Uncovering the Perception of Urban Outdoor Areas Expressed in Social Media

Frances A. Santos[1], Thiago H. Silva[2],
Antonio A. F. Loureiro[3], Leandro A. Villas[1]
[1]Institute of Computing, University of Campinas, Campinas, SP – Brazil
[2]Department of Informatics, Federal University of Technology of Paraná, Curitiba, PR – Brazil
[3]Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, MG – Brazil

***Abstract*—Learning about people's perception that emerges from urban areas has been an interesting multidisciplinary research goal because it has a great potential to ease the hard task of understanding intrinsic characteristics of urban areas. To this end, we propose an approach that explores spatial and semantic aspects in free-text messages shared on location-based social networks (LBSNs) for uncovering and mapping the perception reflected regarding urban outdoor areas. Studying outdoor areas of Chicago, we show that LBSN data carry valuable information about places and could also be used to extract urban perception, helping to better understand urban areas from many aspects. We demonstrate, through a survey with volunteers, that our approach has the potential to correctly capture the opinion considered by the users regarding the reflected perception of those areas, indicating that it could be a feasible alternative for the task under study.**

***Index Terms*—Perception Extraction, Social Media, Text Mining, Outdoor Areas, Dictionary**

I. Introduction

Urban outdoor areas, such as parks, streets, plazas, may offer people the opportunity of having diverse experiences, and, for this reason, could trigger different perceptions among its visitors. For instance, places with touristic attractions can be appreciated by tourists, but it can be considered too crowded and noisy for residents performing their daily routines. Due to that, understand the perception reflected by urban areas, which could improve their description and leverage new services and applications, is a hard task.

Field surveys and sensory walks are traditional approaches used to understand the intrinsic semantic of urban areas, being a powerful tool to capture the detailed opinions of people about places [1], [2]. However, those strategies could be time-consuming because they typically demand a high amount of time of observation and interviews of participants to collect a considerable amount of perception samples. This makes it difficult to perform this type of analysis for a large number of places. To overcome this challenge, we propose an approach that exploits location-based social network (LBSNs) data.

Exploring LBSNs for this task is interesting for various reasons. One of the main ones is that a large number of users act as social sensors sharing a considerable amount of content about urban areas, including their opinions [3]. In addition, the content is rich. It is possible to find general comments on nearly everything. However, while LBSNs can offer a vast amount of data, which can potentially help in the scalability problem of traditional methods for collecting urban perceptions, their exploitation for this purpose is not trivial, since the extraction of useful urban perceptions expressed by users in these systems is a challenging task.

Previous studies have found evidence that it is possible to extract relevant perceptions regarding urban areas from the content of LBSNs shared by users [2], [4]. However, those studies rely on traditional approaches, such as sensory walks, suffering from lack of flexibility and scalability problems. Our study offers an alternative possibility that does not demand traditional approaches, being generic enough to work on virtually any thematic by only exploring public content shared by users on the Web.

For that, we first propose a semiautomatic way to create a hierarchical dictionary of possibly any thematic. Having this dictionary is an important step in extracting the user's urban perception from LBSN data. We demonstrated our methodology by constructing a dictionary, namely UOP-dictionary, which organizes the main words used by people to qualify their experiences in urban outdoor areas of cities. Besides, we present an unsupervised clustering algorithm to identify content shared by users in free-text messages that are spatially and semantically similar, which is fundamental to identify the relevant perception of outdoor areas.

We have used a Twitter dataset to demonstrate the potential of our approach to uncover the perceptions of urban outdoor areas that emerge from LBSNs. Studying Chicago, we show that LBSN data brings valuable information about urban areas and can be used to extract their perceptions, helping to better understand these areas in many aspects. We validate the perceptions extracted from different urban outdoor areas by conducting a survey with volunteers, indicating that our approach yields result very similar to those indicated by users.

The rest of the study is organized as follows. Section II presents the related work. Section III presents our approach to create hierarchical dictionaries. Section IV presents the algorithm to identify spatial and semantic similarities in free-text messages shared by users. Section V map the identified perceptions for Chicago, in addition, it presents and discusses a user evaluation of those perceptions. Section VI discuss implications and possible limitations of our results. Finally, the study is concluded in Section VII.

II. Related Work

In the literature, it is possible to find proposals that aim to extract important features regarding urban areas in different ways, including offline and online ones. Offline sources, e.g., via sensory walks, typically, provide fine-grained data from participants engaged and focused on the given tasks, however, they do not scale easily [2], [5]. To overcome this problem, proposals have been made to extract urban characteristics shared by users in online sources, in order to extract users' perceptions about urban areas for different purposes. The most commonly used online source is social media, especially location-based social networks (LBSNs), which have been shown to be useful for understanding and solving different problems, such as recommending more pleasant routes for pedestrians [6], uncovering the cultural preferences [7], [8] and the city dynamics [9].

Towards that direction, there are some studies focused on understanding the effects that visual, olfactory and auditory features present in the urban environments have on people's perception. Quercia et al. [10] collected people's perceptions about photos taken in streets of London and they found correlation among colors, texture and visual patches of those photos with beautiful, happy, and quiet perceptions.

