# From ITDL to Place2Vec – Reasoning About Place Type Similarity and Relatedness by Learning Embeddings From Augmented Spatial Contexts


Bo Yan
STKO Lab
University of California, Santa Barbara
boyan@geog.ucsb.edu

Krzysztof Janowicz
STKO Lab
University of California, Santa Barbara
janowicz@ucsb.edu

Gengchen Mai
STKO Lab
University of California, Santa Barbara
gengchen_mai@geog.ucsb.edu

Song Gao
Department of Geography
University of Wisconsin, Madison
song.gao@wisc.edu



## ABSTRACT

Understanding, representing, and reasoning about Points Of Interest (POI) types such as `Auto Repair`, `Body Shop`, `Gas Stations`, or `Planetarium`, is a key aspect of geographic information retrieval, recommender systems, geographic knowledge graphs, as well as studying urban spaces in general, e.g., for extracting functional or vague cognitive regions from user-generated content. One prerequisite to these tasks is the ability to capture the similarity and relatedness between POI types. Intuitively, a spatial search that returns body shops or even gas stations in the absence of auto repair places is still likely to satisfy some user needs while returning planetariums will not. Place hierarchies are frequently used for query expansion, but most of the existing hierarchies are relatively shallow and structured from a single perspective, thereby putting POI types that may be closely related regarding some characteristics far apart from another. This leads to the question of how to learn POI type representations from data. Models such as Word2Vec that produces word embeddings from linguistic contexts are a novel and promising approach as they come with an intuitive notion of similarity. However, the structure of geographic space, e.g., the interactions between POI types, differs substantially from linguistics. In this work, we present a novel method to augment the spatial contexts of POI types using a distance-binned, information-theoretic approach to generate embeddings. We demonstrate that our work outperforms Word2Vec and other models using three different evaluation tasks and strongly correlates with human assessments of POI type similarity. We published the resulting embeddings for 570 place types as well as a collection of human similarity assessments online for others to use.


Final draft

## CCS CONCEPTS

• **Information systems** → **Information retrieval**; **Similarity measures**; • **Computing methodologies** → *Machine learning*;


## KEYWORDS

Points of Interest, Similarity, Geo-Semantics, Machine Learning

**ACM Reference format:**
Bo Yan, Krzysztof Janowicz, Gengchen Mai, and Song Gao. 2017. From ITDL to Place2Vec – Reasoning About Place Type Similarity and Relatedness by Learning Embeddings From Augmented Spatial Contexts. In *Proceedings of SIGSPATIAL'17, Los Angeles Area, CA, USA, November 7–10, 2017,* 10 pages.
https://doi.org/10.1145/3139958.3140054


## 1 INTRODUCTION AND MOTIVATION

Semantic similarity and relatedness measures are prominent components of a variety of methods in geographic information retrieval, recommender systems, ontology engineering, and so forth; see [10] for a recent overview.[1] Given the importance of categorization for human cognition [8], place types are one of the three components (location and name being the other two) published by all major gazetteers and POI databases. [2] Place types act as a proxy for functions that a particular place of a given type affords. Intuitively, the presence of a nightclub (irrespective of its name or location) implies a certain exposure to noise during nights, the presence of a younger demographic, singles, a higher potential for drug related crimes, the possibility of getting a drink or snack late at night, and so forth. While each nightclub may differ to some degree, nightclubs share many of their characteristics with bars and the broader category of music venues, while they can neither act as substitute for bakeries nor barbers. Consequently, in the absence of POIs of a certain type, e.g., `Nightclub`, within a search radius, a system should return a

[1]Similarity and relatedness are related concepts, in fact similarity is a subproperty of relatedness but not the other way around. To give an intuitive example, the Griffith Observatory is *related* to Griffith Jenkins Griffith via a `donorOf` relation but the observatory and the person are not *similar*. Many techniques, especially those based on linguistic aspects (including Word2Vec [19]) instead of formal semantics, cannot effectively distinguish between similarity and relatedness. Consequently, we approach them here together. Two of our three evaluation schemata, however, will explicitly focus on (human) assessments of similarity.

[2]In the following, we will use Point of Interest (POI) and place as synonyms.

place of a similar type, e.g., `Bar`. This implies that semantic similarity measures should reflect human assessments of similarity, be it about place types or another topic.

To measure similarity, one may syntactically compare type labels, compute the distance in a place type hierarchy, count common place in their extensions, and so forth. New methods rely on comparing their linguistic meaning by learning word embeddings for all types and then computing their Cosine Similarity. However, such approaches do not consider any spatial information that is implicitly embedded in these place types, such as their co-occurrence patterns. This idea resembles the distributional semantics in linguistics and can be further summarized as: *place can be categorized by their neighbors*. The original counterpart in the linguistics is: *You shall know a word by the company it keeps* [5].

In this work, we embrace the idea of distributional semantics in geographic space and explore the similarity and relatedness of place types using different latent representations with augmented spatial contexts. Spatial contexts are augmented both intrinsically and extrinsically. In order to consider distance in our approach, distance decay and distance lags are used as intrinsic adjustments to augment the spatial contexts. We realize that there is a notable difference between place and space, namely *place is space infused with human meaning* [26], so we take check-in counts, i.e., popularity, as a proxy for human activities into consideration as well. Finally, and to adjust for the fact that place types follow a power law distribution, we also take the uniqueness of types at a certain distance into account. We approach both aspects from an information theoretic perspective, i.e., by measuring *information content*.

**The contributions of this paper are as follows:**

- We illustrate that the commonly used linguistic models alone cannot adequately capture the structure of geographic space such as the distinctive patterns in which places of different types co-occur. Instead, we propose a novel model based on *augmented spatial contexts* that make geographic distance a first-class citizen and adjust these contexts by an information theoretic perspective on the uniqueness of place types within a certain distance as well as their popularity as a proxy for human activities.
- We provide a comprehensive evaluation of different place type embeddings with respect to the top-down Yelp POI category hierarchy. This evaluation essentially brings inductive (bottom-up place type embeddings) and deductive (top-down place hierarchy structure) approaches together.
- We establish two baselines using Amazon's Mechanical Turk Human Intelligence Tasks (HIT) for measuring the similarity and relatedness of place types. Our evaluation result shows that our method has better accuracy than purely linguistically based embeddings, which confirms the importance of explicit spatial contexts. In fact, we demonstrate the remarkable fact that similarity assessments derived from embeddings created exclusively via our augmented spatial contexts, i.e., by merely studying spatial patterns of place types and their relative popularity, correlate strongly with human similarity judgments despite the fact that humans can rely on their rich cultural experience, the meaning of type labels, their background knowledge, and so forth.
- While the resulting place type embeddings can be used for a wide range of tasks that rely on similarity assessments such as commonly used in geographic information retrieval, co-reference resolution and ontology-alignment, as well as recommender system, we introduce a novel perspective, namely *compression*, as an interesting future area of study that deals with the question of whether place types can be substituted or act as proxies for other POI types, e.g., to summarize neighborhoods by a minimal number of place types.
- Finally, we make the embeddings as well as thousands of human similarity assessments from Mechanical Turk available online at http://stko.geog.ucsb.edu/place2vec for future use.

