Relevance Models for Multi-Contextual Appropriateness in Point-of-Interest Recommendation

Anirban Chakraborty
ADAPT Centre, School of Computer Science & Statistics
Trinity College Dublin, Ireland
anirban.chakraborty@adaptcentre.ie

Debasis Ganguly
IBM Research Europe
Dublin, Ireland
debasis.ganguly1@ie.ibm.com

Owen Conlan
ADAPT Centre, School of Computer Science & Statistics
Trinity College Dublin, Ireland
Owen.Conlan@scss.tcd.ie

ABSTRACT

Trip-qualifiers, such as *trip-type* (vacation, work etc.), *accompanied-by* (e.g., solo, friends, family etc.) are potentially useful sources of information that could be used to improve the effectiveness of POI recommendation in a current context (with a given set of these constraints). Using such information is not straight forward because a user's text reviews about the POIs visited in the past do not explicitly contain such annotations (e.g., a positive review about a pub visit does not contain the information on whether the user was with friends or alone, on a business trip or vacation). We propose to use a small set of manually compiled knowledge resource to predict the associations between the review texts in a user profile and the likely trip contexts. We demonstrate that incorporating this information within an IR-based relevance modeling framework significantly improves POI recommendation.

CCS CONCEPTS

• **Information systems → Personalization**; *Information retrieval diversity*; *Recommender systems*; **Probabilistic retrieval models**.

KEYWORDS

Joint Context Modeling, Point-of-Interest Recommendation, Weak Supervision

ACM Reference Format:
Anirban Chakraborty, Debasis Ganguly, and Owen Conlan. 2020. Relevance Models for Multi-Contextual Appropriateness in Point-of-Interest Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401197

1 INTRODUCTION

The definition of contextual (Point-of-Interest) recommendation largely relies on a precise definition of the *context* itself. In our work, we consider that there are two broad sources of context information that a contextual recommendation system can be benefited from. The first of these describes the *present state* of a user at an instant of time, which is typically a combination of features with categorical values, e.g., the location of the user, purpose of the trip (e.g. leisure vs. work), season of the trip (e.g., summer, fall, winter or spring) etc. The location context of a user is an example of a *hard constraint* (recommendations outside the city of a user's current location are of no use), whereas the other non-location type constraints (trip-qualifiers) are examples of *soft constraints* (a relevant POI for a solo trip may also be partially relevant for a trip with friends and so on). The second source of information is the *past state* of a user, which acquired over a sufficient period of time, is likely to broadly capture her general preferences in particular situations. In other words, the past information provides information about the user's general preferences for certain types of items over others, e.g. 'museums' over 'beaches' in particular situations, e.g. when travelling 'solo' (*accompanied-by* feature) for 'leisure' (*trip-type* feature).

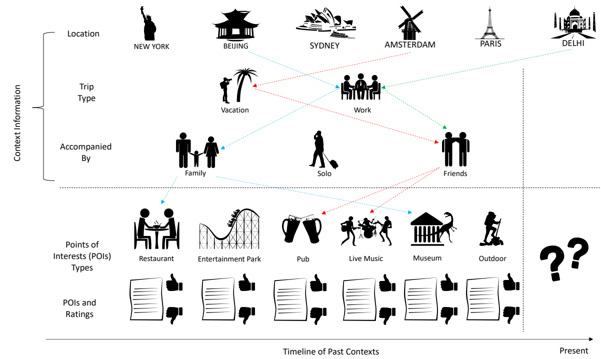

Figure 1: Schematics of Contextual Recommendation

One major challenge to learn a POI's contextual appropriateness is the inevitable absence of explicit annotation of non-location type context (e.g. trip qualifiers, such as 'trip-type' etc.) in the user preference history. To illustrate this point, consider typical user feedback in a location based social networks (LBSNs), such as Foursquare[1] or TripAdvisor[2]. This usually comprises a review text and an explicit rating score in some categorical range (representing the scale from very bad to very good). An important point to note here is that this past information does not contain the trip qualifier information, i.e. the context in which the POI was visited and rated thereafter. Despite the usual lack of availability of non-location or *trip qualifier* information in the preference histories of users, such information may form a part of the present state of the user (i.e. the query). An important research question is then to bridge the

[1] https://foursquare.com
[2] https://tripadvisor.com

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07...$15.00
https://doi.org/10.1145/3397271.3401197

gap between the lack of contextual information from the historical information of user feedback and the constraints imposed by them during the present context.