In the same direction, Naik et al. [11] proposed an approach for predicting the perceived safety of cities exploring images from different sites. Similarly, Porzi et al. [12] demonstrate that computer vision techniques can also identify human perception and predict judgments of safety from images of urban scenes. Dubey et al. [5] explored a global dataset of urban images available online to rank street-level images of urban appearance by also using computer vision techniques. Their results demonstrate that urban perception data at global scale could be extracted from online images.

Quercia et al. [2] explored sensory walks [1] to collect the citizens' perception concerning the smell of the environment. This offline process has enabled the creation of a dictionary with urban-smell related words, which is used to discover messages related to odor perceptions into social media data. A follow-up of this study was presented in [13], where the authors investigated, among other things, the relationship between the predominant color of the image (visual perception) and the smell associated with the image (olfactory perception).

Aiello et al. [4] explore the influence that urban sounds have in the way people perceive places. To discover the urban sounds responsible for triggering the people perceptions and emotions, the authors created a dictionary with sound-related words, which was used to mine LBSN data related to urban sounds. In this way, the streets of London and Barcelona could be mapped with one of the six considered categories (transport, nature, human, music, indoor) of the dictionary created.

As some of the related works, we also take advantage of LBSN data to facilitate the process of understanding aspects of urban areas. However, our study differs from all the previous efforts, because we propose an approach to accomplish this task without requiring time-consuming field surveys or manual steps. Besides, our approach is flexible and generic, enabling, with easy adaptations, the study of different aspects shared in LBSN data, not only about urban areas.

TABLE I
Places reviews dataset statistics.

	Chicago	**London**	**NYC**
Foursquare Tips	5,085	7,261	24,921
Google Places	666	662	753

III. Urban Outdoor Perception Dictionary

As the use of Twitter free-text opinions to extract urban perception is hard because the content can be very noisy, it is important to first explore a less noisy source of opinions of outdoor areas. This is important for learning properly the vocabulary used by users to this end. Some social media sites, such as Google Places, TripAdvisor, and Foursquare, enable users to do reviews about places containing their personal opinions. In these systems, users can share reviews about any place already available in the system, e.g., a specific restaurant, at any time, not necessarily when the user was in the place formulating their opinion about it. Despite having this issue and relatively fewer data publicly available compared to data shared in Twitter, the reviews tend to bring rich details of users' perception, as well as ease the separation of content specific from outdoor places. For this reason, they are explored in this study to learn the vocabulary commonly used by users regarding outdoor areas.

We collected public reviews from Google Places[1] and Foursquare (Tips[2]) written in English language and shared by users in Chicago and New York City (NYC), United States, and London, United Kingdom, before February 2017, dataset called *Places Review*. Both sites define discrete sets of categories that specify the type of places[3] [4], which enables the selection of just reviews about outdoor places. Table I summarizes the number of reviews collected. As we can see, due to the restriction of Google Places API (at most five latest reviews per place), the documents from Foursquare represents most of the places reviews dataset. In total, 39,348 reviews compose the dataset. Based on the places review dataset, we are interested in uncovering what words are frequently used by people to qualify their experiences in urban outdoor areas to build an urban outdoor perception dictionary, namely UOP-dictionary.

In order to combine data and meta-data from both heterogeneous online sources, each review must contain a free-text format, a timestamp representing when the review was created, and a unique identifier of review. More formally, our dataset can be defined as follows:

[1]https://goo.gl/jMQsZm.
[2]The reviews of places in this system receives the name of tips.
[3]Foursquare Venue Category Hierarchy: https://goo.gl/cSFas4.
[4]Google Places Category: https://goo.gl/ACd1AT.

Definition 1: A collection $\mathcal{D}_R$, where each $doc \in \mathcal{D}_R$ is a document determined by a tuple $doc = (id, s, \tau)$, where id is an unique identifier, s is a list of sentences that comprises all content written by the user, $\tau \in \mathbb{R}$ is a timestamp.

In this work, the term sentence is used to refer the preprocessed free-text, where numbers, Uniform Resource Locators (URLs), special characters, punctuation, and stop words were removed, resulting in a vector of single words (tokens).

To build UOP-dictionary, first, we perform part-of-speech tagging in each sentence $s \in doc, \forall doc \in \mathcal{D}_R$, to classify which tag, among noun, adjective, verb, pronoun, etc., is most likely for words of s. We used the Perceptron Tagger from NLTK Toolkit, but others could be used as well. After labeling all words, we extract a set of words considered qualifiers (i.e., labeled with the adjective tag), that occurred at least 20 times in $\mathcal{D}_R$. After that, we asked three supervisors to analyze these set of words independently to generate a subset that effectively qualifies an urban outdoor area in their opinion. Then, we combine the three sets of words, keeping qualifiers present in at least two of them. As a result, we obtain our UOP-dictionary containing 88 English words.

After that, we organize the words of the UOP-dictionary into categories, according to the syntactic and semantic similarity of them, using for that the Word2Vec model. Word2Vec is a neural network based model used to learn the vector representations of words that contain many linguistic regularities and patterns [14]. It takes as input a corpus C (i.e., all sentences in docs of $\mathcal{D}_R$), a window size ws, a minimum count of occurrence of the word $minCount$, and a hyper-parameter m representing the number of features. Then, the model creates a vocabulary of n unique words denoted by W from C, where each word $w \in W$ must occur at least $minCount$ times in C. By using a deep learning method with a single hidden layer, either skip-gram model or continuous bag-of-words model (CBOW), and some activation function only on output neurons (not on hidden layer neurons), either hierarchical softmax or negative sampling, Word2Vec model computes for each pair of words $w_1, w_2 \in W : w_1 \neq w_2$, the probability to find them "nearby" into sentences of C. Two words are considered nearby if they are at most $ws - 1$ positions between them.

