Social Interactions over Location-Aware Multimedia Systems

Yi Yu, Roger Zimmermann and Suhua Tang

Abstract Advancements in positioning techniques and mobile communications have enabled location-based services with a broad range of location-aware multimedia applications. Accordingly, various social multimedia data, relevant to different aspects of users' daily life, is aggregated over time on the Internet. Such location-aware multimedia data contains rich context of users and has two implications: individual user interest and geographic-social behaviors. Exploiting these multimedia landscapes helps mine personal preferences, geographic interests and social connections, and brings the opportunities of discovering more interesting topics. In this chapter, we first introduce some examples of location-aware multimedia data and social interaction data. Then, we report some latest methods related to context detection and location-aware multimedia applications. We further present some analysis of geo-social data. Finally, we point out the trend in the integration of social and content delivery networks. In brief, this chapter delivers a picture of emerging geographic-aware multimedia technologies and applications, with location information as a clue.

1 Motivation and Introduction

Conventionally, content sharing websites [1] and online social networks [2] are separately deployed. Users visit content sharing websites to upload, view, and share

Yi Yu
School of Computing, National University of Singapore, Singapore. e-mail: yuy@comp.nus.edu.sg

Roger Zimmermann
School of Computing, National University of Singapore, Singapore. e-mail: rogerz@comp.nus.edu.sg

Suhua Tang
Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan. e-mail: shtang@uec.ac.jp

their multimedia contents. Users login to social networks to exchange messages and keep in contact with their friends. Recently, various online communities (e.g., Flickr, Foursquare, Facebook, Twitter) have started to provide users with location-based services [3]. In this way, users can record and upload geo-tagged images and videos to these web sites anytime and anywhere with their mobile devices. For example, many fantastic geo-tagged photos taken by users are shared at Flickr. As a result, every day a huge volume of user-generated geo-tagged multimedia data is generated in the Internet.

Location, as an extra information, is playing an important part in complementing content retrieval and recommendation [4, 5, 6, 7, 8, 9]. As shown in Fig. 1, it also serves as an important element to connect content sharing services and online social services, which facilitates personalized, localized, and socialized multimedia content discovery, retrieval, recommendation, and diffusion across diverse user-generated multimedia datasets. In particular, registered users can check in[1] at various venues and contact their friends nearby to share experience with them. Geographic trajectories of users are associated with their preferences and can be used for personalized location recommendation [10]. Check-in information at business venues can be leveraged for geo-fencing services [11, 12, 13, 14], mobile advertising [15, 16], business analytics, and used to analyze the geo-spatial distribution of users and user social behaviors.

A new trend is the integration of social networks and content sharing platforms [4], as follows: Users share their opinions of multimedia contents or recommend multimedia contents on social networking platforms; This helps to spread multimedia contents and events all over the world through the social connections between users [17]; In addition, the geo-spatial distribution of users and social connections between users can be further exploited to optimize the distributed cache [18] of multimedia contents.

The rest of this chapter addresses different parts in Fig. 1. First, we introduce different location-aware media data in Sec. 2. Then, we show some methods related to location inference and geo-fencing in Sec. 3. Next, we present some location-aware multimedia applications in Sec. 4 and the analysis of geo-social data in Sec. 5. We also discuss the integration of social networks and content-sharing networks in Sec. 6. Finally, we conclude this chapter with Sec. 7.

2 Geo-Tagged Multimedia Data on Social Networks

Here, we introduce several typical examples of location-aware multimedia data, e.g., Flickr images, Foursquare check-in, Twitter messages. This demonstrates how user-centric location-aware datasets are associated with multimedia contents. Such

[1] Many social networking services allow users to self-report presence (known as check in) to a physical place and share their locations with their friends. Refer to http://en.wikipedia.org/wiki/Check-in

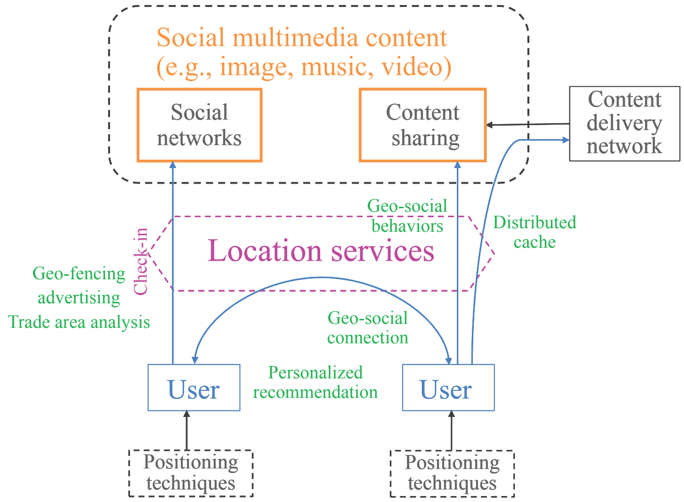


Fig. 1 Connecting social networks and content-sharing platforms via location information.

location-aware social multimedia data with geo-tags can be exploited in later sections to analyze user behaviour, especially user interest.

2.1 Geo-Tagged Photos on Flickr

Location information is important for remembering where a particular photo came from and showing off user's favorite photos to the world over a map. Online photo sharing website Flickr[2] has created the geo-tagging[3] function to let users geo-tag their photos, as shown in Fig. 2. According to the location names, these geo-tagged photos can be classified and displayed on a map.

Flickr acts as a repository of all kinds of photos together with geo-tags. Through crowdsourcing from Flickr's geo-tagged photo collections, geographic discovery can be studied to discover knowledge about different aspects of information on the surface of the Earth, for example, classifying the land-use into classes [19] of academic, sports and residential according to both images and their geo-tags.

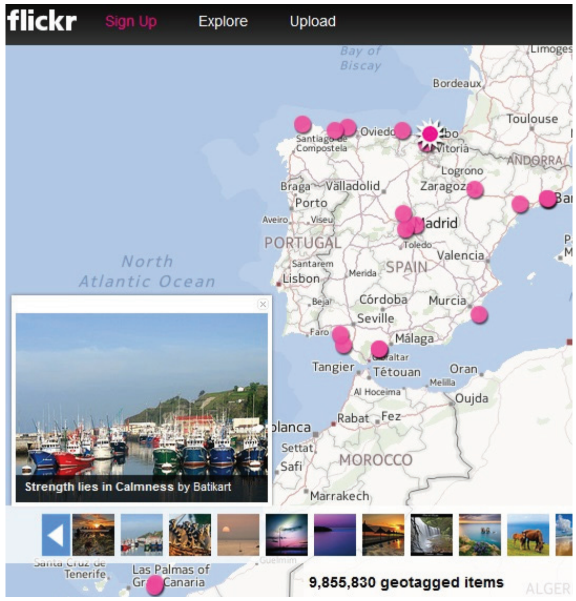


Fig. 2 Geo-tagged Flickr photos shown on the map.

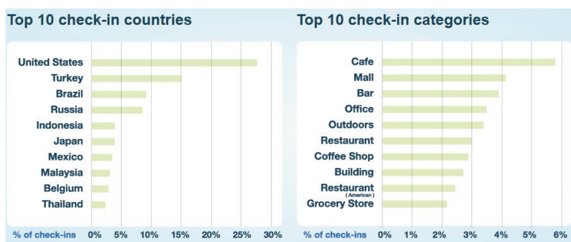


Fig. 3 Top check-in countries and categories in Foursquare, reported by http://gnip.com/foursquare/.