The remainder of this paper is organized as follows. Section 2 summarizes existing work on embeddings and geospatial semantics. Section 3 presents the dataset and provides basic concepts used throughout our work. Section 4 explains in detail how we model the augmented spatial contexts. Section 5 presents three evaluation schemes and Section 6 is evaluation. Finally, Section 7 summarizes the research and points to future directions.

## 2 RELATED WORK

Most research on POI embeddings originates from word embedding techniques using neural network language models [2]. One of the most successful models in this class is Word2Vec, which is composed of Skip-Gram and Continuous-Bag-of-Words, proposed by Mikolov et al. [19, 20]. It uses neural networks that take advantage of the distributional semantics of natural languages. Skip-Gram learns the embeddings by predicting context words given center words whereas Continuous-Bag-of-Words does it the other way around.

Previous works on embeddings related to geographic information can be grouped into two categories. The first category considers the influence of geographic context on word embeddings. In a first attempt to investigate the extent to which geographic context affects the semantics of words, Cocos and Callison-Burch [3] trained word embeddings in geolocated tweets using geographic contexts derived from Google Places and OpenStreetMap (OSM). Their work is similar to ours in a sense that they also realize the importance of geospatial contexts, but the scope of their work remains limited to the linguistic domain. In addition, their result shows that geographic context is not as semantically rich as textual context. In contrast, we will demonstrate that *augmented spatial contexts* are indeed rich in semantic information. Zhang et al. [31] also acknowledges the variation in the semantics of words depending on the geographic space. They propose a vector space transformation under different topic distributions in order to generate a mapping between different geographic contexts. Yet again their approach is focusing on linguistic aspects whereas geographic aspects are not directly considered in their model.

The second category is more similar to our work which models geographic entities directly. Yao et al. [28] and Zhang et al. [30] have a very different focus compared to our study as they utilize embedding techniques in order to detect the spatial distribution of urban land use and uncover urban dynamics. We are focusing on exploring the extent to which different adjustment to the spatial context influences the embedding results. Feng et al. [4] and Zhao et al. [32] learn embedding in order to predict future POI visits or recommend

POIs. This is a byproduct of the original prediction-based Word2Vec models. Our work has a different focus and therefore does not require temporally sequential data, such as check-in sequences of users. Instead, we are interested in the *semantics* of place types and utilize embeddings as a means to construct representations, share them, and to measure (semantic) similarity across types, e.g., in the context of query expansion [10] and extraction [12].

This relates our work to research on geographic information retrieval and geospatial semantics, and here more specifically to the social sensing framework of *semantic signatures* [9] which characterizes place types based on thematic, temporal, and spatial perspectives called *bands* in analogy to spectral signatures. For example, thematic bands for Points Of Interest have been studied by Adams and Janowicz [1] using Latent Dirichlet Allocation to extract topics from unstructured texts about place types. Quercini and Samet [23] proposes a set of graph-based similarity measures to determine the relatedness of a concept to a location in the Wikipedia link structure. These location-related concepts, which are referred to as *local lexicon* in their work, can be seen as signatures to differentiate geographic entities as well. Research on the temporal perspective has also shown promising results. Ye et al. [29] studied the temporal dimensions of places in the context of location-based social networks. McKenzie and Janowicz [17] applied temporal signature to reverse geocoding to adjust rankings returned by a spatial range search based on a temporal distortion model. So far, the spatial perspective, i.e., the question whether one can learn place (type) representations exclusively from spatial patterns, has received less attention. Mülligann et al. [22] used a measure based on combining point pattern analysis with semantic similarity, while Zhu et al. [33] proposes 27 spatial statistical features to characterize different aspects of place types in digital gazetteers. Our work can be seen as a continuation of this line of research and a contribution to the semantic signatures framework by using novel methods such as augmented spatial contexts to overcome the limitations of previous work. In fact, we will show that these contexts (even when taken on their own) are able to reproduce human similarity judgments, i.e., yield strong correlations between human assessments and our model.

# 3 PRELIMINARIES

The individual Points of Interest and their categories used in this research are from the Yelp Dataset Challenge[3]. This dataset covers venues from 11 different cities from four countries (United Kingdom, Germany, Canada, and the United States). We selected Las Vegas as study region, but our methods can be generalized to different cities and place type schema; see [18] for a discussion about regional effects. The Yelp dataset groups their 1030 POI types into 22 root categories, such as `Restaurants`, `Shopping`, `Arts & Entertainment`, `Professional Services`, `Health & Medical`, and so forth. Each POI $l_i$ in the POI set $L$ is composed of three parts, a POI name $n \in N$, a geographic identifier (here, latitude and longitude of a place location modeled as centroid) $g \in G$, and a set of associated POI types $\{t_1, t_2, t_3, ..., t_k\} \subseteq T$.

After analyzing the 1030 place types and their frequencies in Las Vegas, we see a long tail in the rank-frequency distribution (Figure 1). The log-log plot also shows a linear trend. Fitting $log(frequency)$

[3] https://www.yelp.com/dataset_challenge

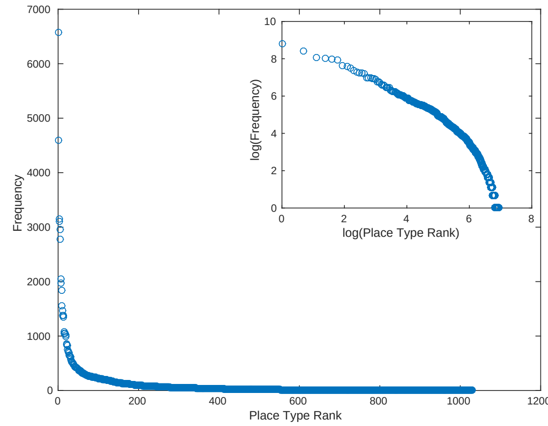


**Figure 1: POI type rank-frequency and log-log plot.**

and $log(rank)$ using linear regression, yields a value of 0.8543 for R-squared which indicates that the model fits strongly to the data and a p-value of 2.2e−16 which indicates that such a scaling effect is highly significant. Simply put, these statistics show that the rank-frequency indeed follows a power law distribution by which a few POI types dominate the data. This is an important motivation for the proposed information content-based frequency adjustment in our augmented spatial contexts discussed in the following section.

# 4 METHODS

In this section, we describe the latent representation method and the augmented spatial contexts. The latent representation originates from natural language processing and has been used successfully in many domains. By acknowledging the difference in context formation between geographic space and linguistic expressions, we introduce three approaches to model the geographic influence in determining latent representations. These methods include, naive spatial context, simple augmented spatial context, and Information Theoretic, Distance Lagged (ITDL) augmented spatial context.

