Discovering Areas of Interest using a Semantic Geo-Clustering Approach

Evaggelos Spyrou[1,2], Apostolos Psallas[2], Vasileios Charalampidis[2], and Phivos Mylonas[3]

[1] Institute of Informatics and Telecommunications, National Center for Scientific Research - "Demokritos", Athens, Greece
espyrou@iit.demokritos.gr
[2] Department of Computer Engineering, Technological Educational Institute of Central Greece, Lamia, Greece
{apsallas,vcharalabidis}@teilam.gr
[3] Department of Informatics, Ionian University Corfu, Greece
fmylonas@ionio.gr

Abstract. Living in the era of social networking, coupled together with great advances in digital multimedia user-generated content, motivated us to focus our research work on humanistic data generated by such activities towards new, more efficient ways of extracting semantically meaningful information in the process. More specifically, the herein proposed approach aims to extract areas of interest in urban areas, utilizing the increasing socially-generated knowledge from social networks. A part of the area of interest is selected, then split into "tiles" and processed with an iterative merging approach whose goal is to extract larger, "homogeneous" areas which are of special (e.g., tourist) interest. In this work generated areas of interest focus on interesting points from the humanistic point of view, thus covering in general main touristic attractions and places of interest. In order to achieve our goals, we exploit two types of metadata, namely location-based information (geo-tags) geo-tags and simple user-generated tags.

Keywords: areas of interest; semantics, geo-clustering; Flickr

1 Introduction

The recent great emerge of social media and social activities, together with advances in digital multimedia user-generated content, shifted research interest to unprecedented domains, like the ones related to the acquisition of information and analysis of the online "footsteps" or presence of users. The latter is often used to produce semi-automatic knowledge about users' whereabouts, interests or even recommend them additional, semantically related information towards covering their information needs. In this framework photographs accompanied by useful metadata information, like tags and/or geo-tags, are considered to be the ideal source of information for the discovery of meaningful, popular trends

with respect to users' behavior. More specifically, location-based info mined from such geo-tagged images offers a great opportunity to analyze users' preferences in their daily lives and complement the knowledge of their social activities through the utilization of associated tags. In this paper we present a novel approach that exploits both tags and geo-tags, towards the discovery of areas of interest.

Still, the above interpretation would be insufficient, in case it ignored the underlying semantics. By introducing a semantic geo-clustering approach we provide a novel analysis framework to merge meaningful areas of interest. In this manner we simultaneously take into consideration both location-based information in the sense of user transitions and user-location relations by incorporating respective semantic knowledge. The proposed method attempts to improve related supervised clustering approaches, by adding the inherent semantics of user tagged images derived from Flickr social network in order to enhance the precise establishment of the analysis classes. More specifically, the herein presented approach is evaluated using a large Flickr dataset consisting of approx. 80K geo-tagged images taken in Athens, Greece.

The rest of the paper is organized as follows: In next Section 2 we present relevant research efforts that also exploit user-generated geodata and metadata from Flickr. The proposed method is presented in Section 3. Then, in Section 4 we present early experimental results, along with the dataset used. Finally, we draw our conclusions and discuss plans for future work in Section 5.

2 Related Work

The motivation of this work is to ultimately "discover" large and somehow "homogeneous" areas of interest, by merging small geographic "tiles", based on sets of tags that have been added spontaneously by Flickr users. We feel that this work in novel and to the best of our knowledge there does not exist one to be compared with in terms of the produced results. However and since the aforementioned areas of interest are mainly constituted by tourist attractions (since tags have been harvested by touristic photos), it is related to research activities that aim to provide recommendations of places and/or trends using information directly from geo-tagged photos of Flickr. In general, metadata extracted from Flickr, with or without the aid of visual information have been extensively used in the literature for various research goals.A survey may be found in [8].

Since tags form the most "primitive" type of used-generated knowledge, they have been used in many research efforts. Tags, date information and geo-tags have been exploited by Chen and Roy [3], who used temporal and location distributions and photo visual similarity to extract mainly periodic events. Discovering trends for tourist attractions was the goal of Van Canneyt et al. [9], whose recommendation system adopted a probabilistic approach, ranking places of interest according to their popularity and user–related temporal information. Data clustering on geo-tagged photos was also the goal of Kisilevich et al. [5] who aimed to determine urban areas of interest by analyzed spatial and temporal distributions of metatdata so as to identify events and ranked places of

interest. Cao et al. [2], proposed a tourism recommendation system. They used mean shift clustering and built a set of representative images and tags for each cluster. Ahern et al. [1] analyzed tags that have been collected from geo-tagged photos of a specific area, and upon a TF-IDF-based approach extracted a set of the most representative ones. Similarly, Serdyukov et al. [6] aimed to predict the location photos were taken, relying solely on textual tags.