2.2 Geo-Social Data on Foursquare

Foursquare[4] implements a location layer for the Internet, which is an intersection of virtual social networks and physical world to help connect people with their friends around the world. In addition, Foursquare provides an API to map location information to geo-categories. Specifically, with a given location (latitude and longitude), Foursquare returns venues nearby with metadata (geo-category etc.). From

[2] https://www.flickr.com/

[3] Geo-tagging is the process of adding geographical identification metadata to various media data. Refer to http://en.wikipedia.org/wiki/Geotagging

[4] https://foursquare.com/

Fig. 3, we can see top 10 check-in countries and top 10 check-in geo-categories in Foursquare.

User-generated geographic data may be shared on social networking platforms. For example, checking-in at a venue via Foursquare, Foursquare will tell you who and what are nearby and broadcast your location to your friends and update your Twitter and Facebook status. Foursquare also can serve as a metadata of local business information. When users check-in at the stores, the check-in data provides a spatial distribution of users visiting these stores, and can be used for analyzing the primary trade areas of these stores [15]. Check-ins in Foursquare also can provide user visit information [10].

Data about the geographic positions of users can be made publicly available, together with their online social connections. For example, many Foursquare users choose to automatically push their check-ins to Twitter messages. Although Foursquare does not provide unauthorized access to user friends list, each tweet provides a URL to the Foursquare website, where information about the geographic location of the venue can be acquired. Twitter provides a public API to search and download these tweets. Then, friendship ties and location information can be acquired from tweets. These datasets are publicly available and can be used to study social and geographic networks of users.

2.3 Location-Aware Messages on Twitter

As music plays an important role in our life, users often tweet music-related topics on Twitter[5]. Through crowdsourcing in Twitter, tweets with geospatial coordinates can be leveraged for estimating artist similarity, popularity, and local music trends. In addition, geographic music listening pattern inferred from all music tweets can be visualized on an electronic map [20].

Some social media systems utilize and provide location information at various accuracy levels and run over different geographical scopes (e.g. a street, a suburb, a city, a country), and work with different social web sources (e.g. Twitter, Facebook, etc.). For example, Crisis tracker[6] is a web-based system that automatically tracks sets of keywords on Twitter, and filters stories based on location information.

3 Location and Context-Awareness

Location-based services have experienced different generations. The first generation location-based services were released around 2000 [3]. Various icons are used to represent different categories of point of interest on an electronic map. The pre-

---

[5] https://twitter.com/

[6] http://irevolution.net/2012/07/30/collaborative-social-media-analysis/

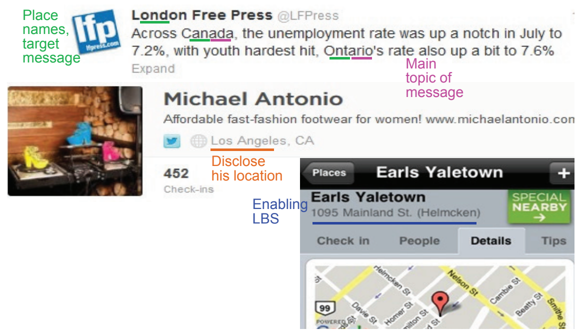


Fig. 4 Location information that can be inferred from social media data.

ferred application was the delivery of nearby points of interest (such as restaurants and bars). Advancements in positioning techniques and mobile communications have enabled the second generation location-based services with a broad range of new and sophisticated applications. Here, we mention two typical applications as examples. (i) Social community platforms like Facebook and Foursquare have enabled location sharing for the mutual exchange of location data between users. A special form of location sharing is the check-in function. It is used to explicitly acquire user locations at certain venues. (ii) Locating people to provide special offers or discounts has attracted much attention in mobile marketing. In this way, advertisers could catch the attentions of users by providing advertisements matching their needs. The area of mobile marketing is the next big thing in the mobile Internet [16]. Particularly, we explain in detail the geo-fencing application [13], which is a promising technique for user-centric mobile location-based services.

3.1 Location Inference from Social Messages

Social media messages contain different types of location information, such as place names appearing in the message, a location from which the message was sent, and so on. Four types of locations, shown in Fig. 4, can be inferred from social web data. *Location in text* is a location type for place names described in a target message (for example, London, Canada, Ontario). *Targeted location* is relevant to the main topic of the target message (for example, Canada, Ontario). *User location profile* is a location type that a user discloses in his profile (for example, Los Angeles). The user's *current location* is a location type that is obtained from location-based service in physical world (for example, 1095 Mainland St.).

When we geo-locate a message, we should consider which location type is appropriate. A framework is proposed in [21] for classifying location elements and a method for their extraction from social web data. This work is related to location inference from text messages. Usually the inputs are the messages and the outputs

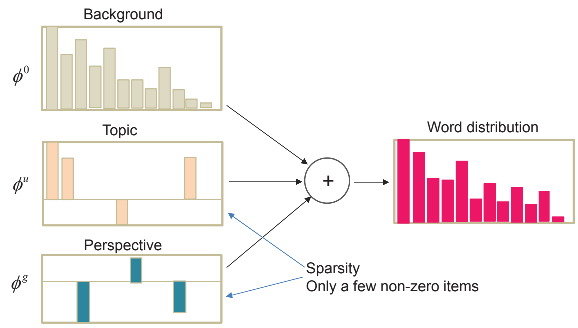


Fig. 5 Modeling term frequency in the log space.

are the locations. There are two components in the system: location name recognition and toponym resolution. The system extracts terms that are possibly location names as location candidates, and resolves whether or not they are location names in the toponym resolution component. A confidence score for each location instance is calculated by multiplying the location popularity and region context scores [21]. After these calculations, the location instance with the highest confidence score is selected as the result of toponym resolution. Finally, the detected location is assigned coordinates.

3.2 *Location Inference from Tweets via the SAGE Model*

Term distribution of tweets written by a given user depends on several factors such as user preference, region distribution and topic distribution. A user has his preferences over regions where he usually spends his time, and preferences over topics that he often tweets about. In addition, at a specific region, the tweets may contain localized keywords such as an airport, a park, a mall, a city, etc. Moreover, the content of tweets may be associated with the topics at a region and can be classified as sports, politics, travel, daily life, etc. Therefore, a tweet is composed of a bag of words from topic, region and background language models. Then, given a tweet, its location can be inferred by using these language models.

We first give some preliminaries in Fig. 5 on how to model term frequency in the log space. For a term v in a model ϕ, its term frequency is β_v, and its log frequency is defined as $\phi_v = \log \beta_v$. Then, the term distribution can be computed by normalizing β_v, and more importantly, it is equivalent to computing the term distribution directly from the log frequency ϕ_v (Eq.(1)).

$$p(v|\phi) = \frac{\beta_v}{\sum\limits_v \beta_v} = \frac{\exp(\phi_v)}{\sum\limits_v \exp(\phi_v)}. \qquad (1)$$

Now consider the sparse additive generative model (SAGE) [22], where several models ϕ^0, ϕ^u, ϕ^g are added together (Eq.(2)). Their addition in the log space is equivalent to the multiplication of term frequency. In this process, ϕ^0 is a basic reference model (β^0 is the term frequency distribution), ϕ^u is the difference between one model and the reference model (β^u is the rate by which term frequency is increased in this model), and ϕ^g is the difference between another model and the reference model.

$$p(v|\phi^0+\phi^u+\phi^g) = \frac{\exp(\phi^0_v+\phi^u_v+\phi^g_v)}{\sum\limits_v \exp(\phi^0_v+\phi^u_v+\phi^g_v)}. \tag{2}$$

The above SAGE model can be used to represent multiple facets involved in automatic generation of text messages. For example, here, use ϕ_0 to denote the log value of term frequencies of a background model. Other components, such as ϕ_u and ϕ_g, are used to describe the topic model and perspective model, which only record the difference from the background model. The SAGE model has two properties. First is sparsity-inducing for a specific model. In other words, only the difference of term frequency of a subset of terms is modeled. For example, in Fig. 5, ϕ_u and ϕ_g only have a few non-zero items. Second is to combine generative facets through simple addition in log space. For each term, the non-zeros values in all models are added together, and then normalized to get the distribution of terms.