As there are n unique words and m features on the hidden layer, after training the Word2Vec model it produces $n \times m$ weights of the hidden layer, which are called word vectors. Such vectors enable to identify if two different words have similar contexts, checking if their word vectors are similar. We empirically defined $ws = 8$, $minCount = 200$, and $m = 300$, where the size of *corpus* is $C = 54,612$, resulting in a model with 217 unique words. The CBOW model and negative sampling worked better in our experiments; therefore, we employed them. Then, we use hierarchical clustering to group words according to the similarity of their word vectors. By experimenting with different linkage criteria and similarity/distance metrics, we find that the complete linkage and cosine similarity produced a better result, for this reason, we kept this configuration.

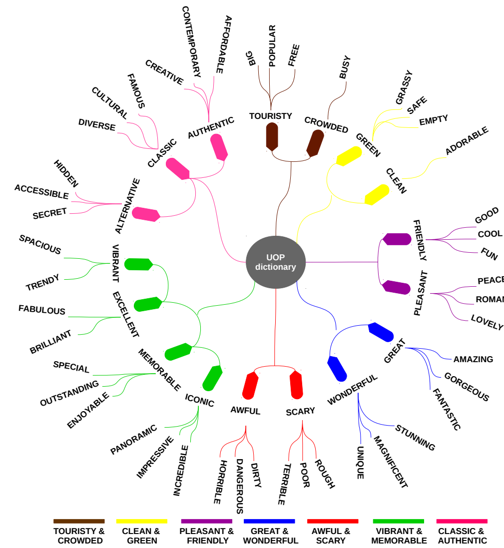


Fig. 1. UOP-dictionary

After these steps we have obtained what we call UOP-dictionary, representing the result of the hierarchical grouping. An illustration of this dictionary is shown in Figure 1. The way this image is organized is based on the dendrogram resulted from the clustering process. After cutting the dendrogram, each color represents a resulting cluster containing qualifiers with higher similarity among them. In the visualization we chose a word that represents a certain branch (subcluster) in a cluster, for instance, the branch Fabulous, Brilliant and Excellent, of the cluster green is represented by the word Excellent. Note that the cluster green also has three other branches, represented by Vibrant, Memorable, and Iconic. To simplify the cluster labeling we chose only two words that describe branches in case there are more. For the green cluster, Vibrant & Memorable is the name that represents it instead of all four words.

As we can see, there are seven main categories (i.e., clusters) that can be used to attribute a perception of some outdoor place. This dictionary is an important step to discover the perception of outdoor places based on noisy LBSN data. All the words that that encompasses the UOP-dictionary are presented in Table II. Next, we present each category based on an extensive investigation of phrases that originated them.

The category Touristy & Crowded represents areas with touristic sights. In addition, this category also represents crowded areas, which is commonly associated with touristic areas but not necessarily. Turning our attention to the category Vibrant & Memorable, we found that it represents the perception of areas that might be able to impress its visitors, for example, with natural and human-made beauties, which are rich in visual characteristics and good for taking photos, contemplation and enjoy the scenery. Studying the category Clean & Green, we found that it represents the perception of areas with significant presence of nature, such as parks and lakes, where green landscape and few (or none) amount

TABLE II
WORDS FOR EACH CATEGORY THAT COMPOSE THE UOP-DICTIONARY.

Category Label	All Words of the Category
Touristy & Crowded	Big, busy, crowded, free, popular, pretty, touristy.
Vibrant & Memorable	Attractive, brilliant, enjoyable, excellent, fabulous, happy, iconic, impressive, incredible, memorable, outstanding, panoramic, rare, spacious, special, trendy, vibrant.
Clean & Green	Adorable, clean, empty, grassy, green, safe, shady.
Pleasant & Friendly	Convenient, cool, cute, different, friendly, fun, good, interesting, lovely, nice, peaceful, perfect, pleasant, quiet, relaxing, romantic, serene, tranquil.
Awful & Scary	Awful, cheap, dangerous, dirty, horrible, loud, poor, rough, scary, simple, terrible.
Classic & Authentic	Accessible, affordable, alternative, authentic, classic, contemporary, creative, cultural, diverse, famous, hidden, secret.
Great & Wonderful	Amazing, awesome, beautiful, calm, colorful, cozy, fantastic, gorgeous, great, magical, magnificent, scenic, spectacular, stunning, unique, wonderful.

of human-discarded waste favor people to enjoy an adorable environment. Perhaps, one might think that the word "shady" should not belong to this cluster. However, we found that it is usually employed to describe shady of trees. This is an example of our dataset: *"Trees are very shady around the ages on the water. So ideal for summer sun breaks"*.