The top part of Figure 1 shows two types of context information of a user, first, the *location* of the user (specifically, a city which the user is currently visiting), and second, the more personal trip-qualifiers (non-location type) information categories which further qualify the location context, e.g. the 'trip-type' (whether vacation or work), 'accompanied-by' (i.e. whether accompanied by family or a solo trip) etc. The vertical line in Figure 1 separates the past context of a user from his present, e.g. the figure shows that the user's current location is Delhi, and that he has visited New York, Beijing etc. in the past. The bottom-left part of Figure 1, constituting a part of a user's history, shows a list of POIs that the user rated positively (or negatively) during her different trips. A path rooted at one of the location nodes and terminating at a particular POI denotes a single trip of a user among her past trips, e.g., in Figure 1, the path shown by the red coloured arrows starting from the node 'Amsterdam' and visiting in sequence the nodes 'Vacation', 'Friends', 'Pub' and 'Live Music' denotes a set of POIs which the user visited (and rated) during her leisure trip to Amsterdam with her friends. Although the complete trip information is shown in the schematic diagram of Figure 1, it is worth noting that the tree is essentially incomplete in real-life situation, i.e. the non-location type contextual information is not present in user ratings. A key research challenge is then to estimate a likely path in the tree from a location to a number of POIs, i.e., *estimate the likely non-location intermediate nodes by utilizing the information from the review text* themselves. After constructing a model of a user's preferences, the challenge in contextual recommendation is to make new recommendations to the user for a new present location (that she has not visited before) with a given set of trip qualifiers, e.g., the path specified in Figure 1 with the green arrows indicates that the user's current location is 'Delhi' which she is visiting for work along with her colleagues. An effective POI recommendation system in this scenario should leverage similar situations in the past (in this example, the user's past non-solo work trips in other locations) in estimating what types of POIs the user had previously rated positively in similar situations, and then use information from these past POIs to recommend a set of similar POIs for the current location.

2 MULTIPLE CONTEXTS IN POI RETRIEVAL

We propose an IR-centric approach to bridge the information gap between a user's past context and the POI descriptions in her current location. While previous IR approaches have addressed the semantic gap between the user review text (queries) and the POI descriptions (documents) [1, 3, 4], these approaches do not make provision for considering the non-location type qualifiers as a part of the queries and eventually considering them in estimating relevance scores of documents. In this section, we describe our proposed pseudo-relevance feedback based IR approach for addressing these additional features.

Definition of Documents and Queries. A user profile is comprised of a set of reviews. Each review comprises a body of text about a POI, a set of tag terms added to it and a score (see the bottom part of Figure 1). It is to be noted that a review in a user profile does not have information about the trip qualifiers, as indicated by the dotted arrows from the upper part of Figure 1 into each review. The current context (query), however, comprises a pair of trip qualifiers of the form (L, Q), which includes a location of the trip, L, and other non-location type qualifiers, $Q = Q_1 \times \ldots Q_c$ (total of c trip qualifier types, each type denoted as Q_i). Specifically, for the TREC-CS [4] dataset used in our experiments, the number of such non-location qualifiers is 3 (i.e. $c = 3$). Each non-location type context q_U is hence a 3-dimensional categorical vector comprised of the attributes Q_1=`trip-type`, e.g. holiday, Q_2=`trip-duration`, e.g. day-trip, and Q_3=`accompanied-by`, e.g. alone or friends etc.