Next, we introduce how a tweet is automatically generated using the SAGE model, based on the term distribution, regional language models, global topics, user preferences etc [23]. A tweet is generated by several steps. In the first step, using both global distribution over latent regions η_0 and user dependent distribution over latent regions η_u, a region r is drawn from the mixed region model $p(r|\eta_0+\eta_u)$. In the second step, using global distribution over topics θ_0, regional distribution over topics θ_r, and user dependent distribution over topics θ_u, a topic z is drawn from the mixed topic model $p(z|\theta_0+\theta_r+\theta_u)$. In the third step, each word w in the tweet is successively generated by drawing from the aggregate distribution $p(w|\phi_0+\phi_r+\phi_z)$, where ϕ_0 parametizes a global distribution over terms, ϕ_r describes the region-dependence of terms, and ϕ_z is a topic-specific distribution of terms.

Although Twitter provides location service, currently only 1% of tweets are geotagged (latitude and longitude). The previous tweet generation model can be used for location prediction of tweets [23], as shown in Fig. 6. Location prediction for a new tweet is based on the words used in the tweet and its user information. User information gives the user dependent distribution over latent regions (η_u). The additive model for region gives a guess of a region r from the model $p(r|\eta_0+\eta_u)$. Words in the tweet are related to regional distribution over topics (θ_r) and user dependent distribution (θ_u) over topics. On this basis, the additive model $p(z|\theta_0+\theta_r+\theta_u)$ for topic gives a guess of topic which maximizes the probability. This probability is associated with the region r. Further maximizing this probability with respect to different regions gives the most proper region for the tweet. This is a rough estimation of the location for the tweet.

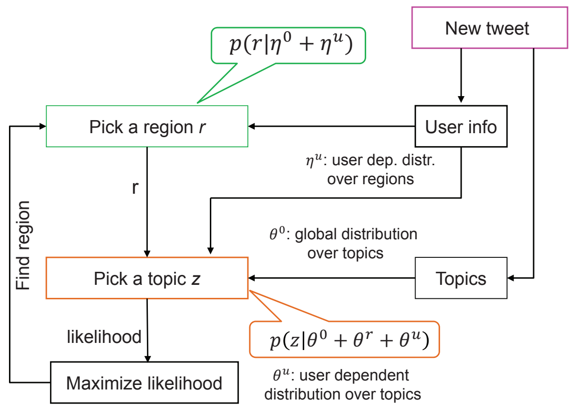


Fig. 6 Location prediction via tweets using the SAGE model.

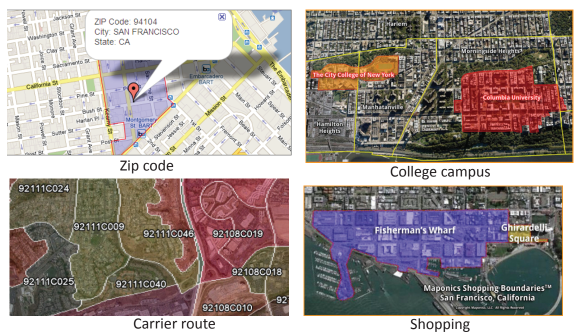


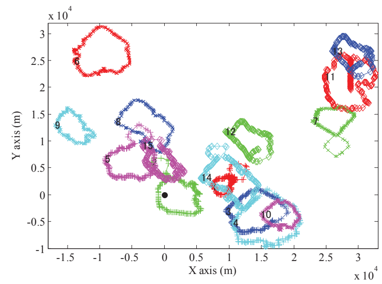
Fig. 7 Semantic geo-boundaries in real life reported from http://www.maponics.com/.

3.3 Context Awareness via Geo-fencing

More and more location-based social services want to locate, reach and interact with users on-the-go and provide various services. To this end, the geo-fencing service [13] (e.g., placecast, sensewhere, zentracker) is introduced to respond to personal user needs, and recent years have seen a growing need for user-centric geo-fencing technique in location-based services.

A geo-fence is a virtual perimeter for a real-world confined geographic area. This area can be the coverage of a particular radio cell or a Wi-Fi access point, or specified by a geographic shape. As a result, geo-fences may have different shapes, e.g., circles, rectangles, polygons, which are specified by geographic coordinates. The basic idea behind geo-fencing is very intuitive: when users enter or exit geo-fences based on geo-fencing-enabled location preferences, notifications are sent out to users or their networks of friends.

Various semantic geo-fence boundaries can be predefined to target very specific geographic areas and customers, visualize business opportunities and help to make



edges of 15 polygons (200)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
285	255	235	196	264	250	240	239	226	226	242	153, 15	152, 20	250	217

Fig. 8 Polygons, points and edges from training dataset of ACM SIGSPATIAL GIS Cup 2013.

more informative decisions. Fig. 7 shows example geo-fences corresponding to zip code boundaries, college campus boundaries, carrier route boundaries, shopping boundaries, respectively.

Geo-fencing is a big feature for user-centric location-based social networks. It mainly deals with pairing a point (a user's coordinate) with a polygon (a semantic geo-fence boundary). In other words, the task is to estimate whether a point is INSIDE or WITHIN a distance of a polygon. Each point has multiple instances each with a unique sequence number, i.e., points can be moving. Each polygon has multiple instances each with a unique sequence number, i.e., the shapes and positions of polygons may change as well. A point may appear in several polygons (in the overlapping area of polygons). Sequence numbers of points and polygons belong to the same space and have no overlapping. Sequence number works as timestamp and a large sequence number means a recent time. When the sequence number of a point is given, all instances of polygons up to that time should be examined. Fig. 8 shows examples of polygons. Here, we can find polygons usually are irregular, and each polygon on average contains around 200 edges. Two polygons (12 and 13) further have inner rings, whose numbers of edges are equal to 15 and 20, respectively.

3.3.1 Efficient Geo-fencing

Geo-fencing is broadly applied in location-based services, e.g., advertisements, child location service. It can be well solved by using the crossing number algorithm [11] (or the winding number algorithm [12]). However, with the rapid increase of geo-spatial datasets, the geo-fencing technique is required to process millions of points and hundreds of polygons or even more in real-time. So, how to efficiently pair points with polygons is becoming a very important task. Here, we introduce a simple but effective and efficient geo-fencing algorithm [14], which is one of top winners in ACM SIGSPATIAL GIS Cup 2013.

