



**HAL**  
open science

# Foto2Events: From Photos to Event Discovery and Linking in Online Social Networks

Eliana J. Raad, Richard Chbeir

► **To cite this version:**

Eliana J. Raad, Richard Chbeir. Foto2Events: From Photos to Event Discovery and Linking in Online Social Networks. SocialCom2014, Dec 2014, Sydney, Australia. hal-01082846

**HAL Id: hal-01082846**

**<https://hal.science/hal-01082846>**

Submitted on 14 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Foto2Events: From Photos to Event Discovery and Linking in Online Social Networks

Eliana J. Raad  
LE2I – CNRS UMR  
Bourgogne University  
Dijon, France  
eliana.raad@u-bourgogne.fr

Richard Chbeir  
LIUPPA Laboratory  
University of Pau and Adour Countries  
Anglet, France  
rchbeir@acm.org

**Abstract**— Online social networking has become the predominant activity in the digital world thanks to multimedia data (mainly photos) sharing (e.g., photos now represent 93% of the top posts on Facebook). Discovering events where users are involved using their own posts and those shared by their friends would be of great importance. In this paper, we address this issue by providing an original approach able to detect, enrich and also link user’s events using photos shared within his online social networks. Using metadata, our approach provides a multi-dimensional gathering of similar photos using their temporal, geographical, and social facets. To validate our approach, we implemented a prototype called *Foto2Events* and conducted a set of experiments on real and generated data. Results show that our approach works well for various metadata distributions.

**Keywords**—image clustering; metadata; event detection

## I. INTRODUCTION

Online social networking has become the predominant activity on the Web. According to the Pew Internet Project’s research related to social networking<sup>1</sup>, 74% of users use online social networking sites and more than their half post original multimedia resources and mainly photos. This clearly shows how photos have become the primary content form of the online sharing, expressing easily a variety of information such as interests, activities and events of the everyday life of people. Only on Facebook<sup>2</sup>, in a single month, 2.5 billion photos were uploaded, most of them identified<sup>3</sup>. Furthermore, online social networking has actually affected the way photos are captured. People are constantly taking photos to instantly share them online. From the coffee mug for *#Breakfast* to the *#Sleeping #Cat*, people are snapping photos of everything. However, they do not realize the wealth of information (metadata) that can be obtained from those photos and which could be used to track their activities. For instance, the project “*I Know Where Your Cat Lives*”<sup>4</sup> raises concerns over online privacy by using public images of cats uploaded to photo sharing websites (e.g., Flickr<sup>5</sup>, Twitpic<sup>6</sup> and Instagram<sup>7</sup>). Location coordinates

embedded in the images’ metadata are used to show where each cat lives and, more importantly, to track their owner’s homes. Living in this connected world, people are voluntarily exposing their every movement and putting more and more of their personal information online. Particularly, sharing real-life experiences with friends at any time and any place has become common on social networks, from social (festive celebrations, concerts, sporting events, etc.) and personal events (family gathering, wedding and other parties, etc.) to the most intimate private moments, conducting sometimes to breach users’ privacy and to scandals (e.g., the well-known example, *#Aftersex* selfies<sup>8</sup> is one of the biggest trends on Instagram where couples post images of their post-coital faces showing themselves lying on bed and even one night stands).

We focus in this paper on detecting events from image collection, since photos posted on social networks represent real stories from users’ life. There are several applications where events detected from photos<sup>9</sup> may be useful:

- Enriching the user profile information from the Social Web: In a previous work [1], we proposed *de-linkability*, a novel technique for preserving individual privacy when sharing multimedia documents. *de-linkability* ensures that user profile information cannot be used to infer his identity. Additional information can be obtained by exploring events of a user or of his contacts.
- Estimating missing metadata in images that do not have social tags or metadata about the image capture [2]. For example, when a user is identified in an event photo containing all metadata, propagation of information can be performed to estimate untagged faces and annotate other photos of the same event.
- Discovering hidden relationships between users in a social network. The Facebook algorithm presented in [3] uses the individual’s network neighborhood to identify people who are dating. By examining people who participated in the same events, new relationships, as well as uncovered relationships can be also discovered.

<sup>1</sup> <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/>

<sup>2</sup> [www.facebook.com](http://www.facebook.com)

<sup>3</sup> <http://www.amara.org/en/videos/dufdmSIJBpYX/info/why-privacy-matters/>

<sup>4</sup> <http://iknowwheretheyourcatlives.com/cat/e3fb32d4a0>

<sup>5</sup> <https://www.flickr.com/>

<sup>6</sup> <http://twitpic.com/>

<sup>7</sup> <http://instagram.com/>

<sup>8</sup> <http://www.complex.com/pop-culture/2014/04/aftersex-hashtag>

<sup>9</sup> Both terms *photo* and *image* will be used interchangeably in the remainder of this paper.

- Preventing identity leakage from events shared online for better privacy protection [4]. When a user is identified in an event, information included in images from this event can be used to disclose sensitive information for the user, or personal data he did not intend to disclose online (e.g., home location).

Recent studies have explored how to detect events in photo collections. Many of these approaches are quite interesting and can be classified in two groups: event clustering approaches and event hybrid approaches. Event clustering approaches [5-11] use spatial, temporal or spatio-temporal information to cluster images. Event hybrid approaches apply specific models [12, 13] and methods such as graph-based algorithms [14, 15]. However, they do not identify links between personal events (e.g., sub-events within a large event) and do not deal with particular events (e.g., *multi-day* and/or *multi-site* events, etc.).