## 4.1 Latent Representation Method

Recent work has shown that the latent representation model Word2Vec can effectively capture the semantic relationships in word spaces based on the distributional semantics assumption [19, 20]. From analyzing the POI type distribution, we know that, similarly to the word frequency distribution [14], it follows a power law distribution. This leads us to taking advantage of the Word2Vec model and its underlying distributional semantics assumption for the study of POI types in geographic space.

We selected the Skip-Gram model, which predicts context POI types given *center* types. Our objective is to approximate the true place type probability distribution from our training data. A typical approach is to use cross entropy to measure the difference between the learned probability and the true probability. Since our data is discrete and we only care about the center place type, the cross entropy can be simplified as:

$$D(\hat{y}, y) = -y_c log(\hat{y}_c) \quad (1)$$

where $\hat{y}$ and $y$ are the learned probability distribution and true probability distribution, respectively. $\hat{y}_c$ is the predicted probability of the context POI types given the center place type (denoted by the index $c$), and $y_c$ is the true probability of the context POI types given the center place type. $\hat{y}_c$ can be further defined as:

$$\hat{y}_c = P(t_1, t_2, t_3, ..., t_m | t_c) \tag{2}$$

where $t_1, t_2, t_3, ..., t_m$ are the context place types and $t_c$ is the center place type. In order to calculate the probability, we apply the Naive Bayes assumption. Note that $y_c$ will always be 1. Finally, we use the softmax function to turn the scores into probabilities and substitute the POI types with vector representations. The objective function is defined as:

$$\text{minimize } J = -log \prod_{t=1}^{m} \frac{exp(u_t^T v_c)}{\sum_{k=1}^{|T|} exp(u_k^T v_c)} \tag{3}$$

where $u_t$ and $v_c$ are the context place type vectors and center place type vectors, respectively; $|T|$ is the cardinality of a POI type, i.e., its *extension*. We implement the model in TensorFlow using Mini-Batch Gradient Descent and Noise-Contrastive Estimation [21].

## 4.2 Naive Spatial Context

An intuitive approach to utilize the structure of geographic space is to naively model the spatial context based on the center place type and context place type co-occurrences. We denote the context place type as $t_{context}$ and center place type as $t_{center}$. This naive method is faithful to the original Word2Vec model and captures the spatial contextual information using a nearest neighbor approach. Unlike natural languages which are sequential in nature, Points of Interest in Yelp are distributed in a 2D geographic space. As a result, instead of using a fixed-size sliding window to construct $(t_{center}, t_{context})$ pairs, we create spatial buffers around each center POI to detect the k-nearest neighbor POIs and record their respective place types as our training pairs. Since each center POI $l_i$ and each context POI $l_j$ can have a set of place types $T_{l_i}$ and $T_{l_j}$ respectively, we use the Cartesian product $T_{l_i} \times T_{l_j} = \{(t_{center}, t_{context}) | t_{center} \in T_{l_i} \wedge t_{context} \in T_{l_j}\}$ to obtain the training pairs for each center POI and candidate context POI. We append these training pairs to the final list of training data $SC_{naive}$ [4] as we iterate through all center and context POIs.

## 4.3 Simple Augmented Spatial Context

Within the naive spatial context the geographic component, namely the distance, is merely used as a criteria to search the neighborhoods and not modeled directly. In this second approach, we augment the naive spatial context by incorporating distance decay and/or aggregated check-in counts (as proxy for the relative popularity or dominance). The rationale behind this approach is that we acknowledge both distance and human activity as essential components in modeling the latent representations of POI types, and, hence, want to study how they can contribute to the final result by modeling them both individually and in combination. Here we define popularity $P_{l_i}$ of a POI $l_i$ as the number of total check-ins associated with $l_i$. By augmenting the spatial context, we increase the number of times a $(t_{center}, t_{context})$ tuple appears in our training dataset with a factor of $\beta$, where $\beta \in \{n | n \in \mathbb{Z}, n \geqslant 1\}$.

[4]We use $SC$ as an abbreviation for Spatial Context and use different subscripts to denote different types of Spatial Contexts.

For incorporating activity alone, the factor $\beta$ is defined as:

$$\beta_{checkin}^{lj} = \lceil 1 + ln(1 + P_{lj}) \rceil \tag{4}$$

where $\beta_{checkin}^{lj}$ is the augmenting factor for the training tuple $(t_{center}, t_{context})$ when the context POI is $l_j$. This is an extrinsic augmentation approach.

For incorporating distance decay alone, we define the augmenting factor as:

$$\beta_{distance}^{lj} = \left\lceil \frac{1 + \frac{\sum_{k=1}^{|L|} P_{lk}}{|L|}}{1 + d^{\alpha}(l_i, l_j)} \right\rceil \tag{5}$$

where $|L|$ is the total number of POIs, $d(l_i, l_j)$ is the distance between center POI $l_i$ and context POI $l_j$, and $\alpha$ is an inverse distance factor, set to 1 in our case. The numerator is a smoothing constant for a given POI dataset. This is an intrinsic augmentation approach.

For combining both distance decay and human activities in the spatial context, the augmenting factor, which combines both intrinsic and extrinsic approaches, is defined as:

$$\beta_{combined}^{lj} = \left\lceil \frac{1 + ln(1 + P_{lj})}{1 + d^{\alpha}(l_i, l_j)} \right\rceil \tag{6}$$

As one can see, the proposed augmenting factors are based on the check-ins of the context POI as well as the distance from the center POI to the context POI, thus incorporating more geographic information in the spatial context. In fact, the naive spatial context is a special case of the augmented spatial context where the factor $\beta$ equals to 1. For the simple augmented spatial contexts, our hypothesis is that the *popularity* of a POI as a context has a positive effect on the center POI whereas the influence of a context POI on a center POI decreases as the distance between them increases. By setting an augmenting factor $\beta$ based on these geographic components, we are stretching the original distribution of POI types in a manner that reveals more latent information in geographic space. To give an intuitive example for our rationale, a single place of the type `Stadiums & Arenas` may dominate a neighborhood while many individual parking spaces and bars only play a supportive function despite their higher frequencies.