Location (*Hard*) Constraint. In an IR-based approach to contextual recommendation, we estimate a similarity function that for a user U takes as input a *query*, constituting instances of

(1) the profile of the user comprising tuples of review text, D, tags, T, and scores, r, i.e., $(D, T, r) \in U$ (for notational convenience, we denote the set of all review documents for the user U as D_U and the set of all tags as T_U, and use the notation P_U to denote $D_U \cup T_U$, i.e., the text found in a user's past history), and
(2) an instance of the user's current context specified by the location and trip qualifiers $(l_U, q_U) \in (L, Q)$.

The objective then is to *rank* a set of POIs (*hard* constrained by $L = l_U$) in decreasing order of their estimated relevance scores within the current context. A way to estimate the relevance scores is to first restrict the set of candidate POIs to only the ones in the specific location (by employing the *hard* constraint), i.e. $S(l_U) = \cup\{d : L(d) = l_U\}$ (L denoting the location attribute of a POI). The next step then makes use of the text in the user profile, P_U, and this candidate set of POI descriptors $S(l_U)$ to estimate the relevance scores,

$$\phi : (P_U, l_U, S(l_U) = \cup\{d : L(d) = l_U\}) \mapsto \mathbb{R}, \tag{1}$$

where the output of the function, ϕ (e.g. with BM25 or a pseudo-relevance feedback method), does not depend on the non-location type qualifiers $q_U \in Q$.

Weak Supervision for Trip Qualifier (*Soft*) Constraints. To incorporate non-location type qualifiers, one needs to learn an association between a word from the review text or the tag vocabulary of a user profile and the likely (historical) context (trip-type, duration etc.) leading to the review. As an example, it should be possible for humans (with their *existing knowledge*) to infer that a review about a pub frequently mentioning phrases, such as 'friends', 'good times', 'tequila shots' etc., most likely corresponds to a visit with friends on vacation (i.e. `trip-type=holiday` and `accompanied-by=friends`). A computational approach towards automatically constructing this association may leverage a knowledge base (e.g. a seed set of term-category associations). One such knowledge resource was compiled in [2], which is composed of the following two different types of manually assessed information.

(1) List of pairs constituting a term and a *single* non-location trip-qualifier with manually judged relevance scores of the form (t, q, a), where t is a term (e.g. food), q is a single category (e.g. holiday) and $a \in [0, 1]$ is a manually judged appropriateness score. An example of a non-relevant pair is (`nightlife, business, 0.1`) with a lower score.

(2) List of pairs of a term with a *joint* context (a 3-dimensional vector of categories) along with a manually assessed binary label (1/0) indicating whether the term is relevant in the given *joint context* or not. As an example, the word 'pub' is assessed to be non-relevant in the joint context of '(`holiday, weekend-trip, family`)', whereas it is relevant in the context '(`holiday, weekend-trip, friends`)'.

We formally denote these two knowledge resources as

$$\begin{aligned} \kappa_s &: (w, q) \mapsto [0, 1], w \in V, q \in Q_i, i \in \{1, \ldots, c\} \\ \kappa_j &: (w, q) \mapsto \{0, 1\}, w \in V, q \in Q = Q_1 \times \ldots Q_c, \end{aligned} \tag{2}$$

where Q denotes the set of *joint* non-location type contexts (*soft* constraints), Q_i denotes a *single* context category, and V denotes the vocabulary set of the review text and tags.

A seed set of such labeled examples of term-context (single or joint) association pairs can then be used to define a modified similarity score function ψ. In contrast to the text-based function of Equation 1, this also takes into account the information from the *soft* constraints of the query context. In particular for a given *soft* constraint vector q_U in the user query, we use embedded word vector representations to aggregate the similarities of each word in the review text/tag of a user profile with the seed words assessed as relevant for a single or a joint context q_U. Formally, $\forall w \in P_U$ we define two functions of the form $\psi : (w, q_U) \mapsto \mathbb{R}$, one each for addressing the single and the joint contexts, as

$$\begin{aligned} \psi_s(w, q_U) &= \max(\mathbf{w} \cdot \mathbf{s}),\ s \in \cup\{t : \kappa_s(t, q_U) > 0\} \\ \psi_j(w, q_U) &= \max(\mathbf{w} \cdot \mathbf{s}),\ s \in \cup\{t : \kappa_j(t, q_U) = 1\}. \end{aligned} \tag{3}$$