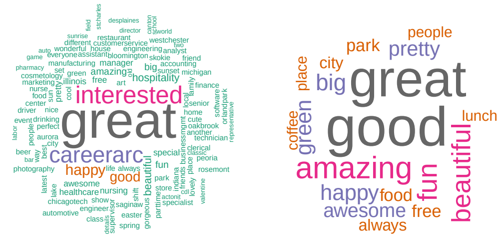
The category Pleasant & Friendly represents the perception of areas that might enable good experiences to its visitors, which might be suitable, for example, to spend time with friends and have romantic encounters. Investigating the category Awful & Scary, we found that it represents the perception of areas that might have triggered bad experiences to its visitors. Perhaps, one might think that the word "cheap" is out of context in the cluster; however, it is commonly related to the kind of stores around the area under evaluation, and whose the word vector is similar to the context of awful and scary areas. For instance, these are examples of our dataset: *"Don't like this place much just good to buy cheap souvenirs"* and *"Chinatown has gotten seedy and irrelevant. Cheap, aggressive street vendors push their counterfeit wares. It lost its charm and allure. Streets have an awful stench."*.

Studying the category Classic & Authentic, we found that it represents the perception of areas that might be able to provide unique experiences to its visitors. Finally, the category Great & Wonderful tend to represent areas that might have the potential to provide outstanding experiences to its visitors.

After this investigation, we confirmed that our methodology is able to organize words in groups that are similar to each other to describe urban outdoor places. By having coherent groups of words, i.e., categories, the task to separate and classify relevant content from noisy LBSN data, such as data from Twitter, becomes more feasible.

IV. EXTRACTION OF PERCEPTIONS

In order to extract the people's perception from messages shared via LBSN about urban outdoor environments, we considered in this study public messages (tweets) of Twitter. Twitter is an online microblogging service, where users can, among other things, share short messages of a maximum size of 280 characters. Using Twitter API, it is possible to gather tweets for areas of interest delimited by bounding boxes, where a fraction of them are geotagged (the ones we consider). The main advantage of the Twitter is that its content can have the current location of users (geolocation), which has been used in many applications to predict or detect events in near real-time [15].



(a) Word cloud of the documents matched to UOP-dictionary.

(b) Word clouds of the documents extracted by Algorithm 1

Fig. 2. Word clouds before and after applying the Algorithm 1.

Considering Chicago as our evaluation scenario, we collected tweets, from January to May 2017, to build the LBSN dataset. The dataset is composed of $76,399$ tweets shared in Chicago. We also represent this dataset as a document collection defined as:

Definition 2: Collection $\mathcal{D}_L$, where each $doc \in \mathcal{D}_L$ is a document determined by a tuple $doc = (id, s, \tau, g)$, where id is an unique identifier, s is a list of sentences that comprises all content written by the user, $\tau \in \mathbb{R}$ is a timestamp, and $g \in \mathbb{R} \times \mathbb{R}$ is the geographic coordinates, expressed by latitude and longitude.

When applying the UOP-dictionary in $\mathcal{D}_L$ several documents not related to user perception about urban outdoor areas tend to be retrieved. Considering data from Chicago, Figure 2(a) illustrates this case thought word cloud, where more centralized words and with larger font sizes are the most frequent ones. We can observe that several frequent words are not related to urban outdoor areas, e.g., "careerarc" and "interested". The reason for that is because the qualifiers that compose our dictionary are not restricted to describe places, but some of them can also be used to describe people, things, and so on.

In order to overcome this limitation, a key step is to group documents that have spatial and semantic similarity, disregarding documents non-related to urban areas, or individual perceptions unrelated to urban outdoor areas. To this end, we propose an unsupervised clustering algorithm to group documents with *spatial* and *semantic* similarity present in the data. Algorithm 1 summarize the main steps of our proposal.

The algorithm expects as input: the UOP-dictionary and Word2Vec model for urban areas, WV, described in the previous section; a dataset containing LBSN data, $\mathcal{D}_L$; a value ϵ representing a distance in meters; a value representing a minimum number of points, $minPts$; and two thresholds, $thresh_1$ and $thresh_2$. The first threshold is used to indicate the acceptable maximum number of documents with coincid-

ALGORITHM 1: Algorithm to group documents with spatial and semantic similarity.

input : UOP-dictionary, WV, $\mathcal{D}_L$, ϵ, $minPts$, $thresh_1$, $thresh_2$.
output: Clusters of documents with spatial and significant semantic similarity.

// $clean$ remove all spatial noise from $\mathcal{D}_L$
$\mathcal{D}'_L \leftarrow clean(\mathcal{D}_L, thresh_1)$;
// $getOutdoors$ filter out from $\mathcal{D}'_L$ all documents created in indoor places
$\mathcal{D}''_L \leftarrow getOutdoors(\mathcal{D}'_L)$;
$\mathcal{D}'''_L \leftarrow []$;
foreach $doc \in \mathcal{D}''_L$ **do**
  // Get the list of sentences for each doc
  $sentlist \leftarrow doc.s$
  // Compute the likelihood using the score function in the word2vec.
  $score \leftarrow WV.score(sentlist)$;
  if $\max(|score|) \geq thresh_2$ **then**
    // Sort points by geolocation.
    $\mathcal{D}'''_L.append(doc)$;
  end
end
$\mathcal{C} \leftarrow$ DSCAN with ϵ and $minPts$ in $\mathcal{D}'''_L$;
$\mathcal{P} \leftarrow match(\mathcal{C}$, UOP-dictionary);

ing GPS location, and the next one to indicate the minimum value of the similarity score required to classify the document semantically.