## 4.4 ITDL Augmented Spatial Context

While the simple augmented spatial context approach models distance and human activities directly, the augmenting factor only applies to the original spatial context using the k-nearest neighbor method. In this sense, the context POIs are limited to the k nearest neighbors regardless of how far or how close they are from the center POI. However, different place types are likely to follow different spatial distributions and form distinct spatial clusters. For example, places of type `Restaurants` may be located closely to many other places of types such as `Hotels`, `Bars`, and `Department Stores`, generating a dense spatial cluster, while POI of type `Police Departments` and other area-serving places will show very different patterns when compared to nearby places (via their types). This spatial variation means that different spatial context information can be captured within different distances. In addition, the distance we are focusing on rapidly increases for such types, so naively setting a single threshold for the search buffer or the number of nearest neighbors will result in homogeneous spatial contexts for

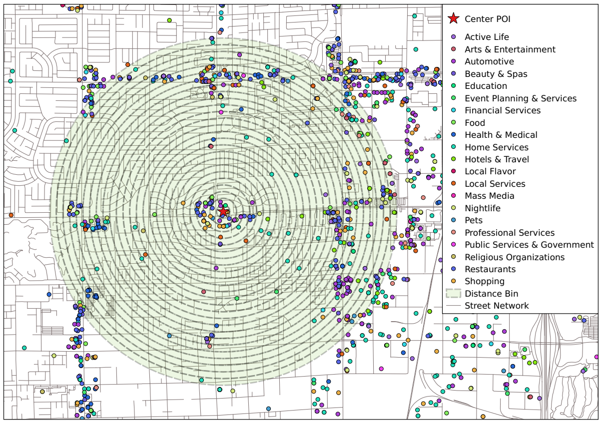


**Figure 2: ITDL augmented spatial context example.**

many different place types, thus sacrificing spatial heterogeneity and numerous distinguishing geospatial semantic characteristics. In light of this, we suggest having multiple different spatial contexts for each POI. Inspired by the use of semi-variograms in spatial statistics such as Kriging, we make use of *distance lags*, i.e., discrete bins, for constructing our spatial contexts. Such binning by a given lag also adjusts for the uncertainty (also called *tolerance*) of place centroids. In fact, previous work shows that the median distance of a POI between different database providers, such as Yelp and Foursquare, is 63 meters [17]. In the following, we will use a lag distance of $h = 100m$.

We use a default distance bin width for each distance lag, thus generating multiple spatial contexts for the same POI. Each spatial context can be used to learn a latent representation that encodes the distributional semantics between the center POI type and the context POI types within said distance bin. Our rationale behind this approach is that due to the nature (and function) of places and their interaction with other places and regions, an all-encompassing spatial context, even augmented with distance decay and human activities, is not sufficient for understanding the overall variation in the geographic patterns. Instead, we propose to first capture the local context by dividing the continuous geographic space, namely the distance, into discrete lags and then combine the semantic information from these different lags to obtain a more holistic global view of each place type; see Figure 2.

Since we aim to capture the spatial interaction between different place types, we want to set the maximum threshold of our spatial context based on this. We define $D_{ti}$ as the set of pair-wise POI distances of the same type $t_i$. For each POI type $t_i$, we calculate the minimum intra-class distance $min(D_{ti})$ and use the maximum of these intra-class distances as our threshold $TS$ for the spatial contexts (here the supremum of the per-type infimums):

$$TS = max(min(D_{t1}), min(D_{t2}), min(D_{t3}), ..., min(D_{tn})) \quad (7)$$

which is the maximum distance value, for at least one type among all place types, to search for context POIs that will not encounter the same type as the center. This $TS$ value helps to capture as much inter-class spatial interaction as possible. Hence, for each center POI, there are $s = \lfloor \frac{TS}{h} \rfloor$ spatial contexts.

For each spatial context, we propose a novel information theoretic, distance lagged augmentation method. The simple augmented spatial context takes into consideration distance decay and human activities, in the ITDL augmented spatial context, however, we focus on the human activities within the local context as well as the uniqueness of each place types per distance bin. The first component that incorporates human activities is defined as:

$$A = -log_2\left(1 - \frac{P_{tj}}{1 + \sum_{k=1}^{|M|} P_{tk}^h}\right) \quad (8)$$

where $P_{tj}$ is the *popularity* (check-in counts) of a place type $t_j$ and $\sum_{k=1}^{|M|} P_{tk}^h$ is the total number of check-in counts of all place types within a distance bin with width $h$. This is a monotonically increasing function with respect to $\frac{P_{tj}}{1+\sum_{k=1}^{|M|} P_{tk}^h}$, which means that if a place type has high *popularity* among all place types within the bin, this component value will be very high. The second component adopts the idea of information content (here, *surprisal*) from information theory to model the uniqueness of a place type given a distance bin:

$$U = -log_2(F_{tj}^h) \quad (9)$$

where $F_{tj}^h$ is the probability of encountering place type $t_j$ in a distance bin. $U$ essentially represents the information content of a place type $t_j$ within a distance bin. Larger $F_{tj}^h$ values will result in reduced information content. Finally, we integrate these two components using a convex combination and our ITDL augmentation is defined as:

$$\beta_{ITDL}^{Ij} = \lceil \omega A + (1 - \omega) U \rceil \quad (10)$$

where $\omega$ and $1 - \omega$ are the weights for the components. Intuitively, this allows us to distinguish unique places (of a certain type) that are highly popular from places that are popular in virtue of their type. Algorithm 1 shows the detailed procedures to construct the ITDL augmented spatial context $SC_{ITDL}$. In order to improve the efficiency of this algorithm, we split the whole task into $s$ tasks that can run in parallel, thus each worker only constructs a spatial context for one distance bin. In short, for the ITDL augmentation method, we use individual context settings to capture extrinsic components such as the popularity and the uniqueness of place types and use multiple spatial context bins combined to capture the intrinsic components such as distance and spatial variation.

## 5 EVALUATION SCHEMES

In this section, we introduce three different ground truths that we establish to evaluate our proposed methods. These ground truth results can also be used to evaluate other tasks involving place type similarity and relatedness. The first ground truth is built from the original Yelp place type hierarchy.[5] We take advantage of this *top-down* hierarchy and evaluate to what degree our *bottom-up* approaches can approximate Yelp's hierarchy. The second ground truth is obtained using Human Intelligence Tasks (HIT) via Amazon Mechanical Turk which is a binary test. The third one is obtained from another HIT which provides similarity and relatedness rankings for different POI types. These three ground truth results, one

[5] https://www.yelp.com/developers/documentation/v3/all_category_list/categories.json

**Algorithm 1:** Constructing ITDL-based Augmented Spatial Contexts $SC_{ITDL}$

**Input** : $L = (N, G, T)$, $s$, $h$, $\omega$
**Output** : $SC_{ITDL}$

$SC_{ITDL}$ := initialize list
**foreach** $l_i \in L$ **do**
    $T_{li}$ := a set of place types associated with $l_i$
    **for** $n = 0; n < s; n{+}{+}$ **do**
        $sc$ := check-in total of all place types in bin $n$
        $sp$ := POI total of all place types in bin $n$
        **foreach** $l_j \in L$ **do**
            $T_{lj}$ := a set of place types associated with $l_j$
            **if** $nh \leqslant d(l_i, l_j) < (n+1)h$ **then**
                **foreach** $t_{ki} \in T_{li}$ **do**
                    **foreach** $t_{kj} \in T_{lj}$ **do**
                        $cc$ := check-in of $t_{kj}$
                        $cp$ := count of $t_{kj}$
                        $A := -log_2(1 - cc/sc)$
                        $U := -log_2(cp/sp)$
                        $aug := \text{ceil}(\omega A + (1-\omega)U)$
                        append tuple $(t_{ki}, t_{kj})$ to $SC^n_{ITDL}$ $aug$ times
                    **end**
                **end**
            **end**
        **end**
    **end**
**end**

using top-down information from Yelp and the other two provided by human judges, provide a comprehensive evaluation for our work.