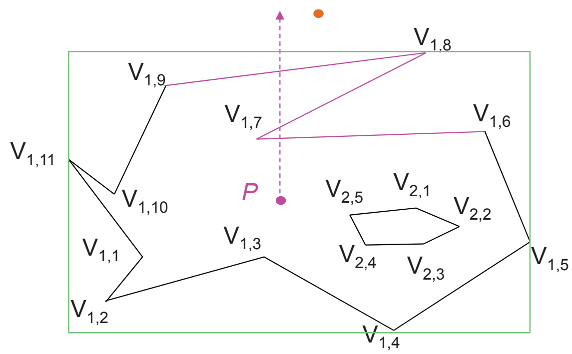


Fig. 9 Geo-fencing: detecting whether a point is inside a polygon.

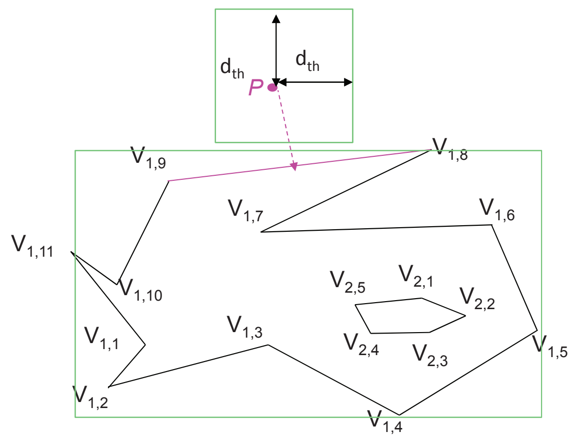


Fig. 10 Geo-fencing: detecting whether a point is within a distance d_{th} of a polygon.

According to the crossing number algorithm [12], the number of intersections for a ray passing from a point to the exterior of a polygon, if odd, indicates that the point lies inside the polygon, as shown in Fig. 9. But this crossing number algorithm requires checking all edges, and becomes inefficient when each polygon contains many edges. Actually, this problem can be simplified by two steps [14]: First, by exploiting the minimum bounding rectangle (MBR) of a polygon, a point outside the MBR of a polygon is surely outside the polygon. An R-tree is further used to quickly detect whether a point is inside the MBR of a polygon. Second, when the point is inside the MBR, instead of an exhaustive search, an edge-based locality sensitive hashing (LSH) scheme is proposed to adapt to the crossing number algorithm. As for the WITHIN detection in Fig. 10, a point might be outside the MBR of a polygon but still within a distance d_{th} of the outer-ring of the polygon. In this case, a rectangle centered at the point, with an edge length being $2d_{th}$ is constructed. If this rectangle does not overlap with the MBR of the polygon, the point is surely not within a distance d_{th} of the polygon. Applying LSH in the WITHIN detection is a little more

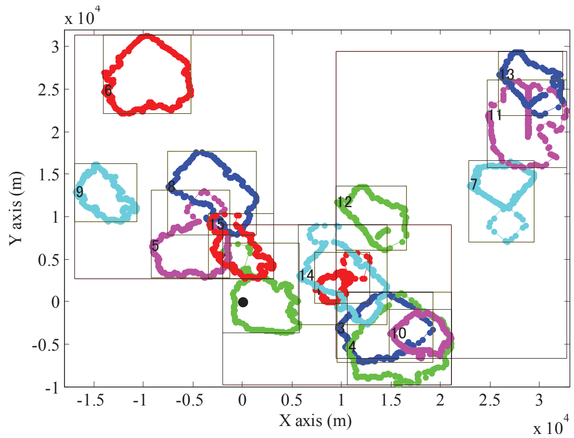


Fig. 11 MBR of polygons are organized in the R-tree.

complex. A probing scheme is suggested to locate adjacent buckets so as to check all edges near to the target point.

Fig. 11 shows an example of the relationship between an input point and its latest instances of polygons. In this figure, each polygon has its own MBR, and 15 basic MBRs are further divided into three groups in a higher level in an R-tree. In this way, MBRs that contain the given point are quickly found instead of exhaustive search. The corresponding polygons are regarded as candidates and are further examined.

3.4 Localized and Personalized Search

Personalization has been a trend of web searching. A method to personalized search is to exploit the location information. As is known, there is a geographic locality in user's interest and culture. So, for the same query, people in different areas may expect different results. These days, search engines can return most relevant local results to users according to the location information in user's profile, while filtering out irrelevant information. For example in Fig. 12, users searching for pasta restaurant in Kyoto and Singapore may get local relevant results. From these search results, it is obvious that Google search engine may personalize results based on users' location information. In this way, location information, as an important dimension, complements multimedia retrieval and recommendation.

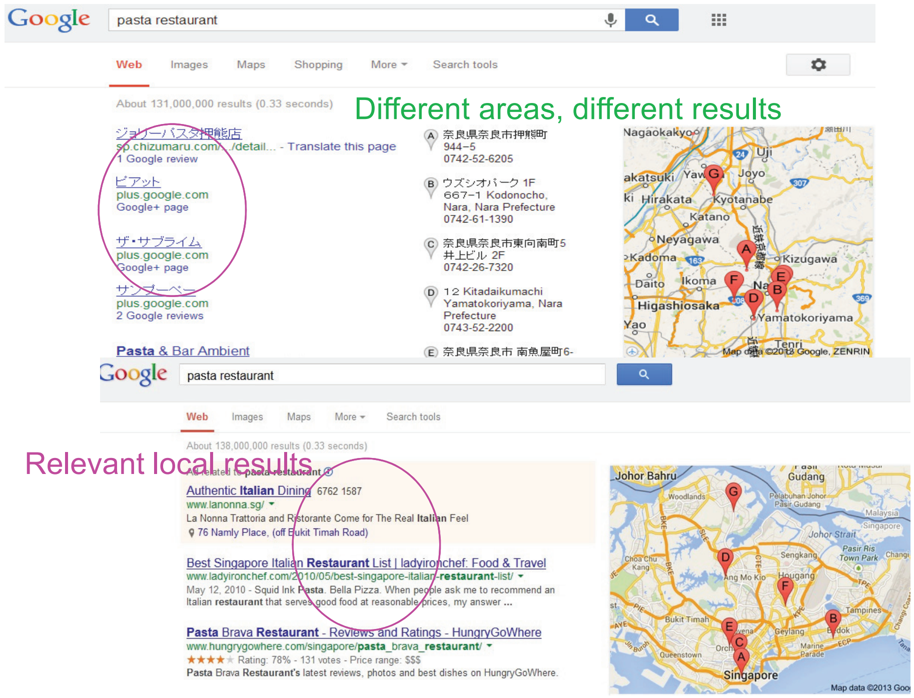


Fig. 12 Personalized search via exploiting location information.

4 Location-Aware Multimedia Applications

Here, we introduce emerging geo-tagged multimedia applications and techniques (e.g., land-use classification, geo-tagged image retrieval). Their common part is to incorporate geographic information as a context in multimedia information processing.

4.1 Music Geo-Listening Patterns

Twitter streaming API can be leveraged to retrieve tweets with geo-spatial coordinates. Further using music-related hashtags helps to extract music listening-related tweets. Then, artist information can be extracted by parsing and analyzing the content of these tweets. Music-related tweets often contain patterns, for example artist name followed by song title. In some cases, artist name might appear as a valid song title, which results in some ambiguity. Generally, pattern-based approach can be used to match potential artist names against the artist dictionary. Track information can be used to help distinguish artist names, by exploiting the musicbrainz database as a knowledge base for artist names and related song titles.