In this paper, we cluster images based on their metadata in the case of online social networks. Metadata gives information about objects, namely persons *who* participated in the event, *where* the event happened, and *when* it happened. The main contributions of this paper are summarized as:

- An event model that estimates what can trigger or end an event, captures the *Where*, *When* and *Who* dimensions of an event, and describes (temporal, spatial, and semantic) relationships between events using image metadata only.
- A clustering algorithm for photos uploaded online based on their metadata attributes, and a rule refinement technique that allows to investigate the multi-\* property of an event (*multi-day* events, *multi-site* events, *multi-person* events).
- A prototype tool called *Foto2Events* to evaluate our approach on real and generated datasets.

The remainder of this paper is organized as follows. In Section 2, we review the related work. Section 3 shows our motivating scenario. Section 4 presents the data model and our event detection approach. Section 5 presents our prototype and experimental tests. Section 6 concludes the paper and discusses future directions.

## II. RELATED WORK

### A. Event Definition

Many different definitions of events from various research fields have been proposed in the literature. In online event detection, most definitions about events refer to their spatial, temporal [16, 17] or spatio-temporal aspects [18-20]. In [19], an event is defined as a 3-tuple  $\langle E, R, t \rangle$ , where  $E$  is a set of entities,  $R \subseteq E \times E$  is a set of dynamic relationships,  $t$  is a continuous time window. Other definitions have provided categories for events [21]: *local* and *global events*, *home* and *away-from-home* events, *routine* and *non-routine* events, etc. *Global events* build collective experiences that permit the sharing of personal experiences as part of a more social phenomenon called *collective events* [11].

### B. Event Detection from Multimedia

A large number of approaches deal with the detection of event from text [15, 22-26]. Recently, numerous studies have investigated the problem of event detection from images. They

can be classified in two groups: event clustering approaches [5-11] and event hybrid approaches [12, 14, 17, 21, 27, 28]. Extracting events from multimedia in terms of photographs or images is much more difficult when compared to text for essentially two reasons: i) Event detection from images requires aggregation of heterogeneous metadata [29]; ii) Linking multimedia data to event model aspects is far more challenging than textual data [30]. In fact, many aspects of an event should be taken into consideration, as described in the multimedia event model presented in [13], such as time, space, actors, granularities, sub-events, etc. In response to this challenge, we focus on methods for the detection of events from images shared on online social networks.

### C. Event Detection in Online Social Networks

New approaches have emerged in the past few years in the area of event detection from images on online social networks. The authors of [11, 22] use the context allied with social media content for this aim. The authors of [22] propose a method for grouping together photographs of similar events using significant terms for each event from Wikipedia and Google. The authors of [11] use the time-space-visual clustering to find the home location or locations of the user. In a subsequent work [31], the authors seek to group personal events coming from different users into richer social events based on two metadata features (time, space). Specifically, they show how event detection can bring together users who participated in the same event and hence propagate social connections among users. The authors of [5] investigate photos posted on social media sites to detect social events. First, they represent the social media and their metadata (time-stamp, location, visual content and text) into a star-structured  $K$ -partite graph. They model relationships between social media and relationships between metadata sets. Then they apply the co-clustering method on the star structure. The authors of [32] detect events by temporally monitoring the social media sharing activity at specific locations. They use metadata attached to photos (tags, description, and geographic location) to enrich the event dataset and infer the topic of events.

None of these works shares a common notion of events. Events are regarded as additional pieces of some metadata of the particular media type managed. In our work, we present an event model that can be adapted to any type of media. The limitations of these approaches with respect to our goal are as follows:

- Insufficiency in exploring all photo metadata: Most of these approaches do not include all the creation metadata (e.g., creator name). Although some approaches consider visual content [5, 11, 32], they extract only low-level features such as color. Opposed to these methods, our proposed approach is able to exploit objects found in images, particularly faces of persons.
- Lack of event linking: Existing approaches do not consider relationships between events neither the multi-\* property of an event (e.g., *multi-day* events, *multi-site* events, *multi-person* events, etc.).
- Restriction of the image extraction to the user's personal profile: Most of the existing approaches [5, 11, 21, 32] limit the personal space to the user's profile.

They do not exploit how photos of others can include information about the user and his events.

Another problem that we address in our approach is the space sparsity. The location data may be over-close in the photo collection (e.g., when almost photos are taken in one city) or too scattered (e.g., when the user frequently travels to different places around the world). We also consider the frequency of photo publication on social networks. Actually, some users tend to share only a few images (under-exposure) while others tend to over-share image data on social media (over-exposure). Another crucial difference that distinguishes our work from other prior related work is that we define rules for the multi-\* property of event.

### III. MOTIVATING SCENARIO

We use a real scenario that motivates the problem. Being an active user on Facebook,  $u_0$  frequently posts photos taken at different events. Her current residence is in the city of *Biarritz* in southern France, as indicated in her Facebook profile. She has four hundreds of connections. Some are family members, friends or colleagues. Others are acquaintance. In our study, we will focus on photos of the last trip she made to Stockholm with 12 of her colleagues in May 2014. Eight of her colleagues ( $u_1, \dots, u_8$ ), who are also in her Facebook social network, have posted photos of the same trip. The departure was from Biarritz airport. In the first case, we show the set of photos captured by  $u_0$  over the four days of the trip. In the following cases, we show how this event was captured by different users over the trip. In case 5, we show the presence of new faces in the photos of  $u_0$ .