## 5.1 Hierarchy-based Evaluation Scheme

The original Yelp categories provide us with a natural way to calculate the similarity and relatedness of different POI types based on their hierarchical structure. There are two major ways to measure (semantic) similarity and relatedness for our tasks: distribution-based measures and knowledge-based measures [7]. While our proposed methods aims to capture the distributional semantics, the evaluation scheme derived from Yelp categories falls into the knowledge-based measures group. Numerous models have been proposed for such measures. In summary, edge-based measures and information content-based measures are two widely-used subgroups. In our study, we choose two measures from each subgroup to form our evaluation scheme. In addition, since the information content-based measures depend on the definition of information content, we also select two different definitions of information content in order to provide a more holistic evaluation scheme. In the end, we have 6 different measurements based on the Yelp hierarchy.

The first edge-based measurement is proposed by Wu & Palmer [27], which is defined as:

$$SIM_{WP}(t_1, t_2) = \frac{2N_3}{N_1 + N_2 + 2N_3} \tag{11}$$

$t_{lcs}$ is defined as the least common superclass of place types $t_1$ and $t_2$. $N_1$ is the shortest path from $t_1$ to $t_{lcs}$. $N2$ is the shortest path from $t_2$ to $t_{lcs}$. $N_3$ is the shortest path from $t_{lcs}$ to root. The second edge-based measurement is proposed by Leakcock & Chodorow [13]:

$$SIM_{LC}(t_1, t_2) = -log\left(\frac{N}{2D}\right) \tag{12}$$

where $D$ is the maximum depth of the taxonomy and $N$ is the shortest path between place types $t_1$ and $t_2$.

For the information content-based measurements, we use the models proposed by Lin [15] and Jiang & Conrath [11]. Their definitions are shown in Eq. 13 and Eq. 14, respectively. $IC$ is the information content of each place type and $t_{lcs}$ is the least common superclass of place types $t_1$ and $t_2$ within the Yelp hierarchy. Jiang & Conrath's method calculates the distance between $t_1$ and $t_2$, so the similarity is equal to $SIM_{JC}(t_1, t_2) = 1/DIS_{JC}(t_1, t_2)$.

$$SIM_{Lin}(t_1, t_2) = \frac{2IC(t_{lcs})}{IC(t_1) + IC(t_2)} \tag{13}$$

$$DIS_{JC}(t_1, t_2) = IC(t_1) + IC(t_2) - 2IC(t_{lcs}) \tag{14}$$

Both models proposed by Lin and Jiang & Conrath depend on the definition of information content, so we also include two different definitions of information content that can be calculated from the place type hierarchy. The information content proposed by Sánchez et al. [24] is defined as:

$$IC_{Sanchez} = -log\left(\frac{\frac{|leaves(t_i)|}{|subsumers(t_i)|} + 1}{max\_leaves + 1}\right) \tag{15}$$

where $|leaves(t_i)|$ is the number of leaves of place type $t_i$ in the hierarchy, $|subsumers(t_i)|$ is the number of place types that are more general than $t_i$ in the hierarchy and $max\_leaves$ is the number of leaves for the root place type. The information content proposed by Seco et al. [25] is defined as:

$$IC_{Seco} = 1 - \frac{log(|hypo(t_i)| + 1)}{log(max\_types)} \tag{16}$$

where $|hypo(t_i)|$ is the number of POI types that are more specific than $t_i$ and $max\_types$ is the maximum number of types in the hierarchy. Combining these definitions of information content with the methods by Lin and Jiang & Conrath, leads to four measures.

By using these semantic similarity measures, we calculate the pair-wise similarity of Yelp place types. Because these six measures differ in terms of what they measure, the resulting scores are also slightly different. Based on the similarity scores, for each place type, we generate a ranking of similar place types from the most similar to the least similar. We obtain six different groups of rankings for each of the POI types in Yelp. To confirm the validity of this evaluation scheme, we use Kendall's coefficient of concordance $W$ to assess the agreement among these six groups of rankings. The average Kendall's $W$ of all (1030) place types [6] among the six measurements is **0.981**, indicating a nearly perfect agreement among measures. Moreover, in our experiment, we use a subset of 93 place types (see Section 6) and the concordance remains stable at **0.979**. This result implies that our evaluation scheme based on the place type

[6] We only consider 570 place types, namely those that have at least 14 instances in our dataset and use various subsets of these 570 types in our experiments.

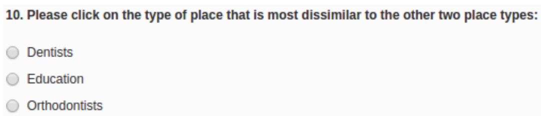


**Figure 3: Binary HIT example.**

hierarchy is valid. To evaluate the result, we mimic the task of geographic information retrieval, e.g. finding the most similar place type based on a given place type. By choosing the first place type in each of the 1030 rankings, we can obtain the result for all six measurements. To evaluate our latent representations, we generate our own rankings of each place type based on the augmented spatial contexts using pair-wise similarity [7] and use Mean Reciprocal Rank (MRR) to test the performance of our methods.

## 5.2 Binary HIT Evaluation Scheme

The hierarchy-based evaluation scheme has some potential drawbacks. First, the hierarchy is created by a small set of people which may lead to a bias. Moreover, in this hierarchy of more than 1000 place types (nodes), the average path length is only 1.73 which indicates that the taxonomy is very shallow. This will result in ties in the rankings generated using the hierarchical structure. Finally, a hierarchy always encodes some underlying ontological commitments, e.g., grouping arts and entertainment in a common class. Hence, in addition to the hierarchy-based evaluation, we utilize Amazon's Mechanical Turk for a binary HIT evaluation scheme.

For the HIT task, we generate 80 triplets with each element in the triplet being a place type. For example, one of the triplets is `(Dentists, Education, Orthodontists)`. [8] The task is to choose the place type from each triplet that is most dissimilar from the other two. For each place type in the triplet, a human judge will make a binary decision; see Figure 3. We published the HIT task on Amazon Mechanical Turk and each of these 80 tests was done by 25 human workers. The final result of each test is determined by the mode answer of the 25 human workers. For instance, the final answer for the test `(Dentists, Education, Orthodontists)` is `Education` as this is the most often excluded type.

To evaluate the latent representations generated by augmented spatial contexts, for each triplet, we calculate the pair-wise similarity score using 2-combination. For example, for the above mentioned triplet, we calculate the similarity scores of three pairs `(Dentists, Education)`, `(Dentists, Orthodontists)` and `(Education, Orthodontists)`. We pick the one with the highest score and return the other place type as the result for this test using our methods. For instance, if `(Dentists, Orthodontists)` has the highest score, then `Education` is the result from our methods. We evaluate the accuracy of different methods on all triplets.