These music related tweets are classified according to artist genre information [20]. More specifically, the genre tags available for each artist are collected from last.fm, and further refined by using a list of known genres from freebase. Then, the artists (and genre tags) are split into k clusters, k ranging from 10 to 20. Next, each tweet is assigned an artist cluster number based on its included artist information. Further exploiting the coordinates of tweets, the number of tweets per artist cluster per area can be computed as music listening pattern and visualized on a world map.

4.2 Geo-Tagged Images for Land-Use Classification

Next example is exploring geo-tagged images for land-use classification. In this application, the problem of geographic discovery, particularly land-use classification, is investigated through crowdsourcing of geographic information from geo-tagged photo collections in Flickr. The geo-tagged photos are represented by their visual or text features to perform land-use classification. This is formulated as a supervised classification problem, in which support vector machine (SVM) [24] is used. Three land-use classes are considered in [19]: academic, residential, and sports. To generate a predicted land-use map, the target area is divided into multiple sub-regions, each separately classified.

Visual feature and text feature are main components in land-use classification model [19]. An intuitive question in the classification is how to model proximate sensing from visual features or textual feature contained in geo-tagged images. Bag of visual words (BoW) with a soft-weighing scheme is used to extract a BoW feature from each image, and a dictionary of 500 visual words is used. Flickr images commonly have user-supplied text associated with them. A dictionary of terms is created based on the words extracted from the title, descriptions, and tags associated with each image. The text analysis is performed at the group level since there is typically not enough text associated with the individual images for effective classification. Each of the text components associated with an image is parsed into a set of terms, and each group of images is represented by a histogram of terms among the dictionary. Then, pLSA (probabilistic latent semantic analysis) [25] is used as a tool to reduce the dimensionality of the term histogram of each image group.

4.3 Geo-Visual Image Similarity

Sometimes, it is necessary to identify geo-tagged images that contain similar views of identical objects so as to retrieve similar images taken in the same location. The geographic location of the photo image is measured where the picture is captured, not where the object is located. So, the position in the geo-tag is not the position of the captured object (but camera position where the object is taken). Images having identical objects are defined as orthologous images, for example, in Fig. 13, three

Fig. 13 Different photos showing the identical Merlion.

photos are similar to each other. Then, an orthologous identity function (OIF) [26] is used to estimate the degree to which two images are similar. OIF is a similarity rating function that uses both the geographic distance and image distance of photos.

4.4 Geo-Location and Context-Based Pedestrian Detection

In previous examples, we introduced classification of immovable objects such as land-use. Now we discuss the classification of movable objects, pedestrian detection related to geo-location.

Pedestrian detection can be conducted by different models with different complexities [27]. In the conventional model, the detection problem can be simply formulated as computing the posterior probability $p(P|V)$, where P denotes the pedestrian label and V denotes visual appearance of image or image batch. The second model adds the geographic location G as $p(P|V,G)$. Different locations will influence both the visual appearance and pedestrian presence probability. Therefore, the third model further exploits the environment context E, considering that different environments will influence the visual appearance of pedestrians. This model involves all factors including geographic location and environment context. Its context-based posterior probability $p(P|V,G,E)$ means the probability that an image contains a pedestrian given visual appearance V, the location G and environment E.

By leveraging a vast amount of web images, a contextual image database is constructed, in which each image is automatically attached with geographic location (i.e., latitude and longitude) and environment information (i.e., season, time and weather condition). Two pre-trained classifiers are exploited: a time classifier to decide whether an image was taken in the daytime or at night, a season classifier to decide which season an image was taken in. There is no any hint on weather condition in image metadata. Therefore, the weather condition is divided into three classes (snow, fog and normal), and a weather classifier is trained as well. By incorporating visual feature, geographic location (i.e., latitude and longitude) and environment context (i.e., season, time and weather condition), a context-based pedestrian detection method [27] can be realized by the probabilistic model discussed above.

4.5 Soundtrack Recommendation for User-Generated Videos via Context Information

Most user-generated videos, taken outdoor, lack suitable soundtracks. Adding a matching soundtrack to a video can make the video much more attractive for sharing. Generally, different geo-locations convey different affective atmospheres. For example, a busy city has a different atmosphere from a majestic mountain view. In this sense, each geo-category is associated with a mood. Based on mood similarity, a soundtrack can be recommended to a video scene.

Table 1 Relationship between geo-categories and moods.

Geographic category	Related mood(s)
Arts & Entertainment	Quiet, calm
Colleagues & Universities	Quiet, calm
Food	Sweet, happy
Great Outdoors	Dreamy
Nightlife Spots	Funny, intense, playful
Professional & Others Places	Aggressive, heavy
Residences	Sweet, sleepy
Shops & Services	Happy
Travel & Transport	Melancholy, bittersweet, funny

Geo-locations can be classified into geo categories through leveraging Foursquare API. A geo-category is further associated with an atmosphere, or a mood at a venue. Table 1 shows a potential mapping from geo-categories to moods [7]. User study is conducted to identify which mood should be associated with each geo contextual category. By using the relationship in this table, a system can automatically rank mood categories for a given geo-location.

The whole soundtrack generation system [7] has two parts: smartphone application and server side. User generated videos are captured by smartphones together with continuous streams of geo-sensor (GPS) information. These geo-sensor data streams are mapped to a set of ranked, textual geo-tags. Geo-tags are further classified to geo-categories via the API provided by Foursquare, and then mapped to mood tags according to a predetermined geo-mood mapping table (refer to Table 1). Mood-tags provide the input into a music retrieval engine, which returns a music soundtrack most matching these tags. Finally, the music soundtrack is associated with the video and the new video is ready for sharing.

The performance of soundtrack recommendation for user-generated videos can be further improved by exploiting visual features as well. Especially, the classification results from geo-feature and visual feature via SVM [24] are late-fused to generate a more robust result, as discussed in [8, 9].

5 Analysis of Geo-Social Data

Social media are user-centric, and designed for the interactions and communications between users all over the world. Social networking platforms provide ways to create and exchange user-generated contents while sustaining human contact at the same time. Social media have different forms and languages, which include, e.g., videos, images, audio songs, comments, reviews, ratings. Users participate in social networking platforms via different devices, e.g., using desktop PC, tablet, smart phone, or game console. Conventionally, user interface approaches address the user's interaction with devices, the interactions between a user and a software or application. In comparison, online social interactions [2] are the communications between users via the help of social interface. They are established through the self-reinforcing activities of participating users. Social interactions reveal what is going on, what an application or site is about, and reflect psychological views of identity, the self, interpersonal relationships, and social structures.

In the following, we show with examples some approaches related to the analysis of geo-social data, especially personal preference mining, social knowledge discovery, and geographic distribution of social activities.

5.1 Analysis of Social Expertise Based on Number of Check-ins

A user visits different venues and generates different check-ins online. These check-ins reflect user's location history in the physical world. Foursquare has a hierarchical category structure which includes 9 top categories and 410 sub-categories. Using the API provided by Foursquare, the geographic trajectory of a user can be converted to a series of geo-categories, which contains user's personal preferences. A user's location history is regarded as a document and categories or sub-categories are considered as terms in the document. By exploiting the TF-IDF (term frequency inverse document frequency) method [28], features can be computed at different levels, using either categories or sub-categories as vocabularies. When computing the similarity between two users in terms of the trajectories, a similarity score can be computed at each level of the hierarchical category, and their weighted sum gives a total similarity between two users [10].