#### Case 0: Multi-day trip event

$u_0$  has photos in Stockholm over the four days (Fig. 1).



Fig. 1. Some photos of  $u_0$  taken in Stockholm over the four days of the trip showing her colleagues (2<sup>nd</sup> and 4<sup>th</sup> photo) and her old friends (3<sup>rd</sup> photo)

#### Case 1: Event started and ended in Biarritz

$u_1$  took a photo at the airport before take-off. During the trip days,  $u_1$  took many photos in Stockholm with  $u_0$ . There are also other photos of  $u_1$  with  $u_5$ . We note that  $u_5$  did not take any photo with  $u_0$ .  $u_1$  took a selfie on the return trip to Biarritz.

#### Case 2: Event started in Stockholm and ended in Biarritz

$u_2$  started taking photos when she arrived to Stockholm. She took many photos with her friends including  $u_0$ .  $u_2$  took one photo of three other persons ( $u_6, u_7, u_8$ ) alone. In another photo, we notice also the presence of  $u_{10}$  for the first time ( $u_{10}$  is not part of the social network of  $u_0$ ). On the return trip,  $u_2$  took a group photo in Biarritz airport with some of her friends.

#### Case 3: Event started in Biarritz and ended in Stockholm

$u_3$  took the first photo in the plane before the flight, and then many photos in Stockholm with others, except with  $u_0$ .  $u_3$  missed taking photo when he was back to Biarritz.

#### Case 4: Event started and ended in Stockholm

$u_4$  posted only photos that she took in Stockholm with  $u_0$  and other users. These users did not take any photos during the trip, and therefore they had posted nothing online.

#### Case 5: Sub-event meeting with old friends

$u_0$  gathered with old friends at their place in Stockholm and took photos with them, while her colleagues were not with her.

This scenario shows how image metadata plays a role in providing the duration of the event, tracking the related sub-events and the event participants. Photos created by  $u_0$ 's friends in Stockholm during the trip, revealed other users who also were present, even if they were not present on any photo created by  $u_0$ . Some of the users posted photos of the trip while others did not at all. The main challenges that arise from this scenario are:

- Grouping together photos of the trip event even though they were created by different users, over many days and in several places,
- Considering the different ways of each trip member to capture the event photos,
- Identifying all participants whether they are in the social network of  $u_0$  or not, and whether they posted about the event or not,
- Merging sub-events during the trip within one event.

## IV. PROPOSAL

### A. Background

#### 1) Data Model

##### a) User

Let  $U$  denote a set of users in an online social network. Formally, each user  $u \in U$  is described as:

$$u: (user_{id}, pf, I) \quad (1)$$

where:

- $user_{id}$ : is the identifier of the user in the online social network (usually the user's *full name*),
- $pf$ : is the set of attributes describing the user's personal information (such as birthday, hometown, etc.),
- $I$ : is the collection of his photos posted to  $SSN_{u_0}$ .

##### b) Star Social Network

In our study, the social network of a user  $u_0$  is modeled as a star graph composed of the set of users directly connected to  $u_0$ . Formally, we define the star social network of  $u_0$  as follows:

$$SSN_{u_0}: (U, L) \quad (2)$$

where:

- $U$ : is a set of users connected to  $u_0$  in the online social network. The collection of photos is restricted to those shared with and published by  $u_0$ ,
- $L$ : is the set of links (e.g., *friends, colleagues, relatives, etc.*) that exist between  $u_0$  and  $U$ .

### c) Image Metadata

A metadata record consists of a set of attributes describing image creation and content. It is generally incorporated into the image data file (e.g., EXIF or Extensible Metadata Platform (XMP) format, etc.) but could also be described separately to fill missing attributes' values (RDF, MPEG-7, etc.). We formally represent a metadata record of an image  $img$  as:

$$meta_{img}: (c, g, t, O) \quad (3)$$

where:

- $c$  represents the creator of the image,
- $g$  represents the geo-location of the image when captured (i.e., latitude and longitude),
- $t$  represents the creation date/time of the image,
- $O$  represents a set of objects of interest identified in the image (i.e. person faces in our case).

### d) Image

Let  $img \in I$  be an image or photo posted by a user  $u_i$  to  $SSN_{i0}$ . Formally, we represent an image as follows:

$$img: (img_{id}, meta) \quad (4)$$

where:

- $img_{id}$ : is the identifier of the image (e.g., URI),
- $meta$ : is the metadata record describing the image.

$img$  is called *story image* or  $s_{img}$  if its metadata contain at least one face in  $O$  or indicate a place in  $g$  (i.e.,  $g$  is not empty) within the captured scene.  $img$  is called *trigger image* or  $t_{img}$  if its metadata contain one or more particular faces or indicate a significant place that cause an event to happen.  $img$  is called *break image* or  $b_{img}$  if its metadata shows the end of the event (e.g., a back-home photo). For example, Fig. 2 shows a story image with its metadata.  $O$  encloses all the faces identified in the image (represented here by square brackets).