## 5.3 Ranking-based HIT Evaluation Scheme

While the binary-based HIT evaluation can complement the Yelp hierarchy task by relying on human judges, the task is relatively easy. Hence, for the ranking-based HIT evaluation scheme, we want to use human judges to generate a ranking result for each place type.

[7] All similarity scores for our place type embeddings are calculated using Cosine Similarity.
[8] See Goodman's deliberation on similarity for a rationale about using triples [6].


**Figure 4: Ranking-based HIT, showing one MTurk result.**

We selected 10 place types and for each place type we selected 7 candidate place types for ranking, so altogether we have 70 POI type pairs. We ask 25 human judges on Amazon Mechanical Turk to rate on a scale of 1-7 the similarity of each of these pairs. Such task can be considered as very challenging in the context of studying semantic similarity [10] and requires more attention to user interface design (Fig. 4) to adjust for some well-known characteristics of human similarity judgments, notably that such judgments are known to be non-symmetric. In addition, we selected a slider-based design to ease visual comparison between pairs; see [6].

After receiving the results, we have rankings of each place type from 25 human judges. In order to check if the rankings are consistent, and, thus, whether the task is meaningful, we use Kendall's coefficient of concordance $W$ to evaluate the agreement score among the judges. The average Kendall's $W$ score over all place types in the test is **0.79** which indicates very high agreement.

In order to evaluate our place embeddings using the proposed augmented spatial contexts, we generate a ranking for each place type based on the pair-wise similarity score. We then calculate the average Spearman's rank correlation coefficient between our rankings and the rankings from the HIT task as the criteria to evaluate the performance of our models.

# 6 EXPERIMENT AND RESULT

In this section, we discuss the experiments to evaluate our work and their results. We also point to an interesting research question that arises from our work. First, we have to define the number of dimensions for the POI type embeddings. Next, we compare our embeddings with the state-of-the-art word embeddings trained from the Google News corpus as a baseline using the proposed evaluation schemes in order to reiterate the necessity of augmenting spatial

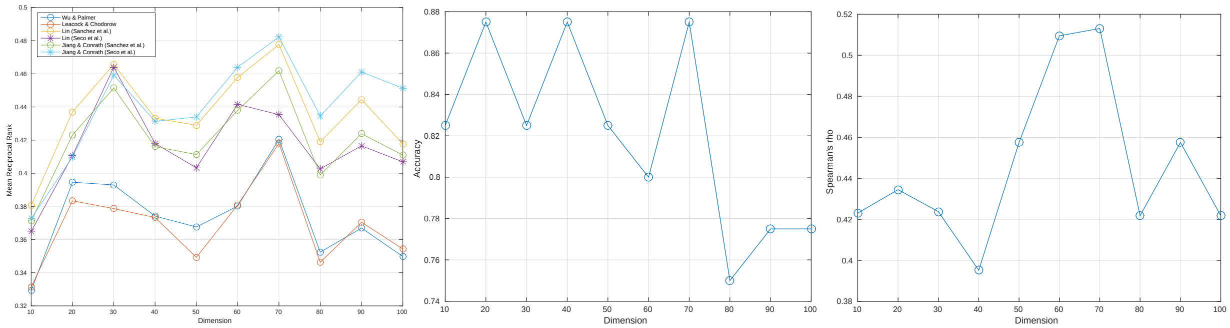


**Figure 5: Left to right, Mean Reciprocal Rank (MRR) for the hierarchy-based evaluation, accuracy for the binary HIT evaluation, and Spearman's $\rho$ for the ranking-based HIT evaluation.**

contexts to obtain richer semantic information from geographic space. In addition, we visualize and analyze different embeddings spaces from different augmented spatial contexts using dimension reduction techniques and present *place type profile* as a visual assistance tool for understanding place type similarity and relatedness. Finally, we briefly look at a very interesting research question that arises from our work, namely whether there is potential for compression by merely using a *subset* of POI types to learn *all* POI types. From an urban planning perspective, this question can also be framed from a summarization perspective, by asking whether there are certain place types that are *indicative* of a neighborhood (when modeled as a set of POI) .

### 6.1 Selecting Dimensions

An important parameter for latent representation models is the number of dimensions for the embedding vectors. As the total number of place types is relatively small compared with the vocabulary size of natural language, we selected dimensions ranging from 10 to 100 with a step interval of 10 to determine the number of optimal dimensions for our model. Since we want to combine both intrinsic and extrinsic information in our spatial context, we focus on using the augmenting factor $\beta^{lj}_{combined}$ in this task, which takes into consideration the influence of geographic distance and POI *popularity*. Figure 5 shows the dimension test result using the Yelp hierarchy-based evaluation scheme, the binary HIT test, and the ranking-based HIT. Although there is a variation in the absolute values of the six measurements, the overall trend is very similar. It shows that using **70** dimensions yields the best overall results and we will use this number for the experiments described below.

### 6.2 Comparison

By introducing the augmented spatial contexts, we want to demonstrate the richness of semantic information latently encoded in geographic patterns. First, to justify the need for POI type embeddings, we compare the evaluation results of the word embeddings trained from the Google News corpus with the place type embeddings trained from Yelp POIs and our augmented spatial contexts. Word embeddings have been used in a variety of information retrieval tasks and have been frequently used as proxies for geographic information retrieval. Many of the word embeddings techniques, however, only consider unigrams, such as the pre-trained Word2Vec embeddings from Google, which means that they are not suitable for many place type names, such as `Auto Repair`. In addition, and as argued above, geographic space is inherently different from word space, and, thus, word embeddings lack the ability to capture spatial interaction among different geographic entities and distance (decay) effects which is a significant factor in measuring place type similarity and relatedness.

In order to support our argument, we compared the word embeddings with the proposed place type embeddings using different spatial contexts, namely one with the naive spatial context and four with the augmented spatial contexts. Recall that there is a weight parameter $\omega$ in the ITDL augmented spatial contexts, to adjust the relative importance of $A$ (activity) and $U$ (uniqueness). We tested our model with $\omega$ values ranging from 0.1 to 1 with 0.1 as step interval. Our $TS$ value is 2644.5 meters, so the total number of spatial contexts for each $\omega$ value for the ITDL approach and a lag of 100m is $s = \lfloor 2644.5/100 \rfloor = 26$. In the end, we can obtain 234 different augmented spatial contexts and learn place type embeddings from each of these contexts using parallel threads. In order to compare the evaluation results, for each $\omega$ value, we test the performance of each of the 26 bins and concatenate the embedding vectors of the top five bins to generate the final place type embedding of 350 dimensions. We use the best $\omega$ values as our final result of the ITDL augmented spatial contexts.

We compared the pre-trained Google Word2Vec result with our place type embeddings using both the hierarchy-based evaluation scheme and the binary HIT evaluation scheme. $SC_{naive}$ is the spatial context without augmentation. $SC_{checkin}$, $SC_{distance}$, $SC_{combined}$ and $SC_{ITDL}$ are the methods detailed in Section 4. Table 1 shows the result of the hierarchy-based evaluation. As mentioned earlier, word embeddings trained using Google News corpus only contain unigrams, so we select a subset (93 place types) as our testing data. All methods are tested using the six measures described in Section 5. Table 2 shows the binary and ranking-based HIT results. The hierarchy and binary evaluations show that the results obtained by using spatial contexts, even without any augmentation, are substantially better than the one purely based on a linguistic perspective, thereby also showing the benefits of our approach over previous work outlined in Section 2. This confirms our hypothesis that geographic space carries rich latent semantic information that cannot be captured by the word space alone.

**Table 1: Mean Reciprocal Rank for the hierarchy-based evaluation.**

| Model | $SIM_{WP}$ | $SIM_{LC}$ | $SIM_{Lin}$ ($IC_{Sanchez}$) | $SIM_{Lin}$ ($IC_{Seco}$) | $SIM_{JC}$ ($IC_{Sanchez}$) | $SIM_{JC}$ ($IC_{Seco}$) |
|---|---|---|---|---|---|---|
| Word2Vec | 0.288 | 0.321 | 0.354 | 0.334 | 0.349 | 0.333 |
| $SC_{naive}$ | 0.412 | 0.398 | 0.474 | 0.442 | 0.455 | 0.478 |
| $SC_{checkin}$ | 0.385 | 0.387 | 0.448 | 0.428 | 0.452 | 0.474 |
| $SC_{distance}$ | 0.381 | 0.396 | 0.458 | 0.426 | 0.443 | 0.458 |
| $SC_{combined}$ | 0.420 | 0.418 | 0.478 | 0.435 | 0.462 | 0.482 |
| $SC_{ITDL}$ | **0.447** | **0.431** | **0.498** | **0.479** | **0.487** | **0.483** |

For the ranking-based evaluation scheme, we dropped the Google Word2Vec embeddings to be able to use bigrams and because using a merely linguistic context already did not perform well for the two simpler tasks. In all three evaluations the ITDL augmented spatial contexts is able to model more semantic information, and, thus, yields better results for the place type similarity tests. With a $\rho$ of **0.7**, i.e., a strong correlation with human judgments, and an accuracy of **0.95** this becomes most apparent for the more difficult HITs. This is a remarkable result as humans utilize substantially richer information to reason about similarity, e.g., the meaning (and similarity) of the type labels, background knowledge, e.g., about cultural and historic reasons why Asian foods are alike, and so forth. Financially, it is worth mentioning that short as well as long-distance bins contribute to these results, e.g., the highest $\rho$ is obtained by a concatenation of bins 4-17-1-5-24 ($\omega = 0.1$), where 24 represents the 100m distance lag at 2400 meters from the center POI.

**Table 2: Accuracy for binary HIT evaluation and Spearman's $\rho$ for ranking-based HIT.**

| Model | Accuracy |
|---|---|
| Word2Vec | 0.750 |
| $SC_{naive}$ | 0.850 |
| $SC_{checkin}$ | 0.700 |
| $SC_{distance}$ | 0.875 |
| $SC_{combined}$ | 0.875 |
| $SC_{ITDL}$ | **0.950** |

| Model | $\rho$ |
|---|---|
| $SC_{naive}$ | 0.56 |
| $SC_{checkin}$ | 0.56 |
| $SC_{distance}$ | 0.57 |
| $SC_{combined}$ | 0.51 |
| $SC_{ITDL}$ | **0.70** |

## 6.3 Place Type Profiles

Although we use the concatenated place type embeddings in our evaluation, individual augmented spatial context can be used separately for analyzing the characteristics of different place types. Here we propose a 3D visualization, namely *place type profile* as a tool to compare different POI types and their semantic relationships. We use t-Distributed Stochastic Neighbor Embedding (t-SNE) [16] to reduce our place type embeddings in each distance bin into two dimensions, then stack each of these 2D space together to build a 3D profile. Figure 6 shows the profiles of selected types generated with $\omega = 0.5$, the x-axis and y-axis are the two components after dimension reduction using t-SNE and the z-axis is the distance bin.

One can see that `Bars`, `Restaurants` and `Hotels` always cluster together no matter which distance bin they are in. `Police Departments` are a certain distance apart in each bin. `Health & Medical` remains far away from all other POI types. This pattern shows that `Bars`, `Restaurants`, and `Hotels` have very similar contexts in each distance bin, which implies that they interact in similar

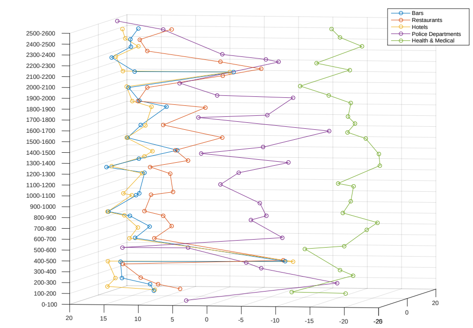


**Figure 6: Place Type Profile with $\omega = 0.5$.**

**Table 3: Place type compression result.**

| Model | Accuracy | $\rho$ |
|---|---|---|
| All Place Types | 0.950 | 0.70 |
| W/O Restaurants | 0.925 | 0.70 |
| W/O Nightlife | 0.925 | 0.70 |
| W/O Professional Services | 0.925 | 0.68 |
| W/O Health & Medical | 0.900 | 0.68 |
| W/O 18 Place Types | 0.875 | 0.59 |

ways with other POI type. We will return to this argument when discussing compression potential next.

## 6.4 Place Type Compression

So far, our experiments are all based on all POI types, which means that we generate our training data for each augmented spatial context using all types and run the latent representation model to retrieve place type embeddings. However, this approach is time-consuming as the number of $(t_{center}, t_{context})$ pairs increases in later distance bins and may also lead to overfitting. In order to obtain more condensed results, we proposed the novel idea of place type compression. Our intuition is that many place types such as `Restaurants` and `Nightlife` are co-located with other types (via their POI) following similar patterns. Hence, our hypothesis is that these types can serve as proxies in the sense that we can omit, for instance, all nightlife places (and places of their 17 subtypes) and still learn good embeddings for *all* types including `Nightlife`. Some place types such as `Professional Services` have weaker interaction patterns with other place types, thus making it harder to represent them by other POI types.

In order to test our hypothesis, we select four different root place types: `Restaurants`, `Nightlife`, `Professional Services`, and `Health & Medical`. We remove each of these place types and their subtypes from the context POI types in our training and run our models using the ITDL augmented spatial contexts. In addition, we run our model by removing all 18 place types aside of those four (there are 22 root place types). The accuracy result of the binary HIT evaluation and the Spearman's $\rho$ result of the ranking-based HIT are shown in Table 3. The result shows that dropping either `Restaurants` or `Nightlife` does not have much effect on the final embeddings while dropping either `Professional Services` or `Health & Medical` will result in a (small) decrease in performance. Consequently, given the 570 studied types, removing even 69 from them, e.g., by removing the `Restaurants` supertype, leaves us with enough proxy types, i.e., types that interact with other types in similar ways. Dropping 18 place supertypes, however, and trying to generate embeddings merely on the 4 remaining supertypes will result in a substantial decrease. This confirms our hypothesis that we can compress our model and still obtain high-quality latent representations of place types.

## 7 CONCLUSION AND FUTURE WORK

In this research we proposed a novel approach, namely augmented spatial contexts, to capture the semantics of place types by learning vector embeddings and using them to reason about place type similarity and relatedness, a common prerequisite for geographic information retrieval. By comparing the place type embeddings generated using the proposed methods with state-of-the-art word embeddings, we were able to show that our information-theoretic, distance lagged augmented spatial contexts substantially outperform the baseline and better capture the latent semantic information. We also established three different evaluation schemes to systematically evaluate the resulting POI embeddings. We published the embeddings as well as the HIT results online to foster reproducibility and in the hope that they will be reusable by others working on vector representations of place types. We used place type profiles as a way to visualize the semantic relationship among different place types. Finally, we outlined the idea of indicative POI types and their usage in compression as a novel research avenue.

In the future, we will explore place type compression in more detail to determine how different combinations of POI types can affect the quality of the overall place type embeddings and will follow up on the idea of using them to summarize neighborhoods. Finally, we focused on geodesic distance here but our methods can be generalized, e.g., using L1 distance (taxicab), in future work.

## REFERENCES


[1] Benjamin Adams and Krzysztof Janowicz. 2015. Thematic signatures for cleansing and enriching place-related linked data. *International Journal of Geographical Information Science* 29, 4 (2015), 556–579.

[2] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. 2003. A neural probabilistic language model. *Journal of machine learning research* 3, Feb (2003), 1137–1155.

[3] Anne Cocos and Chris Callison-Burch. 2017. The Language of Place: Semantic Value from Geospatial Context. *EACL 2017* (2017), 99.

[4] Shanshan Feng, Gao Cong, Bo An, and Yeow Meng Chee. 2017. POI2Vec: Geographical Latent Representation for Predicting Future Visitors. (2017).

[5] John R Firth. 1957. A synopsis of linguistic theory, 1930-1955. (1957).

[6] Nelson Goodman. 1972. Problems and projects. (1972).

[7] Sébastien Harispe, Sylvie Ranwez, Stefan Janaqi, and Jacky Montmain. 2015. Semantic similarity from natural language and ontology analysis. *Synthesis Lectures on Human Language Technologies* 8, 1 (2015), 1–254.

[8] Stevan Harnad. 2005. To cognize is to categorize: Cognition is categorization. *Handbook of categorization in cognitive science* (2005), 20–45.

[9] Krzysztof Janowicz. 2012. Observation-driven geo-ontology engineering. *Transactions in GIS* 16, 3 (2012), 351–374.

[10] Krzysztof Janowicz, Martin Raubal, and Werner Kuhn. 2011. The semantics of similarity in geographic information retrieval. *Journal of Spatial Information Science* 2011, 2 (2011), 29–57.

[11] Jay J Jiang and David W Conrath. 1997. Semantic similarity based on corpus statistics and lexical taxonomy. *arXiv preprint cmp-lg/9709008* (1997).

[12] Junchul Kim, Maria Vasardani, and Stephan Winter. 2017. Similarity matching for integrating spatial information extracted from place descriptions. *International Journal of Geographical Information Science* 31, 1 (2017), 56–80.

[13] Claudia Leacock and Martin Chodorow. 1998. Combining local context and WordNet similarity for word sense identification. *WordNet: An electronic lexical database* 49, 2 (1998), 265–283.

[14] Wentian Li. 1992. Random texts exhibit Zipf's-law-like word frequency distribution. *IEEE Transactions on information theory* 38, 6 (1992), 1842–1845.

[15] Dekang Lin et al. 1998. An information-theoretic definition of similarity.. In *Icml*, Vol. 98. 296–304.

[16] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9, Nov (2008), 2579–2605.

[17] Grant McKenzie and Krzysztof Janowicz. 2015. Where is also about time: A location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures. *Computers, Environment and Urban Systems* 54 (2015), 1–13.

[18] Grant McKenzie, Krzysztof Janowicz, Song Gao, and Li Gong. 2015. How where is when? On the regional variability and resolution of geosocial temporal signatures for points of interest. *Computers, Environment and Urban Systems* 54 (2015), 336–346.

[19] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv:1301.3781* (2013).

[20] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. 3111–3119.

[21] Andriy Mnih and Koray Kavukcuoglu. 2013. Learning word embeddings efficiently with noise-contrastive estimation. In *Advances in neural information processing systems*. 2265–2273.

[22] Christoph Mülligann, Krzysztof Janowicz, Mao Ye, and Wang-Chien Lee. 2011. Analyzing the spatial-semantic interaction of points of interest in volunteered geographic information. In *International Conference on Spatial Information Theory*. Springer, 350–370.

[23] Gianluca Quercini and Hanan Samet. 2014. Uncovering the spatial relatedness in Wikipedia. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 153–162.

[24] David Sánchez, Montserrat Batet, and David Isern. 2011. Ontology-based information content computation. *Knowledge-Based Systems* 24, 2 (2011), 297–303.

[25] Nuno Seco, Tony Veale, and Jer Hayes. 2004. An intrinsic information content metric for semantic similarity in WordNet. In *Proceedings of the 16th European conference on artificial intelligence*. IOS Press, 1089–1090.

[26] Yi-Fu Tuan. 1977. *Space and place: The perspective of experience*. Uni. of Minnesota.

[27] Zhibiao Wu and Martha Palmer. 1994. Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 133–138.

[28] Yao Yao, Xia Li, Xiaoping Liu, Penghua Liu, Zhaotang Liang, Jinbao Zhang, and Ke Mai. 2017. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science* 31, 4 (2017), 825–848.

[29] Mao Ye, Krzysztof Janowicz, Christoph Mülligann, and Wang-Chien Lee. 2011. What you are is when you are: the temporal dimension of feature types in location-based social networks. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 102–111.

[30] Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang, and Jiawei Han. 2017. Regions, periods, activities: Uncovering urban dynamics via cross-modal representation learning. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 361–370.

[31] Yating Zhang, Adam Jatowt, and Katsumi Tanaka. 2017. Is Tofu the Cheese of Asia?: Searching for Corresponding Objects across Geographical Areas. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 1033–1042.

[32] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R Lyu. 2017. Geo-Teaser: Geo-Temporal Sequential Embedding Rank for Point-of-interest Recommendation. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 153–162.

[33] Rui Zhu, Yingjie Hu, Krzysztof Janowicz, and Grant McKenzie. 2016. Spatial signatures for geographic feature types: Examining gazetteer ontologies using spatial statistics. *Transactions in GIS* 20, 3 (2016), 333–355.