From Tags to Trends: A First Glance at Social Media Content Dynamics

Evaggelos Spyrou, Phivos Mylonas

From Tags to Trends: A First Glance at Social Media Content Dynamics

Evaggelos Spyrou[1] and Phivos Mylonas[2]

[1] Technological Educational Institute of Lamia
Lamia, Greece
vspyrou@teilam.gr
http://www.teilam.gr
[2] National Technical University of Athens
Iroon Polytechneiou 9, Zographou, Athens, Greece
fmylonas@image.ntua.gr
http://www.image.ntua.gr

Abstract. Current uncontrolled growth of online, digital multimedia content emphasizes research work on identifying trends on how this content popularity may grow over time wrt identifiable user events and interests. In this paper we analyze user-generated photos uploaded to Flickr[1] in order to extract meaningful semantic trends covering specific geographical areas of interest. Initially, we cluster photos based on their geotagging metadata information and divide large areas into smaller "first level geo-clusters" of fixed size, allowing them to overlap if necessary. Within these first level geo-clusters, we identify semantically meaningful "places" of user interest, by analyzing additional textual metadata, i.e. user selected tags that characterize each place's photos. By post-processing them, we select the most appropriate tags that are able to describe landmarks and events occurring within these places of interest and examine their temporal dynamics over a long period of time.

Keywords: social media, clustering trends, geo-tagging

1 Introduction

Over the last years the massive creation of user-generated multimedia content, mostly in the form of digital still images, along with the arise and expansion of social networking activities, resulted in the generation of the so-called *social media*. The latter summarize all web- and mobile-based technologies used to turn human communication into an interactive process between individuals and communities. In an attempt to define this phenomenon, the work of [11] defines social media as "a group of Internet-based applications that allow the creation and exchange of user-generated content". Being the center of attention, online user-generated, multimedia content met an unprecedented interest increase in terms of its organization and manipulation. Typically, this rather

[1] http://www.flickr.com

new kind of online multimedia content is produced, managed and consumed by users or user communities, who, from the one hand, often spend a lot of time linking themselves together through social networks, but, on the other hand, do not spent similar efforts to semantically characterize, meaningfully organize and thus efficiently exploit its nature.

Being part of a brave new digital world, online information in total is becoming increasingly dynamic and hard to track and identify patterns in it. Obviously, recent emergence of social media and enriched user-generated content aids to the benefit of this observation. For instance, multimedia content shared on microblogging platforms, like Twitter[2], is considered to be extremely unpredictable, since it usually becomes popular and fades away within a very short period of time. In addition, the art of analyzing and identifying patterns of temporal variation with respect to online content in general, forms another difficult task, mainly due to the fact that human behavior, that is inherent behind the temporal variation, is considered to be highly unpredictable and outside of any known model. The latter ranges typically between "random" [12] and "highly correlated" [13] states. Towards this direction, the addition of the notion of collectiveness aids the overall pattern deviation and complexity increase, considering all possible differentiations in interactions between small or larger groups of people. As a result, we noticed that there has been little work about patterns characterizing online user-generated multimedia content and by which different pieces of content compete for attention during this process, constituting the fulfillment of our motivation an extremely intriguing research task.

The rest of this paper is organized as follows. In section 2 we begin by presenting recent research on handling community collected photo metadata, mainly focusing on Flickr social network and other online collections. Then, in section 3 we present the main aspects of our work, which briefly may be summarized in the clustering technique we apply on photos based on their geo-data and the tag-ranking algorithm we apply on each cluster. Early experimental results derived from our first use case are presented in section 4 Finally, in section 5 we draw our conclusions and briefly present our future plans.