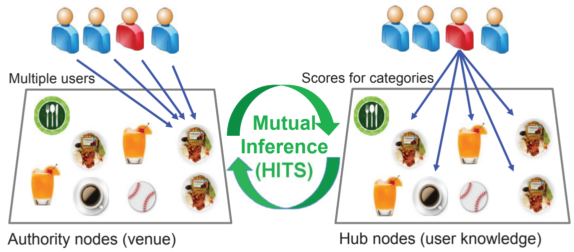


Fig. 14 Iterative model for social expertise discovery.

An example is used to explain how to make use of an iterative model for social expertise discovery [10]. In the model shown in Fig. 14, each user has his scores for different venue categories, and a venue category is associated with multiple users. For a specific category m, a user's knowledge, $u_m.h$, can be represented by the sum of the authority scores ($v_m.a$) of the venues visited by the user (Eq.(3)). On the other hand, the authority score of a venue, $v_m.a$, can be represented by the hub scores ($u_m.h$) of the users who have visited this venue (Eq.(4)).

$$u_m.h = \sum_{u.v \in m} v_m.a. \tag{3}$$

$$v_m.a = \sum_{u \in U} u_m.h. \tag{4}$$

Then, a user with a high score in a category is regarded as a local expert of that category. To identify the local experts of a venue category, for example, Italian food, based on category information recorded in the user's location history, a user's expertise in each category in different cities can be computed by an iterative model, known as mutual inference [29]. In this process, the initial authority and hub scores are set as the number of user's visits.

5.2 Analysis of Business Venues Based on Check-ins

As mentioned before, Foursquare has 9 top-categories and 410 sub-categories. However, in some cases, for example, in trade area analysis, shopping habits are desired [15]. Top 9 categories cannot effectively distinguish check-in patterns. 410 sub-categories can be more valuable in user profiling but with too high dimension.

The LDA (latent dirichlet allocation) method [30] can be used to identify hidden check-in patterns as topics from the histogram of user check-ins in terms of sub-categories. LDA is widely adopted in document topic modeling. It assumes that each document contains a mixture of topics and each topic has certain probability of mentioning a word. LDA identifies topics and calculates the proportion of different topics in each document by examining word distributions in the documents.

Fig. 15 Business attractiveness, the size of an icon is proportional to the popularity of the corresponding business.

More specifically, the distribution of different topics is calculated for a document. Each user is treated as a document, and each topic is regarded as a term. By analyzing the distribution of topics of customers of a store and computing the histogram of the main topic of all customers, the stores can be profiled as well in terms of potential topics [15].

Another example is shown in Fig. 15 for business attractiveness discovery, where the size of an icon is proportional to the popularity of the corresponding business. Consider a number of customers C_1 to C_n and a number of competitor venues V_1 to V_m. a_{ij} represents the number of visits of customer C_i to venue V_j. Then, the probability that a venue V_j is visited can be calculated via

$$P(V_j) = \frac{\sum_{i=1}^{n} a_{ij}}{\sum_{i=1}^{n}\sum_{k=1}^{m} a_{ik}}. \tag{5}$$

This probability is an indicator of local popularity of a business venue. In other words, the probability of venue V_j being visited by all customers in an area reflects its business attractiveness compared with other competitor venues.

5.3 Analysis of User Check-ins in Yelp

In this section, we investigate user behaviors in physical world, by using experimental check-ins data. This data includes 11,537 businesses and 8,282 sets of check-ins from March 2005 to January 2013 in Phoenix, which was provided by Yelp.

Fig. 16 and Fig. 17 respectively show user activity patterns across the 10 most popular categories of Yelp on weekdays and on weekend. Top 7 categories are the same in the two figures, which indicates that user business activities on weekdays and weekend have no significant differences in the area of Phoenix. However, we

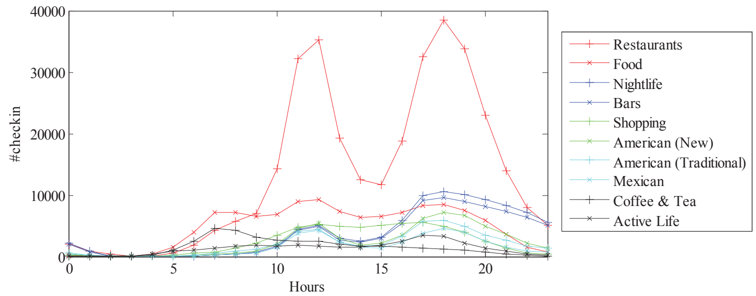


Fig. 16 User activities on weekdays in Yelp.

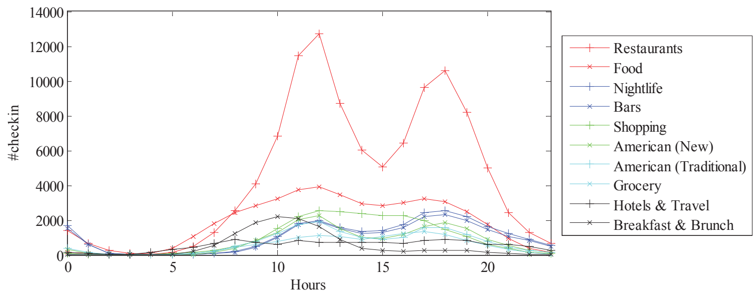


Fig. 17 User activities on weekend in Yelp.

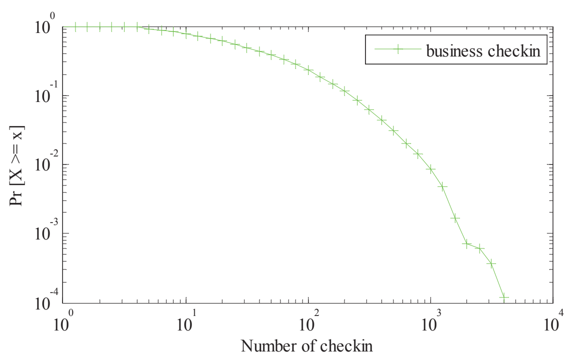


Fig. 18 CCDF of the number of check-ins at business venues in Yelp.

still can find that more people like to travel and go to grocery on weekend than on weekday.

Fig. 18 shows the CCDF (complementary cumulative distribution function) of the number of per-user check-ins. The number of check-ins varies greatly among users. 50% users have a check-in count no more than 20. On the other hand, 1% users have more than 1000 check-ins.

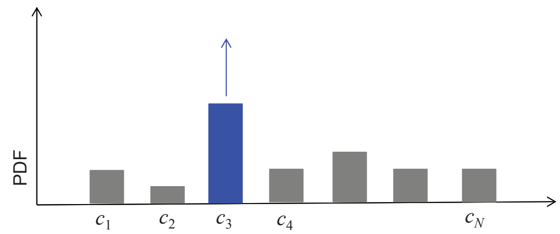


Fig. 19 Distribution of the number of visits to different categories for a user.

5.4 Analysis of Interest Focus and Entropy in Foursquare

User-generated geo-social data contains user behaviors in physical world and also reflects geographic reach and interest of a geo-category context or a multimedia content across the globe. Here, some methods and examples are given to address geographic distribution of social activities related to geographic popularity of photos, tips and videos.