Creator	François Fillon
GPS Position	47.849333 degrees N, 0.249333 degrees W
Create Date	2011:10:31 12:55:27.39

Fig. 2. Example of a story image representation with its metadata

### e) Event

Let an event  $e$  represent a personal story or adventure in a user's life. An event can be represented using one or several images. Formally, an event  $e_{id}$  is defined as:

$$e: (e_{id}, B, F) \quad (5)$$

$B$ : is the image board of the event. It includes an ordered set (sorted according to date/time) of images: a trigger image,

several story images representing the event course, and a potential break image, as illustrated in Fig. 3. We represent the event board  $B$  as a regular expression such as:

$$B: t_{img} s_{img}^* b_{img}^? \quad (5.1)$$

where:

- $t_{img}$  is the event trigger,
- $s_{img}^*$  is the set of story images during the event course,
- $b_{img}^?$  is the potential event break.



Fig. 3. Illustration of an event board

$F$ : is a set of features characterizing the group of users, locations and times in the event. It is represented as:

$$F: (Who, Where, When) \quad (5.2)$$

where:

- $Who$ : is the set of users  $who$  took part in the event  $e$ . In other terms,  $Who = c_B \cup O_B$  where  $c_B$  is the set of creators and  $O_B$  is the set of users identified<sup>10</sup> in the image event board,
- $Where$ : is the set of geographical positions of the images  $g_B$  where the event took place, i.e.,  $Where = g_B$ ,
- $When$ : is the set of creation dates of the images  $t_B$  representing the event period, i.e.,  $When = t_B$ .

The 3W features can be visualized in a 3D graph (Fig. 4).

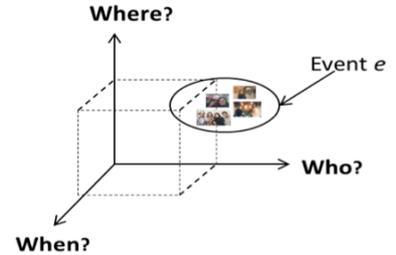


Fig. 4. Representation of event features in a 3D graph

### f) Event Relationships

In our approach, two events  $e_i$  and  $e_j$  are always related to each other with three types of relationships: temporal, spatial, and semantic. This is represented as:

$$R(e_i, e_j): \langle tmp, SPA, SEM \rangle \quad (6)$$

where:

- $tmp$ : is an exclusive temporal relationship that  $e_i$  has with  $e_j$ . We adopted here Allen's relations [33]: *Before, After, Meet, Started by, Overlapped By, Start With, Finish, During, Equal, etc.*

<sup>10</sup> Here, we assume that known faces in photos are already annotated manually or automatically.

- *SPA*: is the set of spatial relationships [34] linking the two events such as:

$$SPA: \langle \text{metric}, DIRECT, \text{topologic} \rangle \quad (6.1)$$

where:

- *metric*: is a metric relation that measures the distance between events. For instance, the metric relation “far” between two events  $e_i$  and  $e_j$  indicates that each pair of points of the segment linking  $e_i$  and  $e_j$  has a distance greater than a predefined threshold.
- *DIRECT*: is a set of directional relations describing the order between  $e_i$  and  $e_j$  according to a direction. In the literature [35], 14 directional relations are considered (e.g., *north*, *south*, *east*, *west*, *left*, *right*, *down*, etc.).
- *topologic*: is an exclusive topological relation that describes the intersection and the incidence between events. We use here the 6 basic relations commonly adopted in the literature [36]: *Disjoint*, *Meet*, *Overlap*, *Cover*, *Contain*, and *Equal*.
- *SEM*: contains the set of semantic and social relationships between the two events. These relationships can express memberships (e.g., *same-club*, *same-friends*, *same-colleagues*, etc.), kinships (e.g., *mother-of*), causality (e.g., *related to*, *due to*, etc.), among others. More details can be found in [37].

## 2) General Concepts

We define four general concepts. We discriminate a place as well as a face in two categories: *common* and *uncommon*.

- **Common place (CP)**: represents a location where  $u_0$  spends the most of his time. This information is provided in the user profile (e.g., home and work locations).
- **Uncommon place (UP)**: represents a location that is away from CP. It is determined based on the metric distance of geographical locations ( $\Delta g_i = g_i - g_{CP}$ ).
- **Common face (CF)**: belongs to a person that often appears in the photos of  $u_0$  or posts photos containing the face of  $u_0$ . We use Jaccard coefficient [38] to capture the degree of co-occurrence of a given person and  $u_0$ .
- **Uncommon face (UF)**: belongs to a person that has only one occurrence in the same photo with  $u_0$  or has repeated occurrences over a specific period.

## B. Event Detection Algorithm

We propose a clustering method based on image metadata to cluster images created by the same user or by different users. Clustering is used to discover personal events based on images shared on social networks. An overview of our Event Detection algorithm is shown in Fig. 5. The suggested algorithm is composed of two steps: first, to define time-space granularities; and second, to group images. A refinement step is proposed to refine clusters by considering the multi-\* property of events. By doing so, we are able to find *multi-day* and *multi-site* events instead of considering them as independent events. Moreover, we are enriching events by aggregating images from multiple persons who participated in the same event.

The input is  $SSN_{u_0}$ , the social network of  $u_0$ , for which we are interested in detecting personal events. The input comprises the user profile information and two sets:  $U$  and  $I$ , which are the set of users connected to  $u_0$  and the set of images created by  $U$ , respectively. The parameter *MinPers* is also given as input to the algorithm. The output is the set of events of  $u_0$  linked to each other by some identifiable relationships. In the following, we present our approach in more details.

### 1) Time-Space Granularities

Our approach takes into consideration the metadata distribution of images in  $SSN_{u_0}$  in order to produce a personalized clustering for images. Therefore, we select time-space granularities based on the collection of images at hand.

#### a) Time Granularity

We construct a function to group images according to their date of capture. We focus on the frequency of posting photos to verify whether photos are available in a specific period of time, or indefinitely over time. The time granularity can be year, month, day, hour, etc. Therefore, we measure the time difference between images. If the time difference is short (less than one day), the time granularity is specified to *half-day*. If the time difference is about one day, the time granularity is specified to *full day*. If the time difference is not in the range of days, but in the range of months, the time granularity is specified to a *month*. Otherwise, if the time difference is larger than one month, we use the granularity level *year* to create time partitions from the image collection.

#### b) Space Granularity

We construct a function to classify images based on their geographical coordinates. Locations can vary at different levels of granularity: *nationally*, *regionally*, *locally* or more precisely *within the same city*. We study first the distribution of locations across countries. If locations are found in different countries, the location’s granularity is very high (*country*). Otherwise, if the same country is found in the image collection, we move to the *regional* level. If locations are in the same region, we examine the *city* level. Finally, if the set of locations are mapped to the same city name, the location’s granularity is set as *street*.

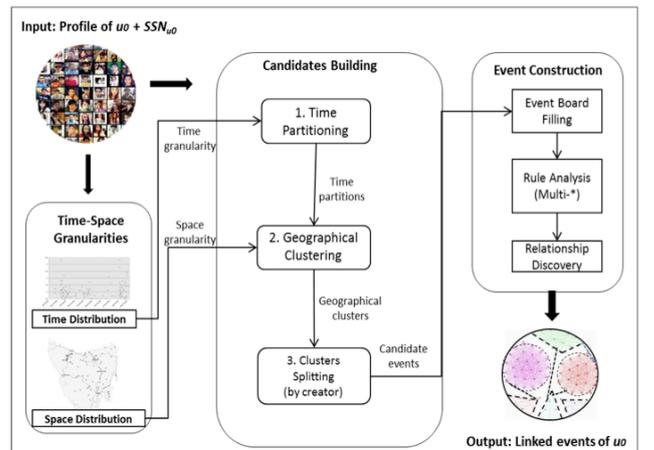


Fig. 5. Description of our approach to detect, refine and link events

## 2) Image Grouping

### a) From Images to Time Partitions

We use the date/time information to form small image sets using the chosen time granularity. The result is a set of partitions, each containing images created at the same time. Let  $t_{Simg_i}$ ,  $t_{Simg(i+1)}$  be the creation date/time of two successive chronologically ordered images. A time partition is defined as:

$$p = \{Simg_i\} / \forall i, t_{Simg(i)} = t_{Simg(i+1)} \quad (7)$$

### b) From Time Partitions to Geographical Clusters

We automatically determine image location by using reverse geocoding<sup>11</sup> with the given GPS-coordinates. For example, reverse geocoding of Latitude 43.459694, Longitude -1.545815 would provide *Biarritz* as a city address. Next, we divide each time partition into geographical clusters according to the chosen space granularity (e.g., country, region, city or street). The result of this step is a set of clusters, each containing images which are created during the same time span and in the same space. A cluster is represented as follows:

$$k = \{Simg_i\} / \forall i, t_{Simg(i)} = t_{Simg(i+1)} \text{ and } location(g_{Simg(i)}) = location(g_{Simg(i+1)}) \quad (8)$$

### c) Splitting Geographical Clusters

Geographical clusters are divided into groups according to their creators. Since a user can be only at one place at a time, images captured by one user at a specific time within a geographical boundary are put in one group. The result is a set of clusters; each cluster  $gr$  includes images created by user  $c$  on a common time span and in the same space.

$$gr = \{Simg_i\} / \forall i, t_{Simg(i)} = t_{Simg(i+1)} \text{ and } location(g_{Simg(i)}) = location(g_{Simg(i+1)}) \text{ and } c_{Simg(i)} = c_{Simg(i+1)} \quad (9)$$

### d) Building Candidate Events

The clusters obtained from the previous step are considered as the candidate events. We determine the event trigger, course and break for each cluster generated to produce *events*. Let  $E_{u_0} = \{e_1, e_2, \dots, e_n\}$  be the set of events obtained so far and ordered by time of their event triggers. We create bag-of-users for each event, by counting once each participant in the event.

## 3) Event Refinement

Refinement is necessary to address some problems: (i) many persons may participate in the same event, take photos of each other's, and upload them, but each in different accounts. (ii) One event can be scattered in different locations (e.g., a multi-city trip). (iii) One same event can persist for many days (e.g., a weekend conference). Therefore, detecting *multi-person* events, *multi-site* events and *multi-day* events is crucial for effective event detection. In order to decide which events should be merged into one, we start by comparing clusters in terms of time and space.

### a) Rule 1. Multi-person events

When two events  $e_i$  and  $e_j$  take place exactly in the same time span and location, we infer that they are two clusters of images taken at the same event by different persons.

$$(When(e_i) = When(e_j)) \wedge (Where(e_i) = Where(e_j)) \rightarrow SameAs((e_i), (e_j))$$

### b) Rule 2. Multi-site events

When two events  $e_i$  and  $e_j$  take place in the same time, but are separated in space, we measure the similarity between the users participating in each event.  $MinPers$  is the minimum number of persons required to assume that the two events have common participants. We set  $MinPers$  to 1. Hence, if  $e_i$  and  $e_j$  have at least one user in common, we infer that they are clusters of images taken at the same event occurring in multiple locations.

$$(When(e_i) = When(e_j)) \wedge (Who(e_i) \cap Who(e_j) \geq MinPers) \rightarrow SameAs((e_i), (e_j))$$

### c) Rules 3 and 4. Multi-day events

When two events  $e_i$  and  $e_j$  are separated in time and space, we define rules based on the event triggers to decide whether the two events are part of the same event (e.g., a multi-day trip). Before performing this step, we discard events where  $u_0$  is not a participant. In fact, those events are interesting to the  $u_0$ 's contacts rather than to  $u_0$ . We call them "social events". What we can learn about  $u_0$  from his social events will be the object of a future work.

In best-case scenarios, an event is triggered immediately by an uncommon place (*UP*). Conversely, when no uncommon place is detected, other data are required to trigger the event. For example, the presence of an uncommon face (*UF*) can cause an event to occur even in a common place. Accordingly, rules can be defined based on the event triggers of two successive chronologically ordered clusters. Let  $e_i$  and  $e_{i+1}$  denote two consecutive candidate events. We merge  $e_i$  and  $e_{i+1}$  when one of the rules described in Table 1 is satisfied.

TABLE I. RULES FOR MERGING EVENTS

	Description	Rules
Rule 3	$e_i$ and $e_{i+1}$ have the same <i>UP</i>	$(Where(e_i) \cap Where(e_{i+1}) = UP) \rightarrow SameAs(e_i, e_{i+1})$
Rule 4	$e_i$ and $e_{i+1}$ have the same <i>UF</i>	$(Who(e_i) \cap Who(e_{i+1}) = UF) \rightarrow SameAs(e_i, e_{i+1})$

## V. EXPERIMENTAL EVALUATION

### A. Prototype

We implemented a prototype, entitled *Foto2Events*, to test, evaluate and validate our event detection algorithm on real and generated data. Our prototype, implemented in Java, has also four components for generating simulated events:

- The *creator* component generates person names which consist of a first name and a surname,
- The *participant* component generates a group of person names corresponding to persons identified in images,
- The *location* component generates a set of geographic coordinates in a specific geographic location,
- The *date* component generates a set of dates within an interval.

<sup>11</sup><https://developers.google.com/maps/documentation/javascript/examples/geocoding-reverse>

### B. Evaluation Metrics

In order to evaluate the effectiveness of our approach, we make use of the entropy measure which looks at how the various events are distributed within each cluster. Following Nayak [39], for an obtained cluster  $C_i$  in the clustering solution, the entropy of a cluster is defined as:

$$E(C_i) = \frac{1}{\log k} \sum_{r=1}^k \frac{n_i^r}{n_i} \log \frac{n_i^r}{n_i} \quad (10)$$

where  $k$  is the number of events generated, and  $n_i^r$  is the number of records of the  $r^{th}$  event that are assigned to the  $i^{th}$  cluster. The entire clustering solution entropy is the sum of the individual cluster entropies weighted according to the cluster size. The formula is given below.

$$Entropy = \sum_{i=1}^k \frac{n_i}{n} (C_i) \quad (11)$$

A perfect clustering solution will be the one that leads to clusters that contain images from only a single event, in which case the entropy will be zero. In general, smaller the entropy value, better the clustering solution is.

### C. Real Dataset

The first experiments use the real scenario described in Section 3. The dataset is built from the Facebook photo collections of 10 users from the UPPA University (<http://www.univ-pau.fr/live/>) who went for a trip during *four days to Stockholm*, Sweden. For each participant, we asked him/her to collect all photos taken during this event and posted to Facebook. This collection (of 1813 photos) will serve as the reference solution. The goal is to test whether all photos captured by the 10 users, in different places and during the four days, were grouped into one cluster.

First, we used the location-tagging, time and date of images' post, and person tags in images as the image metadata. We used *day* as the time granularity and *street* as the space granularity. Then, we applied our event detection algorithm on the social network of one of the participants. We focused only on clusters containing one or more images from the Stockholm trip. For each step of the algorithm, we reported the number of clusters generated and calculated the purity of clusters which is the average of the percentage of images that belong to this trip in each cluster.

Results yields a high number of clusters compared to the expected number of clusters (one cluster) if no refinement is performed. For the same event, we obtain 4 time partitions after *Step 1*, and 42 geographical clusters after *Step 2*. The highest number of clusters is reached after *Step 3* (99 groups). The number of clusters decreased from 48 to 10 after applying *Rule 1* and *Rule 2*. The best clustering result is obtained after applying *Rules 3* and *4*. The interpretation of this result is that when the uncommon place (Stockholm city) was found in the last clusters, the merging was computed. The purity of clusters increased from 33% after *Step 2* to 95% after *Step 3*. The purity of clusters remained nearly unchanged after applying rules. Results indicated that our approach ended up with a large collection of photos taken during a multi-day trip by different persons. Our approach allowed the identification of all the trip members through our rules.