First, the algorithm removes any spatial noise from the dataset, i.e., it filters out from $\mathcal{D}_L$ all documents whose geolocation is the same to other documents if this number is at least $thresh_1$. This step is essential to help to prevent invalid data from being considered, as the probability of this situation happen in practice is low for real data, being more common in automatic processes, such as those used by web robots to perform spam, as discussed in [16]. Based on the dataset $\mathcal{D}_L$, we have determined a threshold $thresh_1 = 10$ to perform this filtering, since most geolocations (about 90%) have less than 10 documents associated with them. This process produces the dataset $\mathcal{D}'_L$.

Next, we filter out from $\mathcal{D}'_L$ documents if their geolocation coincides with any building of the city, generating $\mathcal{D}''_L$. For this, we can use data from the building footprints provided by the city, which are publicly available in Chicago's official open data portal[5]. This step is interesting because help to disregard messages containing opinions not related to outdoor areas, which is our focus in this study.

After that, the algorithm performs a semantic similarity. For each document $doc \in \mathcal{D}''_L$, we explore the Word2Vec model to calculate the likelihood of a doc be a member of a specific class (in our case, the class is urban outdoor perception) [17]. This procedure results in a score, ranging from 0, very unrelated, to 100, very similar, and we explore it to decide if the doc has enough semantic similarity with the model. Based on the dataset $\mathcal{D}''_L$, we have determined a threshold $thresh_2 = 8$ to perform this classification, where most documents (about 75%) have a score fewer than 8. We defined this value by studying the distribution of scores, where was possible to observe a clear change on the curve at the score 8, suggesting that after this point documents tend to be more related to the context of interest. In fact, when evaluating documents with scores less than 8, we start to find documents unrelated to urban outdoor areas. On the other hand, for scores slightly above 8, we practically only find documents related to the urban outdoor context.

[5] Chicago Open Data – https://data.cityofchicago.org.

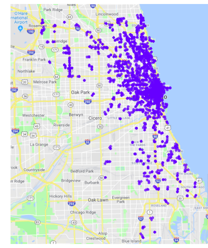
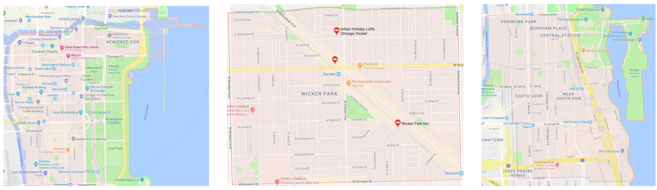


Fig. 3. Spatial distribution of the clusters in $\mathcal{P}$.

Next, documents in $\mathcal{D}'''_L$, representing more semantic similar documents, are clustered according to their spatial similarity, by using Density-Based Spatial Clustering of Application with Noise (DBSCAN) with $\epsilon = 400$ and $minPts = 3$, more details of DBSCAN and its parameters can be found in [18]. We tested with different values of $minPts$, for all scenarios considered, and ϵ did not change significantly. Explore DBSCAN for this task is interesting because it is based on neighborhood density to identify spatial clusters, being able to create clusters with different formats and sizes, and, typically, achieving satisfactory results even in the presence of noises. This clustering process results in a set of clusters $\mathcal{C}$.

Finally, we use UOP-dictionary to label the appropriate category of perception of documents in $\mathcal{C}$. In this way, a given document with at least one word corresponding to the UOP-dictionary can be labeled with one of the seven categories of the dictionary. If two or more words are present in the dictionary, the document is labeled with the corresponding categories, and multiple categories can be assigned to the same document. This helps to reflect the heterogeneity of perception, where the same area may have distinct perceptions according to the opinion of the people. Alternatively, the document may not have words matching with the dictionary, which implies that, despite having semantic content related to the urban area, it is not related to qualifying the area and, therefore, is disregarded. As a result, we have a new set of clusters $\mathcal{P} \subseteq \mathcal{C}$, where the documents are properly labeled with at least one category of UOP-dictionary. Figure 2(b) shows the word clouds for Chicago considering only the documents in $\mathcal{P}$. As we can see, the documents in $\mathcal{P}$ are strongly related to urban outdoor perceptions, being more suitable to be used on perception mapping.

Figure 3 shows the spatial distribution of the clusters in Chicago, where each $doc \in \mathcal{P}$ is represented by a circle with a radius of 50 meters and with its center defined by its GPS location. As we can see, there are several clusters in downtown areas, many overlapped, with sparse clusters scattered around other parts of the city. This is expected behavior because social media users tend to produce and share more content in crowded areas [19].



(a) The Loop (Downtown) (b) Wicker Park (c) Near South Side

Fig. 4. Evaluation scenarios in Chicago.

V. Perceptions of Urban Outdoor Areas

To study our proposed approach, we map the identified perceptions of urban outdoor areas shared on Twitter about some areas of Chicago. Besides, we validate our results by conducting an online survey to gather volunteers' perception about those areas.

A. Studied Areas

In this section, we present and discuss briefly the main characteristics of urban areas evaluated. Due to space limitations, we concentrate our analysis on some areas of Chicago, where there are great concentrations of documents related to urban outdoor perceptions, i.e., documents in $\mathcal{P}$. However, the results discussed for Chicago reflect what is observed in other areas as well.