We investigate the distribution of social media data of LA and NYC, crawled from Foursquare, where 2,728,411 venue photos and 1,212,136 tips are used in the experiments. Since Foursquare is a location-based social networking platform, large volumes of tips and photos are posted in this community. Each user has different interest over all the geo-categories, reflected in the variations of the number of per-category tips or photos. In other words, the distribution of a user's visit in terms of geo-category would likely exhibit a non-uniform distribution, with a large fraction of visits in only a few categories. The distribution of user interest can be measured by two metrics, interest focus and interest entropy.

Fig. 19 shows the distribution of the number of visits of a user in each category, which is usually non-uniform. Let v_{ik} represent the number of visits of a single user i to a category k. Then, interest focus of a user is defined as its highest fraction of visits, as follows:

$$F_i = \max_j \frac{v_{ij}}{\sum_k v_{ik}}. \tag{6}$$

A higher interest focus means the interest of a user is more limited to a specific category.

CCDF of per-user interest focus is shown in Fig. 20, where the visit is defined in terms of tips (messages) or photos. In this figure, we used the top-9 categories. Nearly 50% users have an interest focus greater than 0.5, which indicates that many users have a primary interest (in terms of geo-category).

Interest entropy is the other metric for evaluating how user interests are distributed over different categories. With the fraction of visits to a category in Fig. 19 as a probability, interest entropy is computed as a standard entropy, as follows:

$$H_i = -\sum_k p_{ik} \log_2 p_{ik}, \tag{7}$$

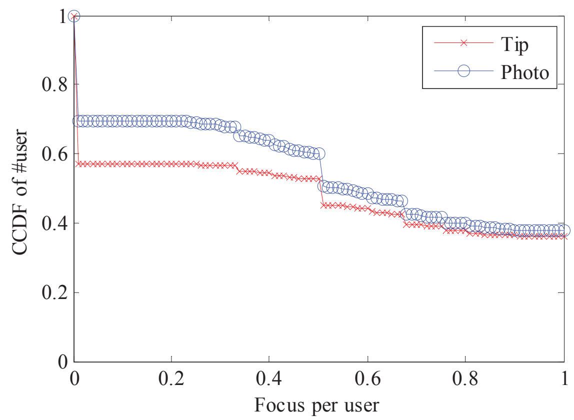


Fig. 20 CCDF of interest focus in terms of photos and tips in Foursquare.

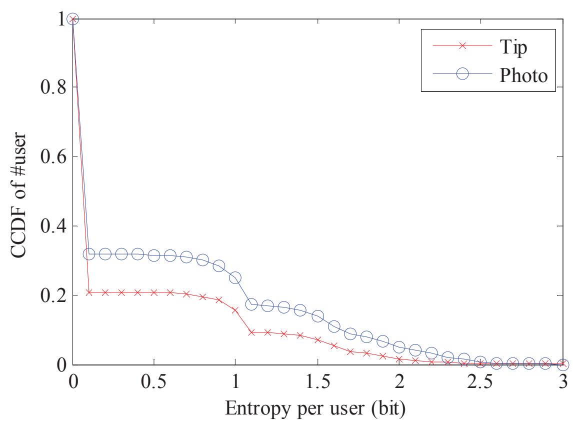


Fig. 21 CCDF of interest entropy in terms of photos and tips in Foursquare.

$$p_{ik} = \frac{v_{ik}}{\sum_j v_{ij}}.$$

It reflects how user interests are distributed over different categories. A higher interest entropy means a more uniform distribution of visits to different categories while a lower value means interests are focused in fewer categories.

Fig. 21 shows the CCDF of per-user interest entropies, in terms of tips and photos. Only 20% users has an interest entropy of photo greater than 1bit, or the number of categories being frequently visited is equal to 2. The interest entropy of tips is lower.

6 Integration of Social and Content Networks

These days, users find videos from the Internet by different methods. Some of the videos are directly searched via the web sites, some other videos may be recommended between users through their social connections. As a result, social connection has a significant impact on video views. In addition, the effect of social sharing is becoming more important as more users are involved in the social networks.

6.1 Geo-Social Networks

Every day, a huge volume of Internet traffic is generated by online multimedia sharing platforms such as YouTube, Flickr, Last.fm. These platforms often rely on content delivery networks [1] to distribute their content from storage servers to multiple locations over the planet. Servers exchange content in a cooperative way to maximize the overall efficiency. Recently, content diffusion is also fostered by web-links shared on online social networks [2]. This may generate large amounts of requests to the provider through the cascading across a user's social links. Content discovery heavily depends on the web search. Web search services like Bing and Google Web Search now are an integral part of our daily life. Google Search alone receives 12.8 billion queries[7] every month from U.S. users. People use web search for a couple of reasons, including listening to music, watching baseball, and making purchase decisions.

Lots of applications and services on the Internet have been developed to make use of location information to meet users' daily needs. The increase of various social media services requires a global platform for sharing user-generated contents, such as videos, images, music, blogs and tweets. Location-enabled tagging for social contents via smart phones and social media services reflects geo-spatial logs of user activities. Users can link their presences and multimedia contents (for example, video, image) to a particular place. Geo-spatial footprints generated by users provide interesting information about the spatio-temporal dynamics of online memes [31], which have important implications for a variety of multimedia systems and applications [8, 20, 32]. On the one hand, geo-spatial data contains personal physical logs of each individual user. On the other hand, it also reflects social behaviors related to the community as data of more users is aggregated. These geo-spatial data could be very useful for studying various lifestyle patterns [33], e.g., public health, cultural identification, urban computing.

[7] http://bit.ly/ThGnOc

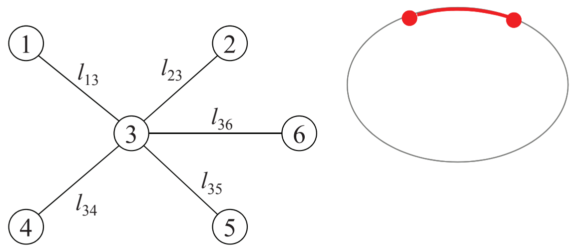


Fig. 22 Friendship and great circle distance.

6.1.1 Graph Representation of Geo-Social Networks

Next we introduce how to model social networks via a graph [17]. Online users are located over the 2-dimensional surface of the Earth. The great-circle distance is adopted as the metric. The distance in Fig. 22 between any two nodes is calculated as a great-circle distance from their geographic coordinates. The social tie between two nodes is represented by a link between them.

A social network can be represented as an undirected graph G, with a node set N and a link set K. When there is a social tie between two users i and j, a link is established between them. The link length is associated with the great circle distance l_{ij}. It is useful to find how friends of a user are geographically distributed. One useful metric is node locality (Eq.(8)). Considering node i and all its neighbors in a set Γ_i, the geographic closeness between two nodes is measured by a function of normalized distance using a parameter β. The average over all nodes in the neighbor set Γ_i gives node locality [17] of node i.

$$L_i = \frac{1}{|\Gamma_i|} \sum_{j \in \Gamma_i} e^{-l_{ij}/\beta}. \tag{8}$$

In this way, the node locality represents the average closeness between a user and his friends, and decreases as the actual distance gets larger. It is useful when exploiting social connections to recommend multimedia contents, as is discussed later.

The node locality can be investigated by using the cascade of Twitter messages. Twitter messages are shown on the author's personal page and also sent to the author's followers. A node is used to represent a user with a geographic location. Then, a directed graph of users can be extracted from the dataset of tweets, and node locality of each user can be calculated.