### D. Simulated Dataset

Now we wish to test our algorithm on different data distributions. So first, we generated our dataset of events using the generation components described previously. Each generation requires as input a *creatorList*, a *participantList*, a *LocationList*, and a *DateList*, and corresponds to one event. Then, we used *day* as the time granularity, and *city* as the space granularity and applied our event detection algorithm. We evaluated the clustering quality performance in terms of entropy. Table 2 shows the sets of event simulation performed. The generation parameters are described as follows: Images within one event can be created by one or more users (Column 2). Images within one event can have common person tags (Column 3): Zero denotes the absence of common persons and one denotes the presence of common persons. An event may take place in a single location or in several locations (Column 4) and may take one or several days (Column 5). At last, we add noise images that do not correspond to any event specification.

TABLE II. SETS OF SIMULATION PERFORMED

Cases	Creator	Participants	Location	Date
1	1	1	1	1
2	1	0	1	1
3	1	1	1	+
4	1	0	1	+
5	1	1	+	1
6	+	1	+	1
7	1	0	1	1
8	+	0	1	1
9	1	1	+	+
10	+	1	+	+
11	1	0	+	+
12	+	0	+	+

Compared to the reference solution, our algorithm was able to group most images belonging to the same event when there is a single creator for each event (Cases 1-4, 5, 7, 9 and 11). Cases 5, 6, and 8 clearly show the benefits of applying *Rule 1* and *Rule 2*, whereas these rules could not be applied to Cases 3, 4, 10, 11 and 12. Additionally, considering similarity of participants was of great importance in the case of multi-site events (Cases 5, 7 and 9).

## VI. CONCLUSION

We presented a clustering algorithm for automatically detecting personal events from photos shared online on the social network of a specific user. We defined an event model that captures the event trigger, course and break, and relationships that can exist between events. In addition, we have defined rules based on the event features (*Who*, *Where*, *When*) for refining cluster results. Using the appropriate time-space granularities, we are able to detect *multi-site*, *multi-day* and *multi-person* events. Our finding suggests that more information about events can potentially be identified by mining, refining and linking images posted to social networks. As a future work, we are planning to use the result clusters of personal events with inference channels to predict if a given image is able to infer the presence of a user in events when the user did not intend to share this information online.