Figure 4 shows the evaluation scenarios, which are bounded by red lines, for three different regions of Chicago: Downtown, Wicker Park and Near South Side. Downtown Chicago (Figure 4(a)) is known as an important commercial and financial center of the city, attracting crowds of visitors and residents with different profiles to their several types of venues. Also, these areas present a wide variety of sounds, visual elements, odors, among others, which can potentially trigger distinct perceptions in people. For these reasons, heterogeneous urban outdoor perception may be favored to occur. The Wicker Park neighborhood (Figure 4(b)) is known as a hub for shopping, eating, and cultural activities in the city. The Near South Side neighborhood (Figure 4(c)) is located just south of the downtown region, and it consists mostly of residential and leisure areas.

B. Perception Maps

With the help of heat maps, the perception level in the city for each category is highlighted according to the number of collective perceptions observed. Figure 5 shows the heat maps for each perception category separately considering the studied areas in Chicago, where the neighborhood maps are overlapped by a grid of cells with 0.002 latitude by 0.002 longitude resolution. The darker the color, the higher the perception strength (where the maximum level is 10).

To quantitatively evaluate the perception level, we calculate a score that measures the influence of each perception in the studied area. Let R_{points} be the fraction of points of a certain perception in the area (denoted by n_{points}) in relation to the total of points of that perception in the city (denoted by N_{points}). Let R_{cells} be the fraction of cells filled with the same perception in the area (denoted by n_{cells}) in relation to the total of cells in the same area (denoted by N_{cells}). The score is:

$$score = R_{points} \times R_{cells} \tag{1}$$

In this way, $score$ reflects the concentration (R_{points}) and coverage (R_{cells}) of the perceptions in the evaluated areas. Besides being useful to rank perception in a certain area, this metric provides extra information on the expressiveness of a certain category also considering all areas of the city.

For Chicago, the Downtown area is considered predominantly Touristy & Crowded ($score = 0.0318$), Pleasant & Friendly ($score = 0.0257$) and Great & Wonderful ($score = 0.0215$), indicating that Downtown Chicago concentrates some of the main tourist attractions of the city, which might favor good experiences to visitors and citizens. Another significant perception category in this area is Clean & Green ($score = 0.0181$), which is mostly concentrated in Chicago river surrounding areas. The Near South Side neighborhood was represented as predominantly Touristy & Crowded ($score = 0.0016$) and Pleasant & Friendly ($score = 0.0016$), where there are high-intensity brown and purple cells spread by typical pleasant areas, such as beaches, parks and theaters, and, in some cases, with overlaps among them. The categories Clean & Green ($score = 0.0006$), Vibrant & Memorable and Classic & Authentic (both with $score = 0.0005$), can be considered secondary perceptions of the neighborhood, with high-intensity cells in distinctive areas.

Turning our attention to Wicker Park neighborhood region, we can see several overlaps among the perceptions Pleasant & Friendly and Great & Wonderful along "N Milwaukee Avenue" and "W Division Street", mainly near of restaurants and bars. This indicates that, for example, shopping or eating in these areas is a pleasant activity to be performed in Chicago, being the strongest perceptions in the neighborhood ($score = 0.0043$ and $score = 0.0040$, respectively). The perception map displays a significant presence of category Touristy & Crowded ($score = 0.0033$), which indicates that the neighborhood attracts a large number of citizens and visitors. However, crowded places may trigger some problems, such as traffic jams, lack of places for parking, long lines, etc, factors that can contribute to offer bad experiences for some people, which may justify the category Awful & Scary to have a $score$ higher than Clean & Safe category (0.0013 against 0.0009).

C. User Assessment

We have conducted an online survey from April 4^{th} to April 20^{th}, 2018, in order to collect the perception of volunteers about urban outdoor areas assessed in the prior section. Our goal with this survey was getting knowledge about the perception of the studied areas to validate our results.

The survey contains three figures of the neighborhoods considered, without our results, as shown in Figure 4, and we

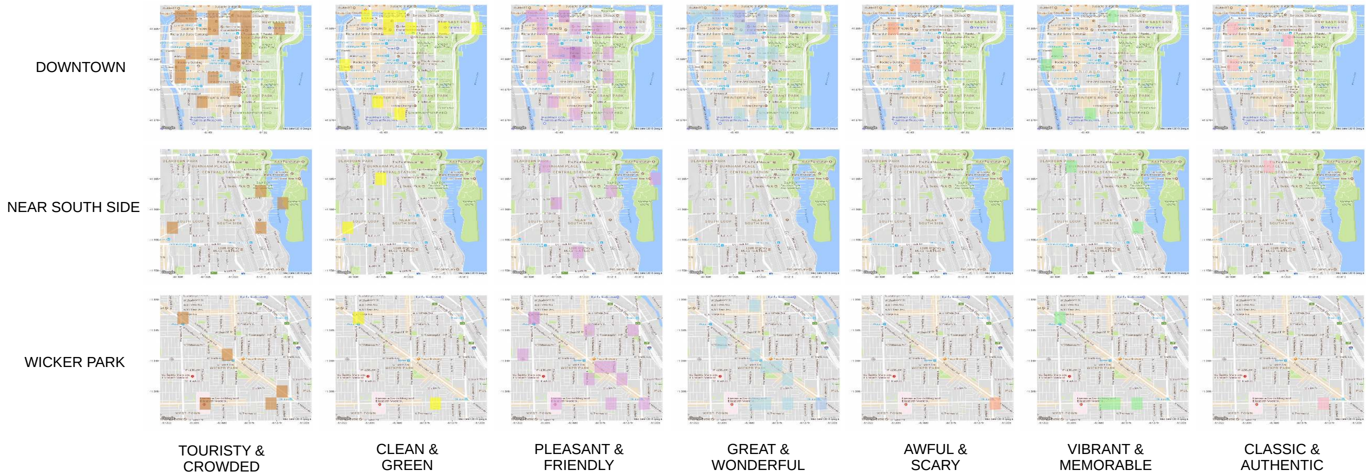


Fig. 5. Perception maps in some Chicago's neighborhoods.

asked the participants to choose one or more set(s) of words, if there was, which better characterized the urban outdoor areas displayed in the figures. The sets of words are the same that comprise the perception categories of UOP-dictionary (see Table II). We also have included two extra sets of qualifying words of urban outdoor areas, which not belong to the UOP-dictionary, to help to test our dictionary. Besides, we offered an empty option to not impose the choice of at least one of the provided sets. For each figure, we also left an option for participants express, via free-text, their thoughts about the area under study. Moreover, we asked basic information about the participants, such as their highest degree of education completed, gender, age range, and the level of knowledge about urban areas of the city.

In total, we recruited 51 volunteers for Chicago, where the majority, 26 participants, knows a lot about of the city, including several less popular ones (high knowledge), 20 participants declared to know the main areas and few less popular ones (medium knowledge), and other 5 just know few popular areas (low knowledge). Considering all participants from Chicago, 23 are women and 28 men, and most of them are adult (16 with age range from 31 to 40, 8 ranging from 41 to 50, 7 with age upper than 50, and the remaining up to 30 years old), and with high education background (21 have bachelor's degree, 11 have master's degree, 2 have Ph.D. or advanced graduate work, and 14 have some college).

For the participants, the Downtown Chicago was the scenario with the higher heterogeneity of perception. According to their opinion, this region is mainly Touristy & Crowded (in the opinion of 62.7% participants), but also is Vibrant & Memorable and Great & Wonderful (52.9% and 39.2% of the opinions, respectively). The similarity between the perception identified by our approach with this result is striking, where the most intense perception also is Touristy & Crowded, followed by Pleasant & Friendly and Great & Wonderful. Despite divergence on second most intense perception, the overall picture is still well captured.

The perception about the region of Near South Side neighborhood, in the opinion of participants, is predominantly Classic & Authentic and Touristy & Crowded (35.3% and 33.3%, respectively), followed by Vibrant & Memorable (29.4%). Note that the participants' opinions and the maps generated for the same area have a very good match. According to our results, Touristy & Crowded category corresponds to a primary perception of the neighborhood, while the Vibrant & Memorable and Classic & Authentic are secondary categories. Although Pleasant & Friendly category is also a primary perception according to our algorithm for Near South Side, and it is not among the most selected by the participants, we can see in several free-text shared by them the use of similar or equal qualifiers of this category (indicated in bold), which reinforces the relevance of such category for this neighborhood:

> *"I **love** this area, I used to live there. It's very **quiet** and **friendly**."*
> *"**Nice**, **quite** area, up and coming."*
> *"Similar to the loop, this is a busy and somewhat touristy area. Lots of **cool** things to do and see here, and **good** classic restaurants."*

The Wicker Park neighborhood according to the opinion of the participants is mostly Classic & Authentic (47.1%) and Pleasant & Friendly (39.2%), followed by Vibrant & Memorable (23.5%) and Touristy & Crowded (19.6%). Our map also identified a very similar perception for this area, being Pleasant & Friendly a primary perception and Touristy & Crowded a secondary one. For participants, this neighborhood also is more Awful & Scary than Clean & Safe (11.8% against 2%) as identified by our algorithm. Although some slight differences, the results show a strong correlation between people's perception and the results obtained by our approach.

VI. Discussions and Limitations

The overall perception captured by our approach in the neighborhoods evaluated seems to surprisingly correct represent the perception of people who know Chicago. Small divergences between our results and the volunteers' opinions are expected because our algorithm represents the perception of many people, and, consequently, the perception of the minority might be misrepresented [9].

Our work has potential to help users to extract knowledge from the city, and thus, to improve their understanding of it, helping them to explore urban areas better, for instance, to identify regions with certain perceptions. In addition, it enables the emergence of new intelligent services, such as personalized route recommendations, which could be offered by systems like Waze or Google Maps. For example, tourists can explicitly request routes to walk through areas with certain perceptions. Even if the route is a bit longer than the shortest one, they might be willing to pay this cost.

It is important to note also that our approach could be applied, with the proper adaptations, in other domains as well. In this study, we focused on urban outdoor areas, but our framework could be explored, for example, for indoor areas. Due to its flexibility, our framework easily enables extensions. For instance, by employing our methodology in other user review datasets, perhaps bigger and more updated from those used in this study, we could potentially enrich our dictionary with new significant words, since the language is in constant evolution and can change differently in different locations.