For each word w (embedded vector denoted as $\mathbf{w}$) in the profile of a user, we compute its maximum similarity over i) all seed words in the case of single context (ψ_s), or ii) with respect to a subset of seed words relevant only for the given context, i.e., the words for which $\kappa(q_U, s) = 1$ in the case of the joint context (ψ_j). Specifically, we use `word2vec` [7] to embed the vector representation of a word.

The reason for using the maximum as the aggregate function in Equation 3 is that a word is usually semantically similar to a small number of seed words relevant to a given context. For example, for the query context 'holiday, day-trip, friends', the relevant seed set constitutes words such as 'base-ball stadium', 'beer-garden', 'salon', 'sporting-goods-shop', etc. However, a word such as 'pub' is similar to only one member of this seed set, namely 'beer-garden', which means that other aggregation functions, such as averaging, can lead to a low aggregated value, which is not desirable in this case.

A Factored Relevance Model with *Soft* Constraints. We combine both the text-based similarity ϕ (Equation 1), and the trip context driven similarity function ψ (ψ_s or ψ_j of Equation 3) into the standard framework of relevance model [6]. With the notations introduced earlier, the extended relevance model is given by

$$P(w|\theta_{P_U, q_U}) = \sum_{d \in D_U} rP(w|d)\psi(w, q_U) \prod_{t \in T_U} P(t|d), \tag{4}$$

where $P(w|d)$ denotes the normalized term frequency of a word w in document d. In addition to addressing the semantic relationship between a user described tag and a term presented in the POI description, the extended relevance model of Equation 4 also takes into account the trip-qualifier based contextual appropriateness of a term w by the use of the $\psi(w, q_U)$ factor. A higher value of this factor indicates that either w is itself one of the seed words in an existing knowledge base or its embedded vector is close to one of the seed words thus indicating its likely contextual appropriateness.

Table 1: *Soft* constraint categories with their values.

Categories	Values
Q_1: `trip-type`	{`business`, `holiday`, `other`}
Q_2: `trip-duration`	{`day-trip`, `longer`, `night-out`, `weekend-trip`}
Q_3: `accompanied-by`	{`alone`, `family`, `friends`, `other`}

To impose the *hard* constraint of the location qualifier l_U, we estimate another relevance model θ_{P_U, q_U, l_U}, by making use of the relevance model estimated only with the *soft* constraints (Equation 4) and the selected subset of location-specific POIs (documents) by following the exposition of [3]. Formally,

$$P(w|\theta_{P_U, q_U, l_U}) = \sum_{d: L(d) = l_U} P(w|d)\psi(w, q_U) \prod_{t \in \theta_{P_U, q_U}} P(t|d). \tag{5}$$

Finally, we linearly combine the two relevance models of Equations 4 (*soft* constraint only) and 5 (both *hard* and *soft* constraints), i.e.,

$$P(w|\theta_U) = \gamma_H P(w|\theta_{P_U, q_U}) + (1 - \gamma_H) P(w|\theta_{P_U, q_U, l_U}), \tag{6}$$

where γ_H is the parameter to control the relative importance of the two relevance models.

3 EVALUATION

Experiment Setup. Our experiments are conducted with the TREC Contextual Suggestion (TREC-CS) 2016 Phase-1 setup [4]. The task requires a system to return a ranked list of 50 POIs from a given query collection, that best fit the user preference history and the user's current context, comprised of $c = 3$ different non-location qualifiers outlined in Table 1. The overall collection comprises over 1.2M of POIs in total, and the number of context queries used in our experiments is 61 (part of the TREC-CS 2016 dataset).