## REFERENCES

- [1] B. A. Bouna, E. J. Raad, C. Elia, R. Chbeir, and R. Haraty, "de-linkability: a privacy-preserving constraint for safely outsourcing multimedia documents," presented at the Proceedings of the Fifth International Conference on Management of Emergent Digital EcoSystems, Luxembourg, Luxembourg, 2013.
- [2] M. Hirota, N. Fukuta, S. Yokoyama, and H. Ishikawa, "A Robust Clustering Method for Missing Metadata in Image Search Results," *Journal of Information Processing*, vol. 20, pp. 537-547, 2012.
- [3] L. Backstrom and J. Kleinberg, "Romantic partnerships and the dispersion of social ties: a network analysis of relationship status on facebook," presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, Baltimore, Maryland, USA, 2014.
- [4] M. Saini, P. Atrey, S. Mehrotra, and M. Kankanhalli, "W3-privacy: understanding what, when, and where inference channels in multi-camera surveillance video," *Multimedia Tools and Applications*, vol. 68, pp. 135-158, 2014/01/01 2014.
- [5] B.-K. Bao, W. Min, K. Lu, and C. Xu, "Social event detection with robust high-order co-clustering," presented at the Proceedings of the 3rd ACM conference on International conference on multimedia retrieval, Dallas, Texas, USA, 2013.
- [6] L. Chen and A. Roy, "Event detection from flickr data through wavelet-based spatial analysis," presented at the Proceedings of the 18th ACM conference on Information and knowledge management, Hong Kong, China, 2009.
- [7] C. S. Firan, M. Georgescu, W. Nejdl, and R. Paiu, "Bringing order to your photos: event-driven classification of flickr images based on social knowledge," presented at the Proceedings of the 19th ACM international conference on Information and knowledge management, Toronto, ON, Canada, 2010.
- [8] S. Papadopoulos, C. Zigkolis, Y. Kompatsiaris, and A. Vakali, "Cluster-Based Landmark and Event Detection for Tagged Photo Collections," *MultiMedia, IEEE*, vol. 18, pp. 52-63, 2011.
- [9] S. Rafatirad, R. Jain, and K. Laskey, "Context-Based Event Ontology Extension in Multimedia Applications," in *Semantic Computing (ICSC), 2013 IEEE Seventh International Conference on*, 2013, pp. 278-285.
- [10] M. Ruocco and H. Ramampiaro, "A scalable algorithm for extraction and clustering of event-related pictures," *Multimedia Tools and Applications*, vol. 70, pp. 55-88, 2014.
- [11] I. Tankoyeu, J. Paniagua, J. Stottinger, and F. Giunchiglia, "Event detection and scene attraction by very simple contextual cues," presented at the Proceedings of the 2011 joint ACM workshop on Modeling and representing events, Scottsdale, Arizona, USA, 2011.
- [12] M.-S. Dao, D.-T. Dang-Nguyen, and F. G. De Natale, "Robust event discovery from photo collections using Signature Image Bases (SIBs)," *Multimedia Tools and Applications*, pp. 1-29, 2012.
- [13] U. Westermann and R. Jain, "Toward a Common Event Model for Multimedia Applications," *MultiMedia, IEEE*, vol. 14, pp. 19-29, 2007.
- [14] Y. Wang, H. Sundaram, and L. Xie, "Social event detection with interaction graph modeling," presented at the Proceedings of the 20th ACM international conference on Multimedia, Nara, Japan, 2012.
- [15] X. Liu, X. Zhou, Z. Fu, F. Wei, and M. Zhou, "Exacting Social Events for Tweets Using a Factor Graph," in *AAAI*, 2012.
- [16] T. Rattenbury, N. Good, and M. Naaman, "Towards automatic extraction of event and place semantics from flickr tags," presented at the Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, Amsterdam, The Netherlands, 2007.
- [17] Z. Xu, X. Wei, X. Luo, Y. Liu, L. Mei, C. Hu, and L. Chen, "Knowle: A semantic link network based system for organizing large scale online news events," *Future Generation Computer Systems*, 2014.
- [18] D. Mahata and N. Agarwal, "Learning from the crowd: an evolutionary mutual reinforcement model for analyzing events," presented at the Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Niagara, Ontario, Canada, 2013.
- [19] A. D. Sarma, A. Jain, and C. Yu, "Dynamic relationship and event discovery," presented at the Proceedings of the fourth ACM international conference on Web search and data mining, Hong Kong, China, 2011.
- [20] G. Valkanas and D. Gunopulos, "How the live web feels about events," presented at the Proceedings of the 22nd ACM international conference on Conference on information & knowledge management, San Francisco, California, USA, 2013.
- [21] S. Sheba, B. Ramadoss, and S. R. Balasundaram, "Event Detection Refinement Using External Tags for Flickr Collections," in *Intelligent Computing, Networking, and Informatics*. vol. 243, D. P. Mohapatra and S. Patnaik, Eds., ed: Springer India, pp. 369-375, 2014.
- [22] H. Abdelhaq, C. Sengstock, and M. Gertz, "EvenTweet: online localized event detection from twitter," *Proc. VLDB Endow.*, vol. 6, pp. 1326-1329, 2013.
- [23] A. Bauer and C. Wolff, "Event based classification of Web 2.0 text streams," *arXiv preprint arXiv:1204.3362*, 2012.
- [24] H. Becker, M. Naaman, and L. Gravano, "Beyond Trending Topics: Real-World Event Identification on Twitter," *ICWSM*, vol. 11, pp. 438-441, 2011.
- [25] S. Kumar, H. Liu, S. Mehta, and L. V. Subramaniam, "From Tweets to Events: Exploring a Scalable Solution for Twitter Streams," *arXiv preprint arXiv:1405.1392*, 2014.
- [26] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller, "Twitinfo: aggregating and visualizing microblogs for event exploration," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 2011.
- [27] M. Brenner and E. Izquierdo, "Social event detection and retrieval in collaborative photo collections," presented at the Proceedings of the 2nd ACM International Conference on Multimedia Retrieval, Hong Kong, China, 2012.
- [28] M. Zaharieva, M. Zeppezauer, and C. Breiteneder, "Automated social event detection in large photo collections," presented at the Proceedings of the 3rd ACM conference on International conference on multimedia retrieval, Dallas, Texas, USA, 2013.
- [29] W. Yan, D. Kieran, S. Rafatirad, and R. Jain, "A comprehensive study of visual event computing," *Multimedia Tools and Applications*, vol. 55, pp. 443-481, 2011/12/01 2011.
- [30] A. Scherp and V. Mezaris, "Survey on modeling and indexing events in multimedia," *Multimedia Tools and Applications*, vol. 70, pp. 7-23, 2014.
- [31] J. Paniagua, I. Tankoyeu, J. Stottinger, and F. Giunchiglia, "Social events and social ties," presented at the Proceedings of the 3rd ACM conference on International conference on multimedia retrieval, Dallas, Texas, USA, 2013.
- [32] X. Liu, R. Troncy, and B. Huet, "Using social media to identify events," presented at the Proceedings of the 3rd ACM SIGMM international workshop on Social media, Scottsdale, Arizona, USA, 2011.
- [33] J. F. Allen, "Maintaining knowledge about temporal intervals," *Commun. ACM*, vol. 26, pp. 832-843, 1983.
- [34] R. Chbeir, Y. Amghar, and A. Flory, "High Expressive Spatio-Temporal Relations," in *FLAIRS Conference*, pp. 471-475, 2002.
- [35] R. Chbeir and F. Favetta, "A global description of medical imaging with high precision," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 33, pp. 752-757, 2003.
- [36] M. J. Enehofer, "Spatial-Query-by-Sketch," in *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pp. 60-67, 1996.
- [37] E. Raad, R. Chbeir, and A. Dipanda, "Discovering relationship types between users using profiles and shared photos in a social network," *Multimedia Tools and Applications*, vol. 64, pp. 141-170, 2013.
- [38] J. Mori, Y. Matsuo, and M. Ishizuka, "Finding user semantics on the web using word co-occurrence information," in *Proceedings of the International Workshop on Personalization on the Semantic Web (PersWeb05)*, 2005.
- [39] R. Nayak, "Fast and effective clustering of XML data using structural information," *Knowl. Inf. Syst.*, vol. 14, pp. 197-215, 2008.