We are also aware of possible some limitations of our proposal. Analyzing samples of the messages shared by users, we found certain phrases containing conflicting perceptions. For example, "... *It is amazing, and dangerous* ..." and "... *these things we build - rusted, broken, gleaming, dingy, dirty, dry, wet, aged, metal, oxidized, plastic, ugly, dull, and flawed though they may be, are beautiful* ...". In those examples, the categories Awful & Scary and Great & Wonderful would be identified. If this type of situation happens many times in the same area, this could favor identifying wrongly two types of perception, depending on the values for the parameters set in the frameworks. We did not find any relevant problems regarding these cases in our results. However, identify and treat this type of phrases beforehand could help to improve our work. Besides, our maps do not take into account the time dimension, i.e., when the messages were shared. We hypothesize that the time might influence some of the results, including in the opinions of the volunteers.

VII. Conclusions

In this study, we present an approach to support the learning and mapping of the perception of urban outdoor areas from a large collection of noisy data expressing users' opinions in LBSNs. The proposed approach is generic enough to be applied in different contexts with small adjustments. Our results suggest that it is possible to identify the perception reflected in urban areas in a scalable way, and, thus, support important mechanisms to help people better understand the semantics existing in different areas of the city. As a future work, we intend to evolve this approach to incorporate other data sources, such as Instagram and Facebook, apply and evaluate our strategy for content in other languages, and analyze more details about the implications of the time dimension on the results.

Acknowledgment

The authors would like to thank the grants Sao Paulo Research Foundation (FAPESP 2015/24949-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (PDSE 88881.132016/2016-01) and CNPq-URBCOMP project (grant 403260/2016-7) by the financial support.

References

[1] V. Henshaw, *Urban smellscapes: Understanding and designing city smell environments*. Routledge, 2013.

[2] D. Quercia, R. Schifanella, L. M. Aiello, and K. McLean, "Smelly maps: The digital life of urban smellscapes," in *Proc. of ICWSM*, Oxford, United Kingdom, 2015.

[3] T. Silva, P. Vaz De Melo, J. Almeida, and A. Loureiro, "Large-scale study of city dynamics and urban social behavior using participatory sensing," *Wireless Communications, IEEE*, vol. 21, no. 1, pp. 42–51, Feb 2014.

[4] L. M. Aiello, R. Schifanella, D. Quercia, and F. Aletta, "Chatty maps: constructing sound maps of urban areas from social media data," *Open Science*, vol. 3, no. 3, p. 150690, 2016.

[5] A. Dubey, N. Naik, D. Parikh, R. Raskar, and C. A. Hidalgo, "Deep learning the city: Quantifying urban perception at a global scale," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Amsterdam, The Netherlands: Springer International Publishing, 2016, pp. 196–212.

[6] D. Quercia, R. Schifanella, and L. M. Aiello, "The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city," in *Proc. of HyperText*. Santiago, Chile: ACM, 2014, pp. 116–125.

[7] W. Mueller, T. H. Silva, J. M. Almeida, and A. A. Loureiro, "Gender matters! analyzing global cultural gender preferences for venues using social sensing," *EPJ Data Science*, vol. 6, no. 1, p. 5, 2017.

[8] T. H. Silva, P. O. V. de Melo, J. M. Almeida, M. Musolesi, and A. A. Loureiro, "You are what you eat (and drink): Identifying cultural boundaries by analyzing food and drink habits in foursquare." in *ICWSM*, 2014.

[9] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh, "The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City," in *Proc. of ICWSM*, Dublin, Ireland, 2012.

[10] D. Quercia, N. K. O'Hare, and H. Cramer, "Aesthetic capital: what makes london look beautiful, quiet, and happy?" in *Proc. of CSCW*. Baltimore, United States: ACM, 2014, pp. 945–955.

[11] N. Naik, J. Philipoom, R. Raskar, and C. Hidalgo, "Streetscore-predicting the perceived safety of one million streetscapes," in *Proc. of CVPR Workshops*, Columbus, United States, 2014, pp. 779–785.

[12] L. Porzi, S. Rota Bulò, B. Lepri, and E. Ricci, "Predicting and understanding urban perception with convolutional neural networks," in *Proc. of MM*. Brisbane, Australia: ACM, 2015, pp. 139–148.

[13] D. Quercia, L. M. Aiello, and R. Schifanella, "The emotional and chromatic layers of urban smells." in *Proc. of ICWSM*, Cologne, Germany, 2016, pp. 309–318.

[14] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. of NIPS*, Lake Tahoe, United States, 2013, pp. 3111–3119.

[15] E. Kalampokis, E. Tambouris, and K. Tarabanis, "Understanding the predictive power of social media," *Internet Research*, vol. 23, no. 5, pp. 544–559, 2013.

[16] D. Tasse, Z. Liu, A. Sciuto, and J. I. Hong, "State of the geotags: Motivations and recent changes." in *Proc. of ICWSM*, Montreal, Canada, 2017, pp. 250–259.

[17] M. Taddy, "Document classification by inversion of distributed language representations," in *Proc. of ACL*, vol. 2, Beijing, China, 2015, pp. 45–49.

[18] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Proc. of KDD*, Portland, United States, 1996, pp. 226–231.

[19] D. Flatow, M. Naaman, K. E. Xie, Y. Volkovich, and Y. Kanza, "On the accuracy of hyper-local geotagging of social media content," in *Proc. of WSDM*. Shanghai, China: ACM, 2015, pp. 127–136.