A more interesting phenomenon is the spreading of YouTube video links via tweets. A cascade over a social network begins when the first user shares some content and becomes the initiator of the cascade. After this event, some of his contacts will share the same content again, and the cascade will recursively spread over the social links.

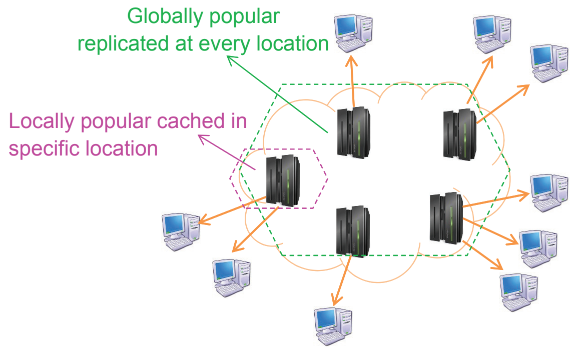


Fig. 23 Content delivery networks handling locally popular and globally popular contents differently.

6.2 Geo-Social Multimedia Content Delivery

Now we introduce some methods related to multimedia content diffusion, for example, geo-social cascades, caching policies and distributed cache. The popularity of multimedia content over the Web can be driven by public media coverage. This type of phenomena often results in globally popular items, which should be widely replicated throughout a content delivery network. Alternatively, content may become popular over social networking platforms because people share it and talk about it. In this way, content may easily spread from a small set of users to a vast audience through social connections, for example, 700 YouTube video links are shared on Twitter every minute [8].

Social sharing also has a large impact on content delivery network. The latter is a system of networked servers holding copies of data items, placed at different geographic locations as shown in Fig. 23. Its performance is influenced by the geographical distributions of the requests. Then, it would be very useful to understand whether an item becomes popular on a planetary scale or just in a particular geographic area. A globally popular content item should be replicated at every location, since it receives many requests from all around the world. On the other hand, when content is only locally popular, it should be cached only in specific locations.

As for standard caching policies [34] used in content delivery networks, each policy assigns a priority $P(v)$ to a video v, and the video with the lowest priority is chosen for deletion when the cache buffer gets full. There are three typical caching policies. In Least-Recently-Used (LRU) policy, $P(v)$ equals $clock(v)$. $clock(v)$ is the last time that the video v is watched, and it involves the simple aging effect. In Least-Frequently-Used (LFU) policy, $P(v)$ equals $Freq(v)$, where $Freq(v)$ is the number of times video v has been requested since it was stored in the cache. The mixed policy combines LRU and LFU, and the priority of video v is given by $P(v) = clock(v) + Freq(v)$, in order to balance both temporal and popularity effects.

[8] http://www.streamsend.com/

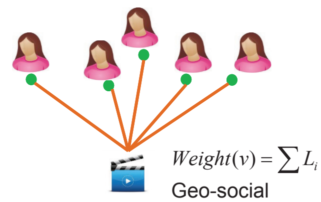


Fig. 24 Geosocial in geographic social networks: users access the same video.

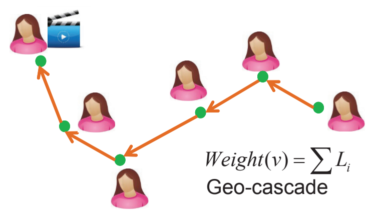


Fig. 25 Geocascade in geographic social networks: users are in the same social cascade.

The above caching policies can be further augmented by exploiting geo-social information. Twitter fosters the popularity of YouTube, since users tend to tweet about videos they like, triggering a spreading of the video. This provides us some opportunities to investigate how geographic information is extracted and used to improve caching of multimedia files. There are two augment catching policies [18] based on the characteristics of the geo-social cascades involving one video. One is Geosocial (shown in Fig. 24), the extra weight of video v, which is added to its priority, is the sum of the node locality values of all the users that have posted a message about the video, even though they are not involved in a social cascade. The other is Geocascade (shown in Fig. 25), the extra weight of video v that is added to the priority is the sum of the node locality values of all the users participating in the video's social cascade. In this way, exploiting social connections helps to find whether a video becomes popular and helps to optimize the cache management.

7 Summary of This Chapter

With an overwhelming amount of social multimedia content on the Internet, it is difficult to find what users are really interested in. For example, a search for "pasta" may return hundreds of millions of social media items. In addition, sometimes, an incomplete query may lead to results of different meanings, where more accurate search requires further information on user preference. For example, a search for "apple" returns the fruit apple and the apple brand. Personalization has been a trend of web searching. Recently, many social networking platforms has provided location-based services, by either explicitly letting users choose their places or implicitly enabling geo-tagging to associate multimedia content with latitude and longitude. Location information has been a very important aspect that helps better understand online social media contents and people activities in physical world.

This chapter takes user location as a clue to discuss a broad topics over location-aware multimedia systems. We have talked about fundamental components related to geographic-aware social media, mobile users, social activities and multimedia content delivery. More specifically, geo-social data contains rich context and has two aspects of implication: individual user interest and geographic-social behaviors. We have shown some examples of geographic-aware social media and social interaction data, and reported latest geographic-aware multimedia applications and methods, for example, how to leverage tweets with geospatial information for mining music listening patterns, how to map geo-categories to moods. We also have discussed some location-enabled advanced topics and approaches. Particularly, we explained geo-fencing in detail, which is a promising technique for user-centric mobile location-based services. We showed some approaches related to personal preference mining, social knowledge learning, geographic distribution of social activities and multimedia content diffusion. To sum up, exploiting location information to mine user preference and social links to predict content popularity will greatly affect the form of content retrieval and delivery, which are attracting, and will continue to attract much research interest.

Acknowledgements The work presented was in part supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

References

1. Pallis, G. and Vakali, A. (2006). Insight and Perspectives for Content Delivery Networks. Communications of the ACM, 49(1), 101–106.
2. boyd, d. m. and Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.
3. Junglas, Iris A. and Watson, Richard T. (2008). Location-based Services. Communications of the ACM, 51(3), 65–69.
4. Brodersen, A., Scellato, S., and Wattenhofer, M. (2012). YouTube around the World: Geographic Popularity of Videos. In WWW, 241–250.
5. Shen, Z., Arslan Ay, S., Kim, S. H., and Zimmermann, R. (2011). Automatic Tag Generation and Ranking for Sensor-rich Outdoor Videos. In ACM Multimedia, 93–102.
6. Ma, H., Zimmermann, R., and Kim., S.H. (2012). HUGVid: Handling, Indexing and Querying of Uncertain Geo-tagged Videos. In ACM SIGSPATIAL, 319–328.
7. Yu, Y., Shen, Z., and Zimmermann, R. (2012). Automatic Music Soundtrack Generation for Outdoor Videos from Contextual Sensor Information. In ACM Multimedia, 1377–1378.
8. Shah, R.R., Yu, Y., and Zimmermann, R. (2014). User Preference-Aware Music Video Generaton Based on Modeling Scene Moods. In ACM MMSys, 156–159.
9. Shah, R.R., Yu, Y., and Zimmermann, R. (2014). ADVISOR–Personalized Video Soundtrack Recommendation by Late Fusion with Heuristic Rankings. Accepted by ACM Multimedia'14.

10. Bao, J., Zheng, Y., and Mokbel, M. F. (2012). Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data. In ACM SIGSPATIