaction of the user [8]. So, we adapt the intra list diversity introduced by [8] for a package of POIs P , we define the intra package diversity:

$$ipd(P) = \frac{\sum_{i, j \in P} 1 - sim(i, j)}{|P|^2} \quad (5)$$

4.2.3 Estimated Appreciation (Prediction)

The estimated appreciation evaluates to what extent a POI i that a user has not yet rated would be interesting. The estimated appreciation of a user $u \in U$ for a POI $i \in I$ is calculated using a simple item-item collaborative filtering:

$$eapp_u(i) = \frac{\sum_{j \in S_i} rating_u(j) \times sim(i, j)}{\sum_{j \in S_i} sim(i, j)} \quad (6)$$

where j designates POIs of the sample S_i , the set of similar POIs rated by the user u , and $rating_u \rightarrow [0, 1]$ associates for a POI the rating given by the user u , divided by the maximum rate. By extension, the estimated appreciation for a user u for a package P is defined by the mean of the estimated appreciation for POIs forming the package:

$$eapp_u(P) = \frac{\sum_{i \in P} eapp_u(i)}{|P|} \in [0, 1] \quad (7)$$

4.2.4 Score of a package

The score of a package evaluates the quality of POIs that form a package for a user u according to the overall popularity, the diversity and the estimated appreciation. The score for a package P for a user u is calculated by:

$$Score_u(P) = C_{eapp} \times eapp_u(P) + C_{opop} \times opop(P) + C_{div} \times ipd(P) \quad (8)$$

where $C_{eapp}, C_{opop}, C_{div}$ are positive Coefficients that modulate respectively, the importance of the estimated appreciation, the overall popularity and the diversity in the score function.

5. RECOMMENDING TOP-K PACKAGES

The construction of top-k packages is done in two steps: first, a set of valid packages are produced in large quantities

with a cardinality $c \gg k$, packages are formed by aggregation around a pivot POI and taking into account the quality criteria. Our approach for forming a set of good valid packages is inspired from the algorithm "BOBO" (*Bundles One-By-One*) introduced by Amer-Yahia *et al.* [1]. We adopted this algorithm to take into account the quality criteria of the packages defined in Section 4.2. The goal of this algorithm is to create c valid packages that respect the budget constraints. It is inspired from $k - nn$ clustering. At each step a POI is chosen as pivot, and a valid package with maximum score is built around that pivot. The pseudo code is described in *algorithm 1*.

BOBO starts with an empty set of packages. Then, a list of candidates pivots POIs is constituted, by ranking I in the descending order of popularity. As long as the number of formed packages is less than c , at each iteration the first POI is taken from the set of *Pivots*, and a package is built around it. This is done by the routine *Pick_bundle* described in algorithm 2. This routine greedily keeps picking the next POI that maximizes the score of the package formed around the pivot, as far as the cost budget and the time budget constraints are satisfied. Let us go back to BOBO's main loop: once a candidate package is created, it is added to "Packages" and its elements are removed from "Pivots" so that they are not longer used.

Algorithm 1: BOBO

Input: I, B_c, B_t , number of packages c
Output: a set c of packages

```

1 Packages  $\leftarrow \emptyset$ 
2 Pivots  $\leftarrow \text{Descending\_sort}(I, opop)$ 
3 while (Pivots  $\neq \emptyset$ ) and |Packages| <  $c$  do
4    $w \leftarrow \text{Pivots}[0]$ 
5   Pivots  $\leftarrow \text{Pivots} - \{w\}$ 
6    $P \leftarrow \text{Pick\_bundle}(w, I, B_c, B_t)$ 
7   Pivots  $\leftarrow \text{Pivots} - P$ 
8   Packages  $\leftarrow \text{Packages} \cup P$ 
9 end
10 return Packages
```

Algorithm 2: PICK_BUNDLE

Input: pivot w, I, B_c, B_t
Output: a package P

```

1  $S \leftarrow w$ 
2 active  $\leftarrow I - \{w\}$ 
3 cost  $\leftarrow c(w)$ 
4 time  $\leftarrow t(w)$ 
5 while (not finish) do
6    $i \leftarrow \text{argmax}_{i \in \text{active}} \text{Score}_u(S \cup \{i\})$ 
7   if (cost +  $c(i) \leq B_c$ ) and (time +  $t(i) \leq B_t$ ) then
8      $S \leftarrow S \cup \{i\}$ 
9     cost  $\leftarrow \text{cost} + c(i)$ 
10    time  $\leftarrow \text{time} + t(i)$ 
11  end
12  else
13    finish  $\leftarrow \text{true}$ 
14  end
15  active  $\leftarrow \text{active} - i$ 
16 end
17 return S
```

Once the required number of packages has been created, they are ranked following their respective scores (Section 4.2.4). Afterwards, we select the k packages having the best scores.

