# A Robust Noise Resistant Algorithm for POI Identification from Flickr Data*

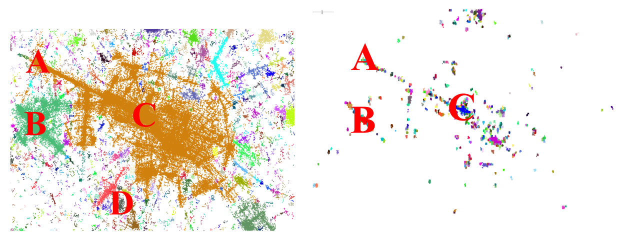
**Yiyang Yang⋆ Zhiguo Gong§† Qing Li¶ Leong Hou U§ Ruichu Cai⋆ Zhifeng Hao‡**
⋆Faculty of Computer, Guangdong University of Technology, Guangzhou, China
§Department of Computer and Information Science, University of Macau, Macau SAR
¶Department of Computer Science, City University of Hong Kong, Hong Kong SAR
‡School of Mathematics and Big Data, Foshan University, Foshan, China

## Abstract

Point of Interests (POI) identification using social media data (e.g. Flickr, Microblog) is one of the most popular research topics in recent years. However, there exist large amounts of noises (POI irrelevant data) in such crowd-contributed collections. Traditional solutions to this problem is to set a global density threshold and remove the data point as noise if its density is lower than the threshold. However, the density values vary significantly among POIs. As the result, some POIs with relatively lower density could not be identified. To solve the problem, we propose a technique based on the local drastic changes of the data density. First we define the local maxima of the density function as the Urban POIs, and the gradient ascent algorithm is exploited to assign data points into different clusters. To remove noises, we incorporate the Laplacian Zero-Crossing points along the gradient ascent process as the boundaries of the POI. Points located outside the POI region are regarded as noises. Then the technique is extended into the geographical and textual joint space so that it can make use of the heterogeneous features of social media. The experimental results show the significance of the proposed approach in removing noises.

## 1 Introduction

Flickr contains more than 8 billions photos from 8.7 million users; in addition, 3.5 million new photos are uploaded to Flickr daily where a substantial number of these photos come from the mobile devices[1]. Such massive amounts of photos with heterogeneous meta-data (geographical and textual tags) are valuable resources to support various mining tasks and have resulted in many research problems, such as Point of Interest (POI) identification [Yang *et al.*, 2014], the POI-based applications [Crandall *et al.*, 2009], photo tag analysis [Zhang *et al.*, 2012] and travel pattern recognition [Zheng *et al.*, 2012].




(a) Low Noise Filtering (b) High Noise Filtering

Figure 1: Clustering results on Paris dataset, the POIs in alphabet order are: (A) `Arc De Triomphe`, (B) `Eiffel Tower`, (C) `Louvre Museum`, (D) `Montparnasse Cemetery`

According to our survey, plenty of travel recommendation works [Cheng *et al.*, 2013; Ying *et al.*, 2013; Popescu and Shabou, 2013; Ying *et al.*, 2014] require a predefined POI database as the input to their recommendation algorithms. It is no doubt that the quality of the POI database is critical to the success of their subsequent processes. In this paper, we propose a robust noise-resistant approach for POI identification using geo- and textual-tagged social media data.

In general, the POI database can be automatically constructed by performing some clustering algorithms over the geo-tagged datasets (i.e., Flickr photos). However, the crowd-contributed data like Flickr often contain large amounts of noises which may generate pernicious effects to the quality of the identified POI. A robust algorithm for POI identification is definitely desired from POI-based applications.

### 1.1 Motivations & Contributions

When visiting a city, people may take photos anywhere, and a large percent of which are POI irrelevant. In this paper, such POI-irrelevant geo-tagged photos are regarded

*We thank the anonymous reviewers for their many insightful comments and suggestions. This work was supported in part by: NSFC-Guangdong JF (U1501254), NSFC (61603101, 61472089, 61472337), NSF-Guangdong (2014A030306004, 2014A030308008), STPP of Guangdong (2015B010108006, 2015B010131015), FDCT of Macau Government (FDCT/116/2013/A3, FDCT/007/2016/AFJ), UMAC RC (MYRG2015-00070-FST, MYRG2017-00212-FST, MYRG2016-00182-FST), GuangdongHPSSP(2015TQ01X140), TPP of Guangzhou (201610010101, 201604016075)

†Corresponding Author: fstzgg@umac.mo

[1] https://en.wikipedia.org/wiki/Flickr

as noises, because they can cause the identified POIs to be significantly deviated from the actual ones. To solve the issue, some techniques are proposed[Ester *et al.*, 1996; Hinneburg and Keim, 1998; Zhao *et al.*, 2009; Purwar and Singh, 2016] , where a density threshold is used to filter out the lower density points as noises. However, such a global setting cannot handle the noise problem well because of the various density values among POIs. Figure 1 shows the situations of the DBScan algorithm[Ester *et al.*, 1996] with two extreme settings of the density threhold. Figure 1(a) is the result of a lower filtering threshold where two urban POIs, (A) `Arc De Triomphe` and (C) `Louvre Museum` in the downtown area are merged together. At the same time, a large amount of irrelevant data points (geo-tagged photos) are included in the result as well. In contrast, if we overly increase the threshold (Figure 1(b)), then it will partition one POI region into many small clusters and at the same time some important POIs (e.g. (D) `Montparnasse Cemetery`) may be omitted just because of their relatively lower density. In this paper, we aim at tackling the problem.