3 Geo–Clustering Algorithm

In this Section we shall present in detail the proposed algorithm. It first divides a large region into square "tiles" (sub–regions), of small, fixed size, then adopts a graph-based representation to model connectivity of neighboring tiles, each described by a set of tags. Tiles are merged and upon an iterative process, a set of larger areas is determined within the initial region. At the following we will use "tile" and "sub-region" interchangeably.

3.1 Notation and Definitions

We first select a region from the urban area of interest. Then, we divide this region into sub-regions. Many approaches have been proposed for this, e.g., in our previous work [7] we used equally-sized, round, overlapping regions. However, this would imply that overlapped tiles would share descriptions (tags), which is not a desired property in the context of this work. Thus, herein we adopt a simpler square grid-based approach, since we focus on the description and merging of sub-regions, each having an empirically set, fixed width, W_R.

Now, let R denote a given region containing a set of photos P. Let also $R_{i,j}$ denote its tiles, each containing a set of photos $P_{i,j}$, thus $\bigcup_R P_{i,j} = P$, a set of tags $T_{i,j}$, containing all tags from photos in $P_{i,j}$ and a subset $D_{i,j}$ of $T_{i,j}$, which constitutes the tag-based region description. In the aforementioned grid, i and j denote the corresponding line and column.

Since we have adopted a square grid, the most intuitive approach is to use 4–connectivity, to define the set of the initial neighboring tiles $N_{i,j} = \{R_{i,j}^{\text{up}}, R_{i,j}^{\text{right}}, R_{i,j}^{\text{down}}, R_{i,j}^{\text{left}}\}$. Obviously, $R_{i,j}^{\text{up}} = R_{i-1,j}$, $R_{i,j}^{\text{right}} = R_{i,j+1}$, $R_{i,j}^{\text{down}} = R_{i+1,j}$ and $R_{i,j}^{\text{left}} = R_{i,j-1}$. Of course, when tiles are merged, the set of neighbors of the resulting sub–region is the intersection of neighbors of the initial tiles.

3.2 Region Description

For each tile, we exploit $T_{i,j}$ to create its semantic representation $D_{i,j}$. We expect that among the user-generated tags, we shall encounter some that describe it by means of locality (e.g., *Thiseio*) or landmark(s) (e.g., *Acropolis*). Even though users tend to add "personal" tags (e.g., a name), we expect that a subset of the most "popular" tags (i.e., selected by the majority of users) will be able to describe a tile in a discriminable way. Thus, for $R_{i,j}$ the region description $D_{i,j}^L$ is the set of the L most "popular" tags, where popularity is measured in terms of the number of users that have used a specific tag within it.

Algorithm 1 Semantic Geo–Clustering

Input: Set of tiles R, set of Descriptions D, Similarity thres. S, Description length L
Output: Final Set of merged regions R

```
1: function GEO_CLUSTERING(R, S, L)
2:     for each R_{i,j} ∈ R do
3:         merge_flag ← TRUE
4:         best_match_N ← ∅
5:         best_match ← S
6:         while merge_flag = TRUE do
7:             for each N ∈ N_{i,j} do
8:                 if JACCARD(D_{i,j}, D^N_{i,j}, L) > S && JACCARD(D_{i,j}, D^N_{i,j}, L) >best_match then
9:                     best_match ← JACCARD(D_{i,j}, D^N_{i,j}, L)
10:                    best_match_N ← N_{ij}
11:                end if
12:            end for
13:            if best_match_N≠ ∅ then
14:                R_{ij} ← MERGE(R_{ij},best_match_N)
15:                merge_flag← TRUE
16:            else
17:                merge_flag← FALSE
18:            end if
19:        end while
20:    end for
21: end function
```

3.3 Region Merging

One of the challenges when comparing two sets is to select an appropriate (dis)similarity measure. Herein we use the Jaccard distance [4], which consists a well-known measure for comparing the similarity and diversity of sample sets. Jaccard similarity $J(A, B)$ between two sets A, B is given by

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \ , \tag{1}$$

where in our case A, B are the sets of tags representing two tiles, extracted using the methodology described in Section 3.2. Using the aforementioned notation, tiles R_A, R_B with descriptions D_A, D_B are merged when a) they are neighbors and b) $J(D^L_A, D^L_B) > S$, where $S \in [0, 1]$ is a user–defined similarity threshold.

The merging process starts from $R_{1,1}$ and continues horizontally. The distance to all its neighbors is checked. It is merged with the tile whose similarity is the max among all those whose similarity is greater than S, if any. For a new tile, its description is calculated based on the union of the sets of tags and the process continues by checking the similarities to its neighbors. In case there does not exist a neighbor with similarity greater than S, the process continues with the next unmerged tile. A graphical example of the tile merging process is illustrated in Fig. 1. Semantic geo-clustering, Jaccard similarity and region merging are presented in pseudocode in Algorithms 1, 2 and 3, respectively.

4 Experiments

For the experimental evaluation of our approach we used an urban image dataset which consists of a total of 79,465 photos collected from the center of the city of

Algorithm 2 Jaccard Similarity

Input: Tile Description D_1, Tile Description D_2, Tile Description Length L
Output: Jaccard Similarity Measure J_S of Descriptions D_1, D_2

function JACCARD(D_1, D_2, L)
  $D_1 \leftarrow$ SORT(D_1) // sort tags based on number of users
  $D_1 \leftarrow$ TRIM(D_1, L) // keep first L tags
  $D_2 \leftarrow$ SORT(D_2)
  $D_2 \leftarrow$ TRIM(D_2, L)
  $J_S \leftarrow |D_1 \cap D_2| / (|D_1| + |D_2| - |D_1 \cap D_2|)$
  return J_S
end function

Algorithm 3 Region Merging

Input: Sub–region R_1, Sub–region R_2
Output: Merged sub–region R_{new}

function MERGE(R_1, R_2)
  CREATE(R_{new}) // Create an "empty" new region
  $D_{\text{new}} \leftarrow$ DESCRIPTION(T_1, T_2) // Create description of new region, based on tags of R_1, R_2
  $N_{\text{new}} \leftarrow N_1 + N_2 - N_1 \cap N_2$ // Create set of neighbors of new region
  for each $N \in N_{\text{new}}$ **do**
    if $N.N_{ij} = R_1$ **or** $N.N_{ij} = R_2$ **then**
      $N.N_{ij} \leftarrow R_{\text{new}}$
    end if
  end for
  return R_{new}
end function

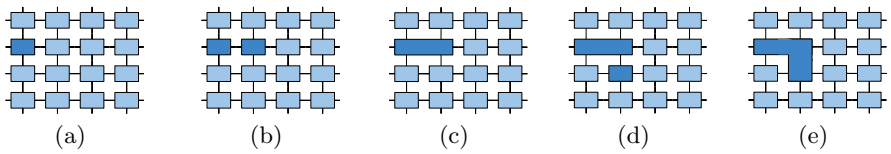


Fig. 1. Merging process: For a given tile (a), at a given step, one of its neighbors is considered as a candidate for merging (b). Their similarity is above the given threshold S, thus they are merged (c). At another step, a neighbor of the new tile is considered as a candidate for merging (d). Their similarity is above S, thus, they are merged (e).

Athens, Greece. All these photos are geo-tagged, dated between January 2004–December 2015 and collected from Flickr using its public API[4]. More specifically we queried Flickr for a region covering what is in general considered to be the center of the Athens, (i.e., where the main touristic attractions are located) and retrieved all geo-tagged photos. This rectangular area is equal to 7.7km^2. Its Northern-Western and Southern-Eastern points have coordinates $(37.9836, 23.7153), (37.9643, 23.7541)$ respectively.

The collected photos have been captured by 5038 users of various nationalities, thus they contain tags of different languages. Although the majority of these tags is in English, we used the Google Translate API[5], in order to translate

[4] https://www.flickr.com/services/api/
[5] https://cloud.google.com/translate/

non-English tags (leaving English ones unchanged). This way, tags which would otherwise act as "noise", became of use. Additionally, we also created a manual stoplist, whose goal was to remove non–relevant (to our goals) tags. For example, many cameras and smartphones automatically add brand, model and settings; also tags such as *Greece* or *Hellas* or even *Athens* are both common and spread to the whole area, thus do not provide any useful information, while also tend to be amongst the most popular. In Fig.2 we illustrate the sets of tags extracted from the tiles that correspond to the Panathenaic (Kallimarmaro) Stadium[6].

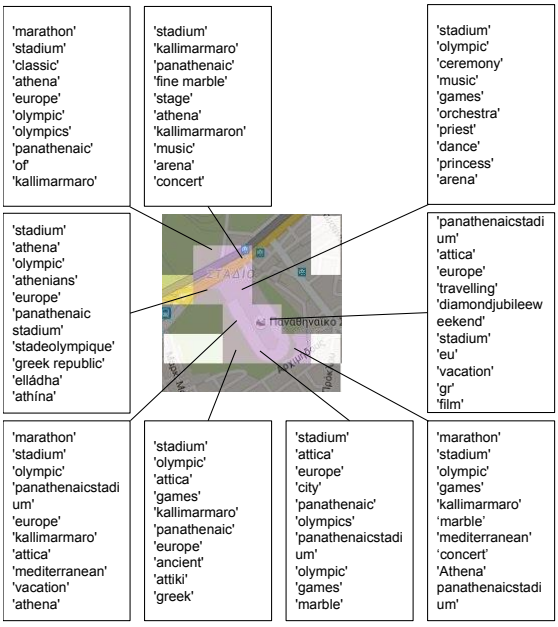
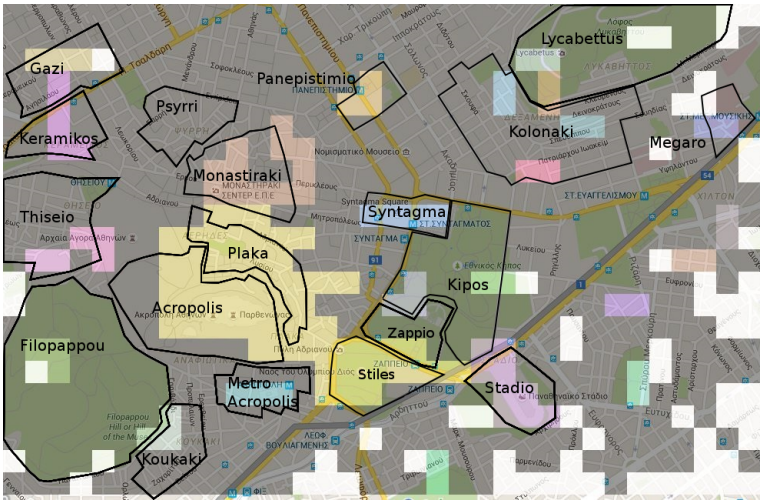


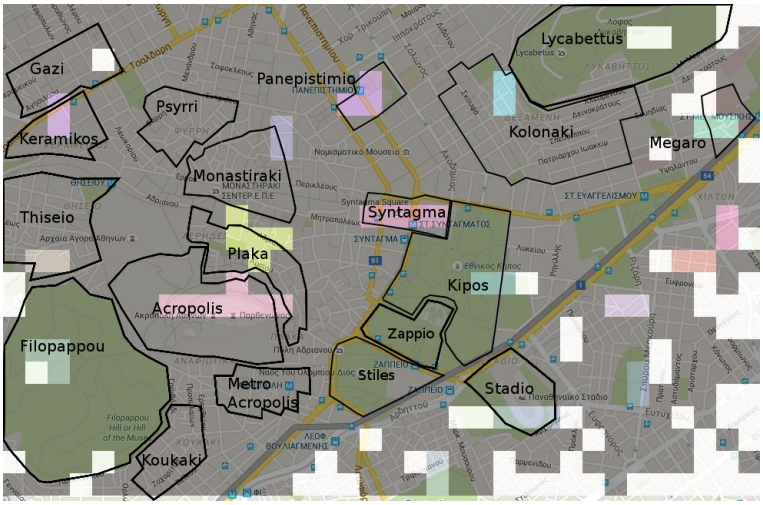
Fig. 2. Sets of tags extracted from merged tiles that correspond to the Panathenaic (Kallimarmaro) Stadium.

The entire region was divided into 770 square regions, of fixed width equal to 100m. The similarity threshold S and the number L of tags to consider were empirically selected. More specifically, after considering several try-and-error rounds where the algorithm failed to produce satisfactory results, either by merging the majority of tiles into a single region, or by performing only a few merges, resulting to small and non-meaningful areas, we identified their optimal values. Furthermore, for the sake of fair evaluation, we constructed an empirical "ground truth" set of areas upon discussion with residents (which were not involved in this work). In the following we depict two early results of our algorithm compared to the constructed ground truth in Fig. 3, where we may observe that even a small

[6] https://en.wikipedia.org/wiki/Panathenaic_Stadium



(a) $S = 0.85, L = 10$



(b) $S = 0.9, L = 10$

Fig. 3. Results compared to an empirically constructed ground truth set of areas.

increase in S may lead to a significantly improved qualitative result. Also, the result of Fig. 3(a) is quite close to the ground truth. We should note that at this early stage of research it is pointless to provide a more quantitative result, e.g., by measuring the coverage of the ground truth. Finally, it is also worth mentioning that the algorithm merges also some additional regions, irrespectively to the ground truth ones; the latter still produce a meaningful merging result, however not of touristic interest, like the ones on the bottom right of Fig. 3(a).

5 Conclusions and Discussion

In this paper we presented an approach which aimed to extract areas of interest in urban areas, using socially-generated knowledge from Flickr. We selected a part of the city of Athens, split it into "tiles"and proceeded with an iterative merging approach whose goal was to extract larger, "homogeneous" areas of the city which are of tourist interest. We showed that our approach succeeds when its parameters are selected appropriately. The generated areas cover in general the main touristic attractions and places of interest of the city center, and their boundaries often are close to those that a local resident may determine if asked. We should emphasize that we used almost "raw" tags, i.e., we did not use any "intelligent" technique to select/filter tags based on their relevance to the problem at hand, apart from the process described in Section 4.

In future work we plan to further improve the proposed algorithm and apply it into larger urban areas (e.g. the whole city of Athens and other major European cities). We also aim to perform an exhaustive evaluation, regarding the sizes of tiles, sets of tags to consider and similarity threshold and investigate whether a set of parameters shows satisfactory performance in more cities. We also wish to evaluate the results by real life users. A possible way is to consider the output of our system as a set of tourist recommendations and focus on user satisfaction. However, such an evaluation has been shown to be a difficult and expensive task, which may involve empirical issues in the process. Thus it should not also involve local residents but also visitors in both the construction of a "ground truth" and the assessment of the system's output.

References

1. S. Ahern, M. Naaman, R. Nair and J.H.-I. Yang, *World explorer: visualizing aggregate data from unstructured text in geo-referenced collections*. Proc. of the ACM/IEEE-CS JCDL, 2007.
2. L. Cao, J. Luo, A. Gallagher, X. Jin, J. Han, and T.S. Huang, *A worldwide tourism recommendation system based on geotagged web photos*. Proc. of IEEE ICASSP, 2010.
3. L. Chen and A. Roy, *Event detection from flickr data through wavelet-based spatial analysis*. Proc. of ACM CIKM, 2009.
4. R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification, 2nd ed.*, Wiley, 2001
5. S. Kisilevich, D. Keim, N. Andrienko and G. Andrienko, *Towards Acquisition of Semantics of Places and Events by Multi-perspective Analysis of Geotagged Photo Collections*. Geospatial Visualisation, pp. 211–233, Springer, 2013.
6. P. Serdyukov, V. Murdock and R. Van Zwol, *Placing flickr photos on a map*. Proc. of ACM SIGIR, 2009.
7. E. Spyrou and Ph. Mylonas, *Analyzing Flickr metadata to extract location-based information and semantically organize its photo content*, Neurocomputing, vol.172, pp. 114-133, Elsevier, 2016
8. E. Spyrou and Ph. Mylonas, *A survey on Flickr multimedia research challenges*. Engineering Applications of Artificial Intelligence, vol. 51, pp. 71-91, Elsevier, 2016.
9. S. Van Canneyt, S. Schockaert, O. Van Laere and B. Dhoedt, Bart, *Time-dependent recommendation of tourist attractions using Flickr*. Proc. of BNAIC, 2011.