6. EXPERIMENTS

6.1 Data Set

The goal of our experiments were: (1) evaluate the relevance of the packages recommended by our algorithm, and (2) evaluate their diversity as well. In order to have a set of POIs constituting potential recommendations, we crawled data from Tripadvisor. Each POI has a thematic category organized in a tree structure, which allows us to construct our similarity measure. In addition, Tripadvisor provides for a POI the number of users rating for a POI, we used it as an indicator of its popularity, for estimating the function $pop : I \rightarrow \mathbb{N}$ defined in Section 4.2.1. For our experiments, we crawled users rating information from POIs in the five most popular cities in France. We exclude POIs that have very few or no reviews. The dataset contains 40635 ratings for 1183 POIs by 18227 users, so as we see the data is very sparse. We associate with each POI its cost and its average time of visit crawled also from Tripadvisor. The average price of POIs was close to €7 and the average time for visiting a POI ranges from 30 minutes to 3 hours. Because of the large sparsity of the underlying user rating matrix, we selected the 20 most active users as our sample for testing the algorithms.

Versions	C_{eapp}	C_{opop}	C_{div}
per	1	0	0
pop	0	1	0
div	0	0	1
pop+div	0	1/2	1/2
per+div	1/2	0	1/2
per+pop	1/2	1/2	0
per+pop+div	1/3	1/3	1/3

Table 1: Different versions of our system

6.2 Evaluation Metrics and Experimental Protocol

The following measures are used for evaluation :

Precision: is calculated as the ratio of recommended POIs that are relevant to the total number of recommended POIs.

Diversity: we extend the intralist diversity introduced by Ziegler *et al* [8] to a set of k packages $\{P_1, \dots, P_k\}$. The Mean Intralist Diversity (MILD) is defined.

$$\text{diversity} = \text{MILD}(\{P_1, \dots, P_k\}) = \frac{\sum_{i=1}^k \text{ILD}(P_i)}{k} \quad (9)$$

F_{PD} : the F-measure is the harmonic mean of precision and diversity :

$$F_{PD} = \frac{2 \times \text{precision} \times \text{diversity}}{\text{precision} + \text{diversity}} \quad (10)$$

Our goal was to test the impact of personalization (per), popularity (pop) and diversity (div) on the quality of recommendations. To this end, we compared several versions of our system, corresponding to different possible combinations of the factors of C_{eapp} , C_{opop} , C_{div} . Each version corresponds to a different combination of the parameters. Versions we tested are summarized in the Table 1. The name of each version indicates the use or not of the different aspects when constituting the set of packages.

We tested our system varying the number of returned packages k , we vary it between four values: $\{5, 10, 15, 20\}$.

	k=5			k=10			k=15			k=20		
	P(%)	D(%)	F_{PD} (%)	P(%)	D(%)	F_{PD} (%)	P(%)	D(%)	F_{PD} (%)	P(%)	D(%)	F_{PD} (%)
per	49.73	40.02	44.35	50.19	41.12	45.20	51.13	42.74	46.56	48.88	43.75	46.17
pop	57.75	47.89	52.36	53.5	51.33	52.39	51.84	49.43	50.59	50.18	48.61	49.38
div	41.01	61.96	49.35	43.01	60.39	50.25	42.68	59.21	49.60	38.08	57.81	45.91
pop+div	53.7	55.73	54.69	53.36	58.04	55.60	51.92	56.14	53.94	49.67	55.81	52.56
per+div	47.85	55.03	51.19	49.88	57.84	53.56	48.1	57.18	52.24	48.22	53.46	50.70
per+pop	59.38	43.25	50.04	54.53	50.57	52.47	53.03	50.37	51.66	51.11	49.25	50.16
per+pop+div	55.09	58.45	56.72	53.66	59.88	56.59	51.75	57.33	54.39	50.39	55.84	52.97
Xie et al	57.24	43.58	49.48	53.32	48.11	50.58	51.65	50.8	51.22	50.45	52.85	51.62

Table 2: Comparison of our different versions with the competitive approach

The cost budget is fixed to €60 and the time budget fixed to 600 minutes. We tested our algorithms under various cost and time budgets with very similar results, so other budgets are omitted for lack of space.