To handle the noise issue, we propose a novel approach based on the Laplacian Zero-Crossing Technique [Canny, 1986]. It is constructed based on the relative change of the geographical density. This idea is motivated by an **observation** that the density of geo-tagged photos encounters a drastic change when crossing the POI boundaries. The Laplacian of the density function indicates the speedup of the gradient ascending and it changes from positive to negative when crossing the boundaries of the POI region (Zero-Crossing of the Laplacian). We attempt to capture such "changes" automatically in the identification, thus thoroughly overcome the problem caused by the global density threshold setting.

Flickr photos are not only tagged with geographical locations, but also texts. The quality of the identified POI can be further improved if the textual features are taken into account. For such a sake, we apply Local Sensitive Hashing (LSH) algorithm [Charikar, 2002] to transform a term vector into a hashing value and exploit Hamming Distance [Manku *et al.*, 2007] to measure the difference between two photos in the textual space. This technique can benefit our algorithm from two aspectives: (1) it significantly reduces the computational cost for the distance evaluation between textual tags (2) with defined distance metric, it enables to extend Laplacian Zero-Crossing algorithm into the Geographical $\times$ Textual joint space.

The contributions of this work are summarized as follows:

1. We propose a novel technique, **Laplacian Zero-Crossing Detecting**, to remove noisy data (POI irrelevant), along the gradient ascent process. The technique is based on the drastic change of the local density increase, thus can thoroughly overcome the global setting problem of the density threshold.
2. We extend the proposed techniques into the joint space crossing both geographical and textual features, which can further improve the quality of POI identification.
3. Experiments on real datasets collected from Flickr demonstrate the effectiveness of our algorithm.

### 1.2 Organization of the Paper

The reminder of this paper is organized as follows. Section 2 and Section 3 respectively introduce the Hill-Climbing algorithm and Laplacian Technology. In Section 4, we describe how to integrate the textual feature into the framework. We thoroughly evaluate our proposed techniques on the photo collections of Flickr in Section 5. We discuss the related work in Section 6 and summarize our work in Section 7.

## 2 POI Identification Using Gradient Ascent Algorithm

Intuitively, a location is attractive if visitors take more photos around it. In other words, the location with the highest density of geo-tagged photos among its neighbors is probably a POI. Formally, given a set of geo-tagged photos $x_1, \ldots, x_N$ in the 2-dimensional Euclidean space $R^2$, the density function of a location $x$ can be estimated through a Kernel $K$ as follows [Cheng, 1995]:

$$T_g(x) = \sum_{i=1}^{N} K(\frac{x_i - x}{g}) \quad (1)$$

where $g$ is the geographical bandwidth parameter (e.g. 100 meters). In this paper, Gaussian Kernel $G$ is exploited for its excellent features (i.e. simple and infinite differentiable, easy gradient estimation). By replacing $K$ by $G$ in Equation 1, the normalized gradient of $T_g$ can be computed as:

$$\nabla \ln T_g(x) = \frac{\nabla T_g(x)}{T_g(x)} = \frac{2}{g^2}(\frac{\sum_{i=1}^{N} G(\frac{x_i - x}{g})x_i}{\sum_{i=1}^{N} G(\frac{x_i - x}{g})} - x) \quad (2)$$

Instead of $\nabla T_g(x)$, we use $\nabla \ln T_g(x)$ in this work, for its fast converging speed and easy estimation [Comaniciu and Meer, 2002].

In our scenario, the city POIs are located at places $x$ with $\nabla \ln T_g(x) = 0$. To look for all POIs in a city, we start the gradient ascent algorithm with any photo $x^0$, then iteratively conduct the following Hill-Climbing Algorithm:

$$x^{j+1} = x^j + \alpha \nabla \ln T_g(x^j), \;\; j = 0, 1, 2...m \quad (3)$$

where $\alpha$ is a parameter for controlling the size of the movement, $x^j$ is the $j^{th}$ state of the gradient ascent movement starting from $x^0$, and $x^j$ will converge to $x^m$ with $T_g(x^j)$ monotonically increasing until $\nabla \ln T_g(x^m) = 0$. Hence, the $x^m$ is a stationary point and is called the location of a POI, while the sequence of successive states $x^j, j = 0, ..., m$ is called the **trajectory** of $x$.

According to the Capture Theorem, the trajectory is attracted by a local maximum if the $x^m$ is close enough to it, the trajectory will eventually converge to a local maximum which serves as the urban POI. Thereby, Hill-Climbing approach normally terminates if the shift distance is smaller than a predefined threshold $\delta_g$ or the number of iterations exceeds the corresponding constraint. Photos (or points) of converging to the same local maximum are assigned into the same POI.