Methods Investigated. The objective of our experiments is to investigate what is the most effective way to include the *soft* constraints of a given query to improve retrieval effectiveness. Our proposed method includes this information as a part of a factored relevance model (Equation 6) with the help of the ψ function, which is either of ψ_s or ψ_j respectively for single and joint contexts. To enable fair comparisons of standard baselines with the method of Equation 6, we extend standard baseline approaches with the *soft* constraints as well, which we describe next.

The choice of the *soft-constraint* similarity function $\psi = \{\psi_l, \psi_s, \psi_j\}$ yields three different versions for each method investigated, corresponding to i) not using the *soft* constraints, or using ii) the single-context or the iii) joint-context based similarities, respectively. This is denoted by an additional parameter for the function ψ in the results (Table 2). The baseline function corresponding to only location (*hard*) constraint corresponds to the constant function $\psi_l : (w, q) \mapsto \{1\}$. In **BM25**, for each query term t, we include the value of $\psi(t, q_U)$ as the weight of that term in the query. For the baseline approach **RLM**, we use traditional 'RM3', including $\psi(t, q_U)$ as the weight of each expansion term t. In **Pop-K**, we use the K most frequently assigned tags (extracted from the profile of a user) to construct a weighted query, the weights of each term being identical to the BM25 baseline. In **NeuMF**, we employ a recent neural matrix factorization methodology [5], which uses a combination of generalized matrix factorization (GMF) and multi-layer feed-forward networks to model user-item (in our case an item corresponds to a tag) interactions. Similar to Pop-K, the K most likely tags (as predicted by the NeuMF model) are then used to construct a weighted query (using the ψ function as its weight similar to the BM25 baseline). Previous research [1] investigated the use of separately computing a similarity score between the query words/tags and document (POI) words/tags (Content + Tag, or **C+T** score) with a predicted likelihood score of the relevance between a query word and a given non-location (*soft*) constraint category. This study trained an SVM-based binary classifier on the joint-context knowledge resource [2] (with relevance labels 0/1) using as inputs the scores for the single contexts [1]. While testing (query), the distance of a 3-dimensional joint context input (obtained from the query) from the classifier boundary is added to the text (tag-word) matched score (higher the distance higher is the likelihood of a tag to be appropriate to the given joint context). We employ this approach as a baseline and denote it with C+T+SVM. Additionally, we also investigate the method of adding the scores obtained from the ψ_s and ψ_j functions with the C+T approach.

The baseline **FRLM**, i.e., the location-only approach of [3], is a degenerate case (setting $\psi = \psi_l$) of our proposed approach. Our proposed approach corresponds to the two cases where we take into account the *soft* constraints by setting $\psi = \psi_s$ and ψ_j.

For our experiments, the parameters of each method (e.g., k and b in BM25, K in Pop-K etc.) were independently optimized by grid search with respect to the metric nDCG@5. The optimal value of the linear combination parameter $\gamma_H \in [0, 1]$ for the FRLM approaches (for the three different cases of ψ) was found to be 0.8.

Results. From Table 2, we first observe that including the trip-qualifier based information in the form of joint context (ψ_j) mostly improves the POI retrieval effectiveness, e.g. improvements are observed for RLM, NeuMF etc. (compare the results between ψ_j and ψ_l for each method). Standard approaches do not benefit much from the inclusion of the trip-qualifiers in the form of single-context driven scores, a plausible reason for which can be attributed to the fact that relevant single-context matches may not lead to the conjunctive relevance for the joint context. However, including even the single context based similarity scores as part of the query term weights in standard IR and RS (recommender system) approaches tend to improve the recall (effectiveness measures such as MAP and nDCG mostly improve at the cost of a decrease in nDCG@5 or P@5). Second, it can be seen that using the *soft* constraint scores as a part of a model is usually more effective than a simple post-hoc combination of these scores with content matching scores (e.g. the relative improvements in FRLM as compared to that of Pop-K or C+T). Third, in contrast to a parametric approach, such as SVM, the proposed similarity function ψ_j (Equation 3) works better. This is because supervised approaches typically require to rely on large quantities of training data to work well. Moreover, the SVM based approach of [1] did not take into account the semantic similarities between words to estimate the trip-qualifier based appropriateness. It is observed that computing similarities with the embedded word vectors turns out to be more effective. Finally, it can be observed that the best results are obtained when the joint-context based similarity function is incorporated into the FRLM model.