2 Related work

The tasks of semantically characterize, organize and efficiently exploit user-generated multimedia content towards their meaningful exploitation are of great importance within recent research community efforts. Starting back in 2009, Cha et al. [14] collected and analyzed large-scale traces of information dissemination derived from Flickr, aiming at answering a set of information propagation questions. More recently, Kalantidis et al. [17] proposed a visual-based photo image retrieval and localization approach, which exploited low-level image characteristics similarities in order to achieve accurate results. In our own most recent work on the topic [19], we proposed a tag-based methodology analyzing large

[2] `http://www.twitter.com`

Flickr user photo collections in order to select the most appropriate set of tags to be descriptive enough so as to describe an entire spatial area. Another interesting approach is [16], where meaningful travel route recommendations are proposed utilizing Flickr's user histories and past actions behaviors. Still, other approaches focus on mobile platforms and try to investigate whether knowledge extracted from massive content user contribution and interaction may offer any kind of added-value services [15]. Lately, research interest has been given also on statistical approaches to the problem, i.e. Yang et al. [18] developed a k-spectral centroid clustering algorithm so as to identify temporal patterns in online media.

Focusing more on the challenging geographical aspect of the problem, Lee et al. [7] created overlapping geographical clusters for each tag, calculated geographical similarity for pairs of tags, and then introduced similarities for both tags and geographical distributions. Rattenbury et al. [8] extract semantics such as places and events from tags and unstructured text-labels, observing that event tags follow certain temporal patterns, while place tags follow certain spatial patterns. In the same manner, Abbasi et al. [5] identify landmarks using tags and Flickr groups, without exploiting geospatial information, aiming to find relevant landmark-related tags, whereas the work presented in [6] analyzes tags associated with geo-referenced Flickr images and uses a TF-IDF approach to generate knowledge as a set of the most "representative" tags for an area. Finally, Serdyukov et al. [9] adopt a language model which lies on the user collected Flickr metadata and aims to annotate an image based on these metadata and place photos on a map, i.e. provide an automatic alternative to manual geo-tagging.

3 Content Metadata Processing

Although detailed in their nature, most of the above research efforts fail to deal with the problem of identifying and analyzing meaningful semantic trends that are inherent within multimedia content over long periods of time. Our approach presented herein, acts complementary to most of them, since, given a limited set of specific geographical areas of interest, it utilizes textual metadata information to extract meaningful semantic trends and compute their temporal fluctuations.

3.1 Place Clustering

The first step of our methodology consists of a necessary clustering initialization. For reasons relating to data manipulation efficiency, we choose to use the *kernel vector quantization* (KVQ) approach of Tipping and Schölkopf [10] to implement this kind of clustering process. Taking into account this encoding method, maximal distance between clusters may be regarded as the maximum *distortion* level. Using KVQ we have a guaranteed upper bound on distortion and the number of clusters is adjusted accordingly.

Given a point $x \in X$, we define *cluster* $C(x) = \{y \in D : d(x, y) < r\}$ as the set of all points $y \in D$ that lie within distance r from x. The *codebook* $Q(D)$ we obtain by applying KVQ has the following properties. (i) $Q(D) \subseteq D$, that is,

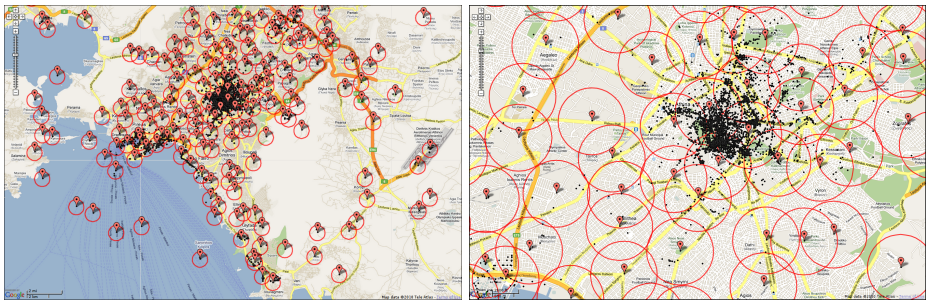

Fig. 1. A map of Athens depicting all geo-clusters. By black dots, red markers and red circles we mark photos, geo-cluster centers and geo-cluster boundaries, respectively.

codebook vectors are points of the original dataset. (ii) The *maximal distortion* is upper bounded by r, that is, $\max_{y \in C(x)} d(x,y) < r$ for all $x \in Q(D)$. (iii) The *cluster collection* $\mathcal{C}(D) = \{C(x) : x \in Q(D)\}$ is a *cover* for D, that is, $D = \bigcup_{x \in Q(D)} C(x)$. However, it is *not* a partition as $C(x) \cap C(y) \neq \emptyset$ in general for $x, y \in D$, that is, clusters are *overlapping*.