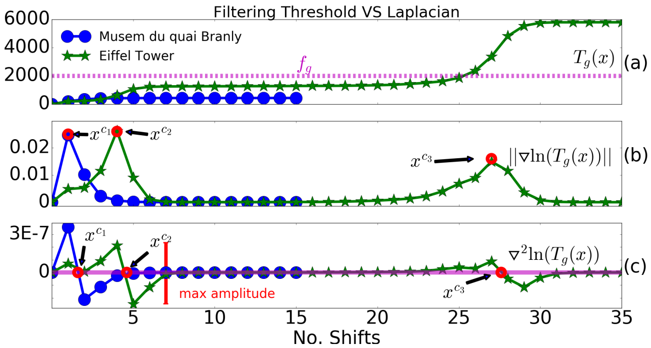



Figure 2: Trajectories to `Eiffel Tower` and `Musée du quai Branly` respectively: (a) Geographical Density $T_g(x)$ (b) The Magnitude of Geographical Shift Vector $||\nabla \ln T_g(x)||$ (c) Geographical Laplacian $\nabla^2 \ln T_g(x)$

As discussed in the Introduction Section, the POIs may include a large percent of noise photos, which significantly bias the identified results. To remove the noises, works [Ester *et al.*, 1996][Hinneburg and Keim, 1998] set a threshold $f_g$, such that all photos with $T_g(x) < f_g$ are discarded as noises. This technique is extensively employed in the Hill-Climbing approach[Cheng, 1995; Comaniciu and Meer, 2002; Crandall *et al.*, 2009]. However, it is difficult to set a right value for $f_g$ as discussed in the Introduction Section. In next subsection, we propose a method for adaptively detecting noises based on the change of the local density.

## 3 Laplacian Technique for Noise Reduction

To understand the principles of our proposed technique intutively, in Figure 2 we plot the values of $T_g(x^j)$, $||\nabla \ln T_g(x^j)||$ via the shift movement iteration $x^j$ ($x$-axis ) for two trajectories which converge to POI `Eiffel Tower` and `Musée du quai Branly`, respectively.

As shown in Figure 2 $(a)$, the density of `Eiffel Tower` is much larger than that of `Musée du quai Branly`. If a global density threshold is used to remove the noise photos as in [Ester *et al.*, 1996; Hinneburg and Keim, 1998], it is probably that the less popular POI (`Musée du quai Branly`) will be missed since its density (e.g. 500) is under the filtering threshold $f_g$ (e.g. 2000). Generally speaking, landmarks located in downtown areas often have a much larger density value than those ones located in the marginal areas. To solve the problem, we propose to adaptively filter the noises according to the intensive change of the density along the trajectories in the Hill-Climbing process.

In a Hill-Climbing process, the magnitude of $\nabla \ln T_g(x^j)$ indicates the increase rate of $T_g(x)$ at the $j$-th step. Normally, $T_g(x^j)$ is monotonically increasing until $x^j$ converges to the POI. It is intuitive that the gradient ascent movement will encounter a steepest ascent when passing the boundary of the POI. We define the **Influence Region** $IR(x^m)$ of a urban POI $x^m$ as an area within which the $x^m$ and its neighbor photos share similar densities. Thereby during the Hill-Climbing process $T_g(x^j)$ changes steadily if it locates inside the $IR(x^m)$, while it changes drastically when it is **shifting** into the region. We attempt to capture such changes as the critical points to separate noise and non-noise photos.

According to our observation, two types of trajectories are distinguished: (1) The trajectory converges to a POI steadily. (2) The trajectory passes one or more drastic changes in terms of the density increase and finally reaches a POI.

In the former situation, the Hill-Climbing is a steady process because the photo and its whole trajectory is located inside the "region" of the POI. While in the latter case, the rapid intensive change of $T_g(x^j)$ could be interpreted as the boundary of the POI. As shown in Figure 2, each intensive change (e.g. $x^{c1}$, $x^{c2}$ and $x^{c3}$) is probably a boundary towards POI and separates (classify) the significant and relatively less significant photos in terms of the geographical density. In other words, the boundary-crossing can be utilized to automatically distinguish the POI relevant photos from noisy ones.

Determining such kind of boundaries is a problem that aims to find the critical point $x^c$ which locates in the boundary between sparse area (i.e., steady change of $T_g(x^j)$) and dense area (i.e., intensive change of $T_g(x^j)$) in a trajectory. For the first type of trajectories, there is no critical point because of the steady climbing process, which means it is unnecessary to perform the noise filtering. Regarding the second type, only part of the trajectory is included in the influence region while the critical point services as the separation between POI relevant photos and noisy ones.