Table 2: Three variants for each method - w/o trip qualifiers (ψ_l), with single (ψ_s) and joint contexts (ψ_j). Significance of the best method (ψ_j) with respect to the most effective baseline is denoted by '*' (t-test with $p = 0.05$).

			Evaluation Metrics			
	Method	ψ	nDCG@5	nDCG	P@5	MAP
Baselines	BM25	ψ_l	0.2747	0.2889	0.3934	0.1326
	BM25	ψ_s	0.2609	0.2889	0.3869	0.1335
	BM25	ψ_j	0.2641	0.2916	0.3639	0.1355
	RLM	ψ_l	0.2615	0.3091	0.3574	0.1437
	RLM	ψ_s	0.2583	0.3107	0.3475	0.1443
	RLM	ψ_j	0.2692	0.3189	0.3639	0.1496
	Pop-K	ψ_l	0.2488	0.2811	0.3410	0.1280
	Pop-K	ψ_s	0.2529	0.2861	0.3639	0.1321
	Pop-K	ψ_j	0.2568	0.2908	0.3574	0.1362
	NeuMF	ψ_l	0.1626	0.2480	0.2361	0.0937
	NeuMF	ψ_s	0.1491	0.2466	0.2131	0.0935
	NeuMF	ψ_j	0.1698	0.2457	0.2393	0.0923
	C+T	ψ_l	0.2499	0.2800	0.3967	0.1330
	C+T	ψ_s	0.2623	0.2841	0.4066	0.1383
	C+T	ψ_j	0.2688	0.2979	0.4000	0.1484
	C+T	SVM	0.2656	0.2833	0.3770	0.1330
	FRLM	ψ_l	0.2919	0.3418	0.3934	0.1616
Proposed	FRLM	ψ_s	0.2956	0.3435	0.4033	0.1637
	FRLM	ψ_j	**0.3075***	**0.3498***	**0.4098**	**0.1687***

Conclusions. We proposed a word embedding based approach to compute the similarity of a given trip qualifier (part of the query) with a POI description by employing weak supervision from a knowledge resource. We observed that modeling the trip-qualifier contexts jointly turns out to be the most effective in comparison to not using these qualifiers at all, or modeling them independently.

ACKNOWLEDGMENTS

This work was supported by the ADAPT Centre for Digital Content Technology, funded under the Science Foundation Ireland Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

REFERENCES

[1] Mohammad Aliannejadi and Fabio Crestani. 2018. Personalized Context-Aware Point of Interest Recommendation. *TOIS* 36, 4 (2018), 45:1–45:28.
[2] Mohammad Aliannejadi, Ida Mele, and Fabio Crestani. 2017. A Cross-Platform Collection for Contextual Suggestion. In *SIGIR 2017*. 1269–1272.
[3] Anirban Chakraborty, Debasis Ganguly, Annalina Caputo, and Séamus Lawless. 2019. A Factored Relevance Model for Contextual Point-of-Interest Recommendation. In *ICTIR 2019*. ACM, New York, NY, USA, 157–164.
[4] Seyyed Hadi Hashemi, Charles LA Clarke, Jaap Kamps, Julia Kiseleva, and Ellen M Voorhees. 2016. Overview of the TREC 2016 contextual suggestion track. In *Proceedings of TREC*, Vol. 2016.
[5] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *WWW 2017*. 173–182.
[6] Victor Lavrenko and W. Bruce Croft. 2001. Relevance Based Language Models. In *ACM SIGIR 2001 (SIGIR '01)*. ACM, New York, NY, USA, 120–127.
[7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *NIPS'13*. 3111–3119.