Geo-clustering Let P be a set of photos, each photo $p \in P$ represented by (p_{lat}, p_{lon}), where p_{lat} and p_{lon} define its capture location, i.e. *latitude* and *longitude*, respectively. We geo-cluster P by applying KVQ in metric space $(\mathcal{P}, d_g)$ with scale parameter r_g, where $\mathcal{P}$ is the set of all possible photos and metric d_g is the *great circle distance* [17]. Given a photo $p \in P$,we define a *geo-cluster* as $C_g(p) = \{q \in P : d_g(p,q) < r_g\}$. That is, the set of all photos $q \in P$ that lie within geographic distance r_g from p. Similarly, given the resulting codebook $Q_g(P)$, define the *geo-cluster collection* $\mathcal{C}_g(P) = \{C_g(p) : p \in Q_g(P)\}$. Contrary to other clustering techniques, the number of clusters is automatically adjusted to the maximal distortion r and not user pre-defined.

We apply KVQ on each cluster, separately and we obtain a set of places $C_l(p) = \{q \in C_g(p) : d_g(p,q) < r_p\}$. For a geo-cluster $C_g(p)$ we define the set of places as $C_l(C_g(p)) = \{C_l(p) : p \in C_g(p)\}$, where $r_p = L_f \cdot r_g$, L_f denoting the ratio of the radius of a place to the one of a geo-cluster. In Figure 1, we illustrate a map of Athens depicting all geo-clusters at two different zoom levels, for $r_g = 700m$. We should note the density of photos in the city center and particularly in the area of the *Acropolis*. Photos taken even $1km$ away from a landmark may be included in the same cluster. The total number and position of clusters is automatically inferred according to supplied data. In Figure 2 we illustrate a geo-cluster and all places within it. We should note that places cover only a small fragment of the geo-cluster in such a non-popular part of the city.

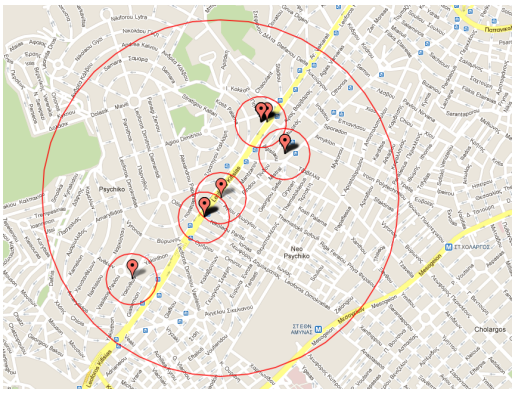

Fig. 2. A geo-cluster and all places extracted within it, using KVQ. Red markers mark photos. The radius of a geo-cluster is $r_g = 700m$, while the one of a place is $r_p = 100m$.

3.2 Exploiting Metadata Intelligence

As it has already been mentioned, our approach relies both on geo-tags and on the most common textual tags. In Flickr, as in many similar websites, users are able to *tag* their photos with a few descriptive words, according of course to their own perception. They often accompany this set of tags with a *title*. Titles are typically more reliable, but more complex to analyze. Tags can also be helpful, despite being rather noisy. A simple step to reduce noise is to first filter all tags of the entire dataset through a manually created stop-list. In previous work [17], this stop-list has been created using terms that are too generic (e.g. *paris*, *france*, *holidays*), describe the conditions of the photo-shoot (e.g. *night shot*, *black and white*), or are typically irrelevant to the content of the photos (e.g. *nikon*, *geo-tagged*). In this approach, current stop-list does contain tags that describe countries/cities and locations in general, for reasons that will clarify in the following. Moreover, when a photo is uploaded to Flickr, it is typically accompanied by its associated camera metadata. These of interest for the current work include *date taken*, the *name* of the owner and its geo-location.

In the framework of this approach we focus on a set of images taken within specific city limits. First, we apply the already discussed clustering algorithm. In this manner, we obtain a set of geo-clusters. We then re-apply the clustering algorithm on each one of them, using an appropriate percentage for the radius c_r of a geo-cluster, namely p_r, thus obtaining a set of places for each geo-cluster. For the time being, our analysis relies on the properties of the content of each place in a separate manner, without taking into consideration its neighbors. Now, let's assume that a given place L_j contains a set of photos $P = \{p_i\}$. Let T be the set of all tags in this place. For a given set of photos P_k, we will denote the set of all tags these photos have been tagged with, by $\mathcal{T}(P_k) = \{t \in T : t \in P_k\}$. Then, $\mathcal{T}(P_j)$ is the set of all tags of place L_j. For each place and using the dates its photos have been taken, we are able to create a cumulative distribution of

its "popularity" through time, for a given date d as $F_p(d) = |\{p_i \in P : d_i \leq d\}|$, where by $|\bullet|$ we denote set cardinality.

Continuing, we shall work on these specific sets of tags and try to understand how tags and geo-tags vary through time, for different types of places. First, we group similar tags, using the well-known and widely adopted Levenshtein distance [20], denoted by d_L in the following. This distance is computed for two given tags t_i, t_j, which are considered similar and are *grouped*, if $d_L(t_i, t_j) < T_{lev}$, where T_{lev} is an appropriately chosen threshold. We treat each such group $\mathcal{T}_d = \{t_i, t_k \in \mathcal{T}(P_j) : d_L(t_i, t_k) \leq T_{Lev}\}$ as a single tag which will be denoted as "representative". The representative tag is the most frequent of the group and "inherits" the dates of all tags belonging to its group. We then calculate the cumulative distribution of each representative tag, for a given date d as $F_g(d) = |\{t_i \in \mathcal{T}_g : d_i \leq d\}|$.

3.3 Defining Trends

Landmarks It is well acknowledged that the most significant places of interest of a given city, are denoted by the term *landmarks* and may include buildings, statues, squares, archaeological sites and so on. In other words, landmarks do denote the most popular places of a city for its visitors. Since "popularity" is definitely a vague term, which in turn may not be able to be measured precisely, in this work, we tend to define and estimate it in a threefold sense explained in the following. First, one should expect that a large amount of photos is taken in the close spatial area surrounding a potential landmark. Second, these photos are generally taken by a large number of people (Flickr users in our case), since landmarks are places of general interest. Finally, since a popular landmark is generally of interest all year, we expect that photos taken thereby should be distributed uniformly through time, or, in general, under a "predictable" distribution. By that, we mean that one should expect e.g. in the Athens case, an increasing number of photos taken between June and August, i.e. exactly when tourist season reaches its peak and a decreasing one between August and April, and so on.

Events Merriam-Webster dictionary defines *events* as "competitive encounters between individuals or groups carried on for amusement, exercise, or in pursuit of a prize". By the term *events*, in this work, we consider events like concerts, festivals, musicals, theatrical performances and so on. We also consider athletic events such as a marathon race, a football game or even Olympic Games as a whole. Considering its nature, duration of an event typically varies. Some such as a football game or a concert may last a few hours, while a festival and some athletic events may span across many days. Since events often attract interest, one should expect an increased number of photos during an event's lifetime, concentrated either to a small spatial area, e.g. a football stadium, or to a significantly larger one, e.g. the Marathon route, which extends itself over more than 40 kilometers.

Places of "no-interest" Since Flickr defines itself as an online media hosting website, it is addressed to all kinds of users and particularly to artists, or tourists. Consequently, one should expect to find many photos of non-landmarks, or non-events in the considered dataset. Typical examples of photos that may be denoted as of "no-interest" for our work may depict a house, a family meeting, etc. We make the assumption that these photos have been taken at non-popular places (i.e. at least when considering the notion of popularity from the general public point of view!); such places generally contain less than 10 photos, typically taken by a single user and tagged with the same tags. At this point we should note that this kind of content/photos may also appear in so-called popular places, but due to their limited number, they may be acceptably considered as "noise" and have a limited effect on the overall analysis process.

4 Experimental Results

4.1 Preparatory Phase

In order to extract trends from Flickr and for the sake of a meaningful demonstration, we used a limited urban image dataset consisting of a total of $18,355$ geo-tagged images from the city of Athens, Greece. These photos have been collected from Flickr using a geographic query that covers a window of the city's center. For each image we have also downloaded all available textual and location metadata. Applying the discussed clustering algorithm produced initially 193 geo-clusters, with a radius of $c_r = 700$m. We then clustered each geo-cluster, using a radius of $p_r = 100$m. the result of this process was the production of 73 places. From each one of the 2123 places, we collected and re-ranked all accompanying tags and by applying a rather aggressive stop-list, we obtained a set of semantically meaningful tags.

4.2 Trends in Athens

Landmarks in Athens There may be no doubt that in Athens, *Acropolis* is by all means the most popular landmark. Thus, in order to present our observations on trends within landmarks category, it was a wise choice to select a geo-cluster in the area of Acropolis, that contained 11298 photos, divided in 19 places. These photos contain more than 100K tags. We analyze this geo-cluster using the proposed approach and a set of 2232 representative tags occurs. In Table 1 we present the most popular tags, grouped as described in Section 3.2. We also include the corresponding frequencies of each group. Tags such as *Athens*, *Greece*, *Acropolis*, *Parthenon* and so on, are distributed in time in a predictable sense, as expected by the assumptions we made. Their distributions are depicted in Fig. 3. We may note that each user has taken 25 photos on average.

Events in Athens Typically an event refers to a specific thing that occurs once at a specific time and place. Hence, given the set of Athens digital images, an event happening in the city satisfied following rules:

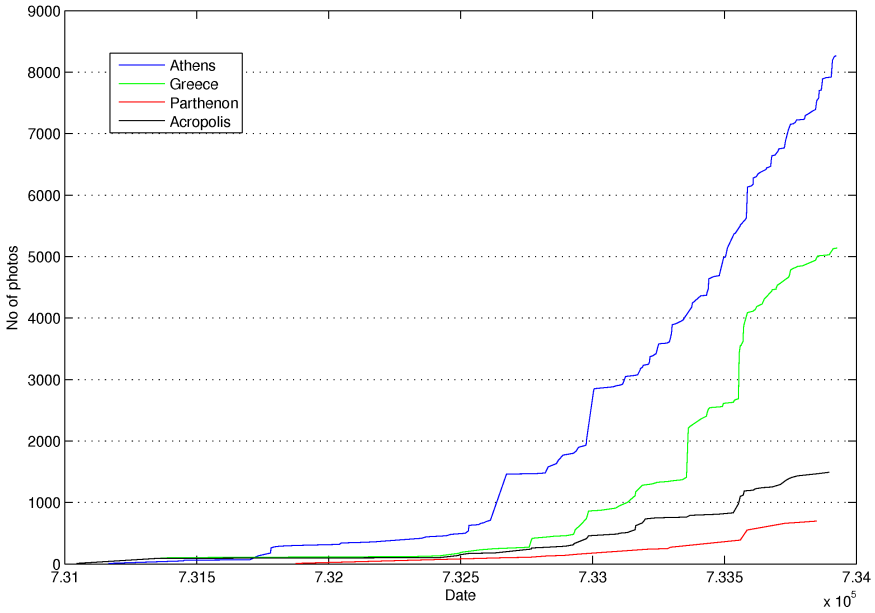

Fig. 3. The cumulative distribution of the most frequent representative tags for a place in the area of Acropolis from 2004 to June 2009. Dates are expressed in MATLAB's serial date number format.

- content of photos should be semantically consistent and since we deal with tags, the latter should be semantically similar
- the group of its photos should have been taken within a certain time period
- the group of its photos should have been taken around the same geo-location

We selected a geo-cluster in the Olympic Centre of Athens. Following the proposed approach, this geo-cluster was divided in 17 places. We select tag "Athens 2004" and a group of tags represented by tag "Athens". Their distributions through time are depicted in Fig.4. We may observe that tag "Athens" is distributed through time in a predictable sense, i.e. such as in Section 4.2, while "Athens 2004" is forming a so-called "spike" in time, as expected by the assumptions already analyzed.

Places of "no-interest" in Athens In Fig. 1, one would have noticed that in Athens, there exist many geo-clusters with only few photos, each containing a sole place. We randomly select one of them and depict it in Fig. 5 without any major loss of generality. Its corresponding representative tags are depicted in Table 2. We may observe that these tags suggest that this cluster represents a small photo collection of a family gathering during Easter time. Furthermore, it is also identifiable that discussed photos have been taken over a timespan of only 3 days.

Athens, Aten, Atenas, Atene, Ateny, Athen, Athene, Atheny, Athina, athen, athena, athenes, athens
Acropole, Acropoli, **Acropolis**, Akropoli, Akropolis, acropole, acropoli, acropolis, akropoli, akropolis
Grece, Grecia, Grecja, **Greece**, Greek, Greeca, grece, grecia
Parthenon, parthenon

Table 1. Groups of most frequent tags, for a place near the Acropolis. Representative tags are in bold.

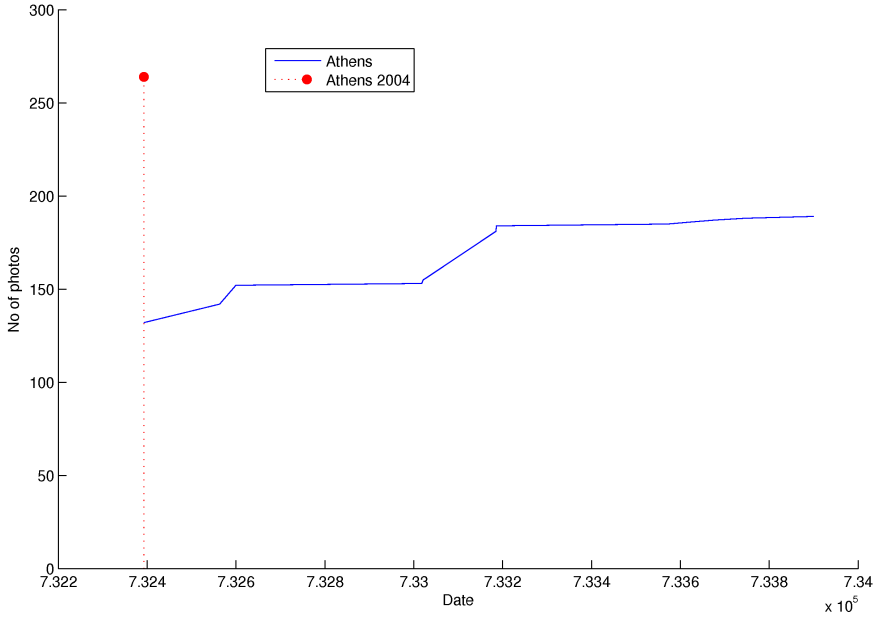

Fig. 4. The cumulative distribution of tags "Athens" and "Athens 2004" for a place in the Olympic Aquatic Centre of Athens from 2004 to June 2009. Dates are expressed in MATLAB's serial date number format.

5 Conclusions and future work

In this paper we have presented our first attempt to manipulate inherent social media dynamics deriving from multimedia content and provide initial results from our work on the meaningful analysis of the latter with respect to the observed underlying trends. We have shown that by being able to exploit different kinds of metadata information (i.e. textual, geo-, temporal and visual chunks of data), one is able to identify meaningful content trends over a large period of time. These trends would facilitate user interaction with generated and already

Kitties, barbecue, easter, home, pasxa

Table 2. Representative tags, for a place of no particular interest.

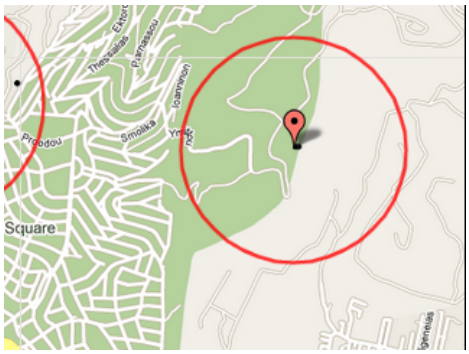

Fig. 5. A geo-cluster of no particular interest with 13 photos, all taken in the same location.

stored content, allowing them to capture at a glance landmarks and events of their interest.

Among our future plans is to further test the proposed approach to larger user-generated datasets, like the one utilized within the VIRaL[3] framework, which currently[4] includes 2.221.176 digital images collected from 39 cities around the world. In addition, we plan to further build on the results of this work by studying tag behavior within a given "place" of interest, as well as by semantically combining "places" based on a meaningful semantic popularity criterion to be defined.

References

1. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. May, P., Ehrlich, H.C., Steinke, T.: ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow through Web Services. In: Nagel, W.E., Walter, W.V., Lehner, W. (eds.) Euro-Par 2006. LNCS, vol. 4128, pp. 1148–1158. Springer, Heidelberg (2006)
3. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
4. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. In: 10th IEEE International Symposium on High Performance Distributed Computing, pp. 181–184. IEEE Press, New York (2001)

[3] `http://viral.image.ntua.gr`
[4] as of May 2012

5. R. Abbasi, S. Chernov, W. Nejdl, R. Paiu and S. Staab, *Exploiting Flickr Tags and Groups for Finding Landmark Photos*, Advances in Information Retrieval, Springer, 2009.
6. S. Ahern, M. Naaman, R. Nair and J.H.I. Yang, *World Explorer: Visualizing Aggregate Data from Unstructured Text in Geo-Referenced Collections*, 7th ACM/IEEE-CS conf. on Dig. libraries, 2007.
7. S.S. Lee, D. Won and D. McLeod, *Tag-Geotag Correlation in Social Networks*, ACM Workshop on Search in Social Media, 2008.
8. T. Rattenbury, N. Good and M. Naaman, *Towards Automatic Extraction of Event and Place Semantics from Flickr Tags*, 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2007.
9. P. Serdyukov, V. Murdock and R. Van Zwol, *Placing Flickr Photos on a Map*, 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2009.
10. M. Tipping and B. Schölkopf, *A Kernel Approach for Vector Quantization with Guaranteed Distortion Bounds*, Artificial Intelligence and Statistics, pp.129–134, 2001.
11. A. Kaplan and M. Haenlein, *Users of the world, unite! The challenges and opportunities of Social Media*, Business Horizons, pp. 59-68, 53 (1), 2010.
12. R.D. Malmgren, D.B. Stouffer, A.E. Motter, L.A.N. Amaral, *A poissonian explanation for heavy tails in e-mail communication*, PNAS, pp. 18153-18158, 105 (47), 2008.
13. A.L. Barabsi, *The origin of bursts and heavy tails in human dynamics*, Nature, 435:207, 2005.
14. M. Cha, A. Mislove, K.P. Gummadi, *A Measurement-driven Analysis of Information Propagation in the Flickr Social Network*, Proceedings of the 18th international conference on WWW, pp. 721-730, New York, NY, U.S.A., 2009.
15. C. Zigkolis, Y. Kompatsiaris, A. Vakali, *Information analysis in mobile social networks for added-value services*, W3C Workshop on the Future of Social Networking, 2009.
16. T. Kurashima, T. Iwata, G. Irie, K. Fujimura, *Travel route recommendation using geotags in photo sharing sites*, Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10, pp. 579-588, 2010.
17. Y. Kalantidis, G. Tolias, Y. Avrithis, M. Phinikettos, E. Spyrou, P. Mylonas, S. Kollias, *Viral: Visual image retrieval and localization*, Multimedia Tools and Applications, 1-38, 2011.
18. J. Yang, J. Leskovec, *Patterns of Temporal Variation in Online Media*, in ACM International Conference on Web Search and Data Minig (WSDM), Hong Kong, China, 2011.
19. E. Spyrou, Ph. Mylonas, *Placing User-Generated Photo Metadata on a Map*, 6th International Workshop on Semantic media adaptation (SMAP 2011), Vigo, Spain, 2011.
20. V.I. Levenshtein, *Binary codes capable of correcting deletions, insertions, and reversals*, Soviet Physics Doklady 10, pp. 70710, 1966.