### 3.1 Zero-Crossing Detecting of Laplacian

To detect such critical point(s) in a trajectory, we propose to exploit the Laplacian Zero-Crossing of the logarithm-normalized density function $\ln T_g(x)$. This idea is motivated by the edge detection problem in image processing [Canny, 1986]. Formally, the Laplacian operator of $\ln T_g(x)$ is defined as:

$$\nabla^2 \ln T_g(x) = \frac{4}{g^2} \frac{\sum_{i=1}^{N} G(\frac{x_i - x}{g})(x_i - x)^2}{T_g(x)} - (\frac{2}{g^2} + ||\nabla \ln T_g(x)||^2) \quad (4)$$

Therefore, the Laplacian operator of $\ln T_g(x)$ can be easily computed by reusing the value of $T_g(x)$ and $||\nabla \ln T_g(x)||$. To apply the operator to the boundary detecting in the Hill-Climbing algorithm, we compare the sign of $\nabla^2 \ln T_g(x^j)$ with that of $\nabla^2 \ln T_g(x^{j+1})$. The gradient movement is said to cross the boundaries of a POI if we have:

$$(\nabla^2 \ln T_g(x^j) \geq 0) \wedge (\nabla^2 \ln T_g(x^{j+1}) < 0) \quad (5)$$

In case that two or more points, say $x^{j_1}, x^{j_2}, ..., x^{j_K}$, satisfy the Laplacian Zero-Crossing along the Hill-Climbing trajectory, we define the boundary point as the one $x^{j_k}$ which makes the value of $||\nabla^2 \ln T_g(x^{j_k}) - \nabla^2 \ln T_g(x^{j_k+1})||$ (amplitude) maximum.

In Figure 2 $(b)$ and $(c)$, we plot the value of $||\nabla \ln T_g(x)||$ and that of $\nabla^2 \ln T_g(x)$ via the Hill-Climbing shift, respectively. Point $x^{c_1}$, $x^{c_2}$, and $x^{c_3}$ are detected as the Laplacian zero-crossing points of $\nabla^2 \ln T_g(x)$, which corresponds to the local maximum points of $||\nabla \ln T_g(x)||$, respectively.

**Algorithm 1** Detection Influence Region By Laplacian

**Input:** POI set $P$ with trajectories.
**Output:** POI set $P^*$ with pruned trajectories
1: **for** each POI $p$ in $P$ **do**
2: **for** each trajectory $t$ of $p$ **do**
3: compute normalized Laplacian $\nabla^2 \ln T_g(x)'$ by Equation 4
4: compute the zero-crossing points of normalized Laplacian by Equation 5
5: **if** there exists one or more zero-crossing points **then**
6: find the zero-crossing point with maximum amplitude $x^{c_{max}}$
7: prune the trajectory $t$ from $x^{c_{max}}$ to the end as $t^*$ and add $t^*$ as trajectory of $p$
8: **else**
9: add $t$ as trajectory of $p$ without pruning

Algorithm 1 demonstrates how to find the Influence Region by the Laplacian. The normalized Laplacian is computed (line 3), according to which the zero-crossing points are detected. In Line 4 the critical point with maximum amplitude is retained as the boundary point since it indicates the most drastic change in density increment. An example is given in Figure 2 $(c)$, the trajectory to `Eiffel Tower` has two critical points $x^{c_2}$ and $x^{c_3}$, and the $x^{c_2}$ with maximum amplitude is selected to tune the trajectory (Line 6).

## 4 Extending the Techniques into the Joint Space

Besides the geographical feature, Flickr photos are tagged with rich texts too. Recently, works [Bui *et al.*, 2015; Vu and Shin, 2015] attempt to identify POI boundaries using texts, in which the POI boundaries are determined by the drastic change of the text similarity between the detecting point and the POI center. However, both works require that the POI database, including the POI center (exact location) and its textual description (extracted from Wikipedia), should be available in advance. In this work, we extend the proposed Laplacian Zero-Crossing technique to the joint space of geographical and textual dimensions. This idea is motivated by the observation that the photos (or other geo-textual-tagged social media like Tweeter) are often described using similar texts within the same region of a POI. These two features together can further improve the quality of the identified POIs.

In the joint space of $G \times S$ ($G$: the geographical space; $S$: the textual space), let $x = (x_{:g}, x_{:s})$ denote the point in the space. Similar to works[Comaniciu and Meer, 2002; Yilmaz, 2007], we define the Multi-Variate Kernel in the joint space as:

$$G_{g,s}(x_i - x) = G_g(\frac{x_{i:g} - x_{:g}}{g})G_s(\frac{x_{i:s} - x_{:s}}{s}) \quad (6)$$

where $g$ and $s$ are the bandwidth parameters in the two spaces respectively, $x_i = (x_{i:g}, x_{i:s})$ and $x = (x_{:g}, x_{:s})$ are the joint vectors of the geographical and textual features.

A joint vector $x = (x_{:g}, x_{:s})$ is of length $2 + d$, where the first two dimensions consist of the latitude and longitude of the geo-tagged photos and the last $d$ dimensions represent the total number of terms in the vocabulary. The $x_{:s}$ records the $TF \cdot IDF$ [2] values of the tags in $x$. So, the textual feature of a photo is represented as a $d$-dimensional vector of terms. The distance of two text vectors $x_{i:s} - x_{:s}$ is defined as the Hamming Distance. Before computing the Hamming Distance, it is necessary to convert the text vector into a Local Sensitive Hash (LSH) [Charikar, 2002; Manku *et al.*, 2007] string: a binary hashing value of length $l$ ($l \ll d$). Hamming Distance between the hashing values gives the approximative textual difference between the respective vectors (photos). LSH technique is adopted for several reasons: (1) the cost to textual distance computation is reduced: only $l$ bits comparison; (2) LSH considers the significance (weight) of the textual tag as well; and (3) the time of neighborhood search in textual space can be further reduced if the corresponding indexing technology is employed.

With the Multi-Variant Kernel (c.f. Equation 6), the density function in the joint space is defined as:

$$T_{g,s}(x) = \sum_{i=1}^{N} G_{g,s}(x_i - x) \quad (7)$$

With the Equation 7, we correspondingly have the normalized gradient in the joint space as:

$$\nabla \ln T_{g,s}(x) = (\nabla_{x_g} \ln T_{g,s}(x), \nabla_{x_s} \ln T_{g,s}(x)) \quad (8)$$

where $\nabla_{x_g} \ln T_{g,s}(x)$ is derived as:

$$\frac{2}{g^2}\left(\frac{\sum_{i=1}^{N} G_g(\frac{x_{i:g} - x_{:g}}{g})G_s(\frac{x_{i:s} - x_{:s}}{s})x_{i:g}}{\sum_{i=1}^{N} G_g(\frac{x_{i:g} - x_g}{g})G_s(\frac{x_{i:s} - x_{:s}}{s})} - x_{:g}\right) \quad (9)$$

and $\nabla_{x_s} \ln T_{g,s}(x)$ is derived as:

$$\frac{2}{s^2}\left(\frac{\sum_{i=1}^{N} G_g(\frac{x_{i:g} - x_{:g}}{g})G_s(\frac{x_{i:s} - x_{:s}}{s})x_{s:i}}{\sum_{i=1}^{N} G_g(\frac{x_{i:g} - x_g}{g})G_s(\frac{x_{i:s} - x_{:s}}{s})} - x_{:s}\right) \quad (10)$$

The Laplacian of the joint density function can be easily computed as:

$$\nabla^2 \ln T_{g,s}(x) = \nabla^2_{x_g} \ln T_{g,s}(x) + \nabla^2_{x_s} \ln T_{g,s}(x) \quad (11)$$

With the equations above, we can re-write the Hill-Climbing and Laplacian Zero-Crossing algorithms in the joint space. Geo-tagged photos are assigned into the same POI cluster if both of their geographical and textual features converge consistently. In Figure 3, we plot the density Laplacian of the Trajectory to `Eiffel Tower` in the Joint Space. Besides, in Figure 3 (b), the $Top$ 6 $TF \cdot IDF$ tags of `Eiffel Tower` are listed: if the feature from the textual space is taken into account, we will probably select $x^{c3}$ as the critical point of the trajectory since it reflects the density change in the joint space. Further, as we will show in our experiments, the Laplacian Zero-Crossing technique is more robust resistant to the noises.

[2] $TF$ and $IDF$ follow the standard definitions as in information retrieval. $TF$ denotes the frequency of a term in a document. $IDF$ denotes the Inverse Document Frequency of the term.

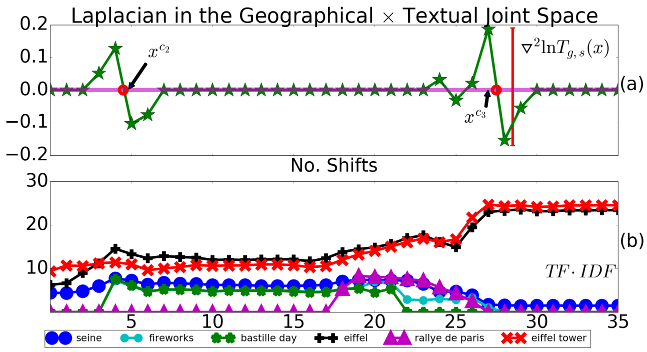



Figure 3: Trajectory to `Eiffel Tower` in the Geographical × Textual Joint Space: (a) Joint Laplacian $\nabla^2 \ln T_{g,s}(x)$ (b) Top 6 Textual Tags in terms of the $TF \cdot IDF$ value

Table 1: Photos collections and MBRs dataset

| City | No. photos<br>No. points | Latitude and<br>Longitude | No.<br>MBRs |
|---|---|---|---|
| **Paris** | 796,427<br>216,300 | $\begin{bmatrix} 48.815 & 48.903 \\ 2.223 & 2.474 \end{bmatrix}$ | 989 |
| **New York** | 1,067,964<br>274,428 | $\begin{bmatrix} 40.499 & 40.931 \\ -74.258 & -73.713 \end{bmatrix}$ | 492 |
| **Rome** | 343,917<br>96,738 | $\begin{bmatrix} 41.790 & 41.989 \\ 12.368 & 12.624 \end{bmatrix}$ | 254 |

## 5 Evaluation

In this section, we extensively evaluate our proposed techniques using three photo datasets collected from Flickr. We use the Flickr API[3] to collect the photos (including geographical locations, textual tags, and taken time) of a city. Specifically, we use the name of a city as the search query and keep only relevant photos (i.e., the geo-location is inside of the boundary of city) in our collection. The statistic of our dataset collections is listed in Table 1. We select three representative cities: New York, Paris and Rome. It should be noticed that the number of photos is much more than the number of points, because several photos might be taken at a single point.

In order to evaluate the identified Influence Region (IR), we use the POI information extracted from OpenStreetMap[4] as the ground truth: the POIs are retrieved if labeled with "tourism"[5]. For each POI in the ground truth, we use its **Minimum Bounding Rectangle** (MBR) as the actual region. The number of MBRs is shown in fourth column of the Table 1.

The objective of evalution is to measure the similarity between identified Influence Region (IR) and the ground truth MBR. To achieve it, it is necessary to define metrics: (1) $F_1$ score: $P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN}, F_1 = \frac{P*R*2}{P+R}$ where $P$ and $R$ denotes the precision and recall respectively. (2) Normalized Mutual Information (NMI): $NMI(MBR, IR) = \frac{I(MBR;IR)}{[H(MBR)+H(IR)]/2}$, where $I$ is the mutual information and $H$ is the entropy. Readers may refer to [Manning *et al.*, 2008] for the detail of the metrics. For both metrics, a higher value indicates the better identification quality.

In this paper, several baseline approaches are evaluated in our experiments ($G$ : geographical space, $S$ : textual space):

1. $DBScan$[Ester *et al.*, 1996] is evaluated as well since it is widely used in geographic based mining tasks[Ester *et al.*, 1996; Zheng *et al.*, 2009; 2012; Ying *et al.*, 2014; Purwar and Singh, 2016].
2. $HC$ refers to the Mean Shift approach[Cheng, 1995; Hinneburg and Keim, 1998; Crandall *et al.*, 2009; Zhao *et al.*, 2009]. $HC_g$ (Equation 2) and $HC_{g,s}$ (Equation 8) refer to the Mean Shift in the $G$ space and the $G \times S$ joint space respectively.
3. $Laplacian$ refers to our proposed noise resistent framework. It is feature independent where $Laplacian_g$ and $Laplacian_{g,s}$ denote its applications to the $G$ space and the $G \times S$ joint space respectively.

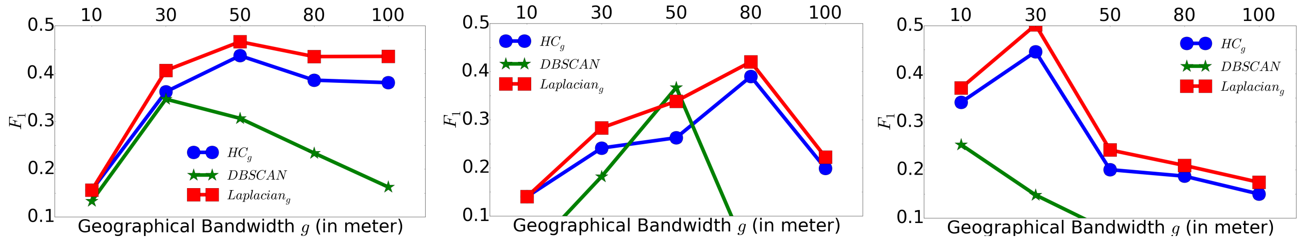
For all approaches, we select five different values for the geographical setting: $10, 30, 50, 80, 100$ ($g$ in meter). For the $G \times S$ joint space based approaches $HC_{g,s}$ and $Laplacian_{g,s}$, since we adopt Local Sensitive Hash (LSH) to accelerate the textual distance computation, the length of Hash value ($l$) is set to 128, while the textual bandwidth ($s$ in Hamming Distance) varies among $30, 50, 60, 80, 100$. The global threshold (for $min\_phts$ in $DBScan$ and $f_g$ in $HC$) are set to 50 through our preliminary experiment.

In our implementation, R*tree is used as the index and costs $O(logN)$ for each geographical range query where $N$ is the number of points, thus the complexity of DBScan is $O(NlogN)$. The complexity of Hill-Climbing is $O(NMlogN)$ where $M$ is the number of average steps (shifts) for a point to coverage. The complexity of our approach is $O(NMlogN + NM)$, it incurs an additional cost of $NM$ because of the zero-crossing detection in each trajectory.

First, we evaluate the approaches in the geographical space. By varying the geographical parameter $g$, the performances of the approaches in terms of $F_1$ on three Urban datasets are shown as in Figure 4. $Laplacian_g$ is better than $HC_g$ in all settings because it can automatically find the central area of the POI while the less significant points are labeled as noises. Compared to $HC_g$, our $Laplacian_g$ gains 12% improvement on average in terms of $F_1$ over the three data collections. In NMI comparison, $Laplacian_g$ outperforms $HC_g$ and $DBScan$ in all settings. On average, 3% improvement can be achieved if Laplacian Technique is applied to $HC_g$. The detail of NMI comparison is not listed because of the space limitation. The experiments on the geographical space demonstrate that the proposed Laplacian Technique is an effective approach against noises.

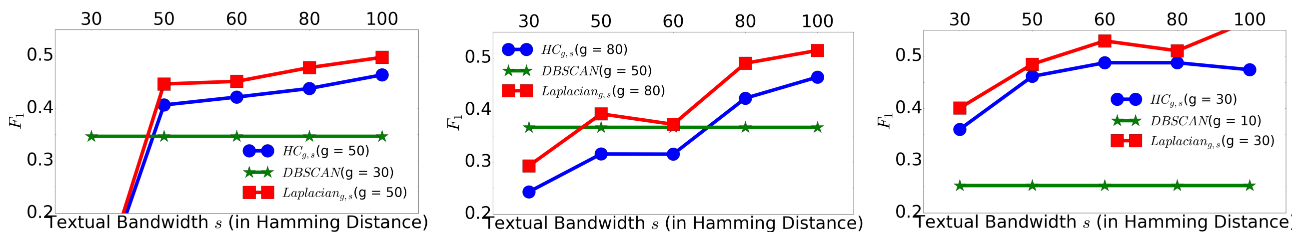
Second, for $HC_g$ and $DBScan$ we fix the geographical parameter $g$ to the setting which leads to the best $F_1$: e.g. Paris $g = 50\ meters$, New York $g = 80\ meters$, Rome $g = 30\ meters$, and further evaluate the performance by varying the value of $s$ (textual bandwidth). In the $G \times S$ joint space, in terms of the metric $F_1$ (Figure 5), our approach improves

[3] http://www.flickr.com/services/api/
[4] www.openstreetmap.org
[5] visited on 2016/11/01




(a) $F_1$ on Paris Dataset (b) $F_1$ on New York Dataset (c) $F_1$ on Rome Dataset

Figure 4: $F_1$ Evaluation By Varying Geographical Bandwidth $g$




(a) $F_1$ on Paris Dataset (b) $F_1$ on New York Dataset (c) $F_1$ on Rome Dataset

Figure 5: $F_1$ Evaluation By Varying Textual Bandwidth $s$

Table 2: Performance of $Laplacian_{g,s}$ Over $HC_{g,s}$

| $g$(meter) | Paris | | New York | | Rome | |
|---|---|---|---|---|---|---|
| | $F_1$ | $NMI$ | $F_1$ | $NMI$ | $F_1$ | $NMI$ |
| **30** | 11.05% | 2.99% | 3.52% | 1.57% | 9.88% | 1.29% |
| **50** | 6.65% | 1.98% | 8.01% | 1.82% | 12.75% | 1.99% |
| **80** | 10.00% | 2.65% | 18.03% | 2.80% | 25.44% | 1.98% |
| **100** | 8.95% | 2.21% | 22.25% | 2.26% | 27.49% | 2.60% |

the identification quality in almost every setting. According to the range setting of textual bandwidth ($s$), a small value of $s$ would lead to small POI regions. The corresponding improvement made by our algorithm is limited. However, if we set the textual bandwidth $s$ large enough (larger than 50), the performance of our proposed approach is outstanding.

As a summary, Table 2 shows the Performance of Joint $Laplacian_{g,s}$ over $HC_{g,s}$. For each setting of $g$, we compute the relative improvement over $HC_{g,s}$ by varing textual bandwidth $s$. It is obvious that our approach works much better if multi-features are used.

## 6 Related Works

Based on location-aware photos, the works [Kennedy and Naaman, 2008] identify the POIs by information in the geographical space. Yang et al. [Yang *et al.*, 2011; 2014] propose to integrate multiple information into the identification process. The authors in [Crandall *et al.*, 2009; Kurashima *et al.*, 2010] detect the landmarks (POI) by Mean Shift[Cheng, 1995]. Zheng et al. in [Zheng *et al.*, 2009] use a density-based approach to find POI hierarchy. In [Zheng *et al.*, 2012], the authors attempt to detect the user travel patterns by making use of geographical (DBScan [Ester *et al.*, 1996] based), temporal and textual features. The authors in [Bui *et al.*, 2015][Vu and Shin, 2015] propose different algorithms to identify the POI boundaries while the POI location information is required as the input.

Zheng and Xie [Zheng and Xie, 2011] construct a framework to recommend personalized traveling sequences by collaborative filtering (CF)-based model. Ying et al. [Ying *et al.*, 2014] design a system to recommend POIs in a "random-walk" manner. Liu et al. [Liu *et al.*, 2013a] solve the same problem by learning the geo-influence of the users' behaviors while [Liu *et al.*, 2013b] focus on the POI category information. Coincidentally, the authors in [Cheng *et al.*, 2013] and [Yuan *et al.*, 2013] respectively study the successive and time-aware recommendation. Based on the Web Knowledge, Adam et al. [Rae *et al.*, 2012] design a POI information recognization framework. Ying et al. [Ying *et al.*, 2013] propose a framework to extract users' traveling patterns. The works mentioned above require a geo- and text-ware POI database which can be automatically constructed by our algorithm.

## 7 Conclusion

In this paper, we propose a novel noise resistent framework for discovering Urban POIs based on textual-tagged footprints. We provided a technique to identify the POI influence region by Laplacian. Besides we also demonstrated how to integrate the geographical and textual information of the photos into the discovering process. Finally, extensive experiments are conducted to verify our claims and discussions. We believe that our techniques are particularly useful in modern travel recommendation works.

## References


[Bui *et al.*, 2015] Thanh-Hieu Bui, Yong-Jin Han, Seong-Bae Park, and Se-Young Park. Detection of poi boundaries through geographical topics. In *ICBigComps*, pages 162–169. IEEE, 2015.

[Canny, 1986] John Canny. A computational approach to edge detection. *PAMI*, 8(6):679–698, 1986.