6.2.1 Competitive Approach

To evaluate also the effectiveness of the proposed system, we compare our results with the package recommendation method proposed by *Xie et al.* [6], which is the closest work to ours. The authors also computed the estimated appreciation of POIs using an item-item collaborative filtering approach, without taking in consideration the popularity of POIs and the diversity aspect of recommendations.

6.3 Results and discussions

Results of our versions compared to the competitive approach according to precision, diversity and F_{PD} are reported in Table 2. In all our versions, we can notice a high influence of the popularity of POIs with respect to the precision. It is important to underline that the popularity is a significant factor as well as the personalization. In fact, in most cases, the "pop" version leads to a better precision than the "per" version and the "pop +div" version better than the "per+div" version. These results are in accordance with [5], which highlights the importance of the popularity and its effect on the relevance of recommendations. Varying the number of packages, the "per+pop" version always performs the best precision and outperform the algorithm of *Xie et al.*, due to combining the personalization and the popularity. Without surprise, the "div" version is the one who performs the best diversity compared to all others. However, it has also the worst values for precision. Concerning the F-measure between precision and diversity, we notice that the "per+pop+div" realizes the best compromise, and outperforms the competitive algorithm. This version tends to promote a large diversity, performs better than *Xie et al.*, and is not significantly different in precision comparing to the "per + pop" version. Thus, the "per+pop+div" is the best approach when considering both precision and diversity.

7. CONCLUSIONS AND FUTURE WORK

Motivated by applications of trip planning, we studied the problem of recommending packages consisting of sets of POIs. Our composite recommendation system consists of ranking packages according to a score function, where the score of a package depends on the estimated appreciation, the overall popularity and the diversity of POIs constituting the package. We formalized the problem of generating top-k packages recommendations that are under budgets,

where a cost and a time is incurred by visiting each recommended POI and the budgets are user specified. We developed an algorithm for retrieving the top-k packages with best scores. The evaluation of our system using a real world dataset crawled from the website Tripadvisor demonstrates its quality and its ability to improve both the relevance and the diversity of recommendations. We plan now to realize a study of the proposed system with real users on a situation of mobility, where the localization context will take an important role on the recommendation process. Furthermore, it will be interesting to compare the preferences of the user between a recommender system providing classical ranked lists and our composite recommender system. A mobile application is currently being developed, so that we can test it later with real users and study the exploratory dimension achieved by the composite approach and its impact on users satisfaction.

8. REFERENCES

- [1] S. Amer-Yahia, F. Bonchi, C. Castillo, E. Feuerstein, I. Mendez-Diaz, and P. Zabala. Composite retrieval of diverse and complementary bundles. *Transactions on Knowledge and Data Engineering*, 2014.
- [2] A. Angel, S. Chaudhuri, G. Das, and N. Koudas. Ranking objects based on relationships and fixed associations. *EDBT*, 2009.
- [3] L. Castillo, E. Armengol, E. Onaindía, L. Sebastián, J. González-Boticario, A. Rodríguez, S. Fernández, J. D. Arias, and D. Borrajo. Samap: An user-oriented adaptive system for planning tourist visits. *Expert Systems with Applications*, 2008.
- [4] M. De Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu. Automatic construction of travel itineraries using social breadcrumbs. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, 2010.
- [5] H. Steck. Item popularity and recommendation accuracy. *RecSys*, 2011.
- [6] M. Xie, L. V. Lakshmanan, and P. T. Wood. Breaking out of the box of recommendations: From items to packages. *RecSys*, 2010.
- [7] M. Xie, L. V. Lakshmanan, and P. T. Wood. Comprec-trip: A composite recommendation system for travel planning. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, 2011.
- [8] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. In *World Wide Web*, 2005.