[Charikar, 2002] Moses S Charikar. Similarity estimation techniques from rounding algorithms. In *STOC*, pages 380–388. ACM, 2002.

[Cheng *et al.*, 2013] Chen Cheng, Haiqin Yang, Michael R. Lyu, and Irwin King. Where you like to go next: Successive point-of-interest recommendation. In *IJCAI*, 2013.

[Cheng, 1995] Yizong Cheng. Mean shift, mode seeking, and clustering. *PAMI*, 17(8):790–799, 1995.

[Comaniciu and Meer, 2002] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *PAMI*, 24(5):603–619, 2002.

[Crandall *et al.*, 2009] David J. Crandall, Lars Backstrom, Daniel P. Huttenlocher, and Jon M. Kleinberg. Mapping the world's photos. In *WWW*, pages 761–770, 2009.

[Ester *et al.*, 1996] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, pages 226–231, 1996.

[Hinneburg and Keim, 1998] Alexander Hinneburg and Daniel A. Keim. An efficient approach to clustering in large multimedia databases with noise. In *KDD*, pages 58–65, 1998.

[Kennedy and Naaman, 2008] Lyndon S. Kennedy and Mor Naaman. Generating diverse and representative image search results for landmarks. In *WWW*, pages 297–306, 2008.

[Kurashima *et al.*, 2010] Takeshi Kurashima, Tomoharu Iwata, Go Irie, and Ko Fujimura. Travel route recommendation using geotags in photo sharing sites. In *CIKM*, pages 579–588, 2010.

[Liu *et al.*, 2013a] Bin Liu, Yanjie Fu, Zijun Yao, and Hui Xiong. Learning geographical preferences for point-of-interest recommendation. In *KDD*, pages 1043–1051, 2013.

[Liu *et al.*, 2013b] Xin Liu, Yong Liu, Karl Aberer, and Chunyan Miao. Personalized point-of-interest recommendation by mining users' preference transition. In *CIKM*, pages 733–738, 2013.

[Manku *et al.*, 2007] Gurmeet Singh Manku, Arvind Jain, and Anish Das Sarma. Detecting near-duplicates for web crawling. In *WWW*, pages 141–150. ACM, 2007.

[Manning *et al.*, 2008] Christopher D Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to information retrieval*, volume 1. Cambridge university press Cambridge, 2008.

[Popescu and Shabou, 2013] Adrian Popescu and Aymen Shabou. Towards precise POI localization with social media. In *MM*, pages 573–576, 2013.

[Purwar and Singh, 2016] Archana Purwar and Sandeep Kumar Singh. DBSCANI: noise-resistant method for missing value imputation. *J. Intelligent Systems*, 25(3):431–440, 2016.

[Rae *et al.*, 2012] Adam Rae, Vanessa Murdock, Adrian Popescu, and Hugues Bouchard. Mining the web for points of interest. In *SIGIR*, pages 711–720, 2012.

[Vu and Shin, 2015] Dung D Vu and Won-Yong Shin. Low-complexity detection of poi boundaries using geo-tagged tweets: A geographic proximity based approach. In *SIGSPATIAL Workshop*, page 5. ACM, 2015.

[Yang *et al.*, 2011] Yiyang Yang, Zhiguo Gong, and Leong Hou U. Identifying points of interest by self-tuning clustering. In *SIGIR*, pages 883–892, 2011.

[Yang *et al.*, 2014] Yiyang Yang, Zhiguo Gong, and Leong Hou U. Identifying points of interest using heterogeneous features. *ACM TIST*, 2014.

[Yilmaz, 2007] Alper Yilmaz. Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection. In *CVPR*, pages 1–6. IEEE, 2007.

[Ying *et al.*, 2013] Josh Jia-Ching Ying, Wang-Chien Lee, and Vincent S. Tseng. Mining geographic-temporal-semantic patterns in trajectories for location prediction. *ACM TIST*, 5(1):2, 2013.

[Ying *et al.*, 2014] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wen-Ning Kuo, and Vincent S Tseng. Mining user check-in behavior with a random walk for urban point-of-interest recommendations. *ACM TIST*, 2014.

[Yuan *et al.*, 2013] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat-Thalmann. Time-aware point-of-interest recommendation. In *SIGIR*, pages 363–372, 2013.

[Zhang *et al.*, 2012] Haipeng Zhang, Mohammed Korayem, Erkang You, and David J. Crandall. Beyond co-occurrence: discovering and visualizing tag relationships from geo-spatial and temporal similarities. In *WSDM*, pages 33–42, 2012.

[Zhao *et al.*, 2009] Qi Zhao, Zhi Yang, Hai Tao, and Wentai Liu. Evolving mean shift with adaptive bandwidth: A fast and noise robust approach. In *ACCV*, pages 258–268, 2009.

[Zheng and Xie, 2011] Yu Zheng and Xing Xie. Learning travel recommendations from user-generated gps traces. *ACM TIST*, 2(1):2, 2011.

[Zheng *et al.*, 2009] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW*, pages 791–800, 2009.

[Zheng *et al.*, 2012] Yan-Tao Zheng, Zheng-Jun Zha, and Tat-Seng Chua. Mining travel patterns from geotagged photos. *ACM TIST*, 3(3):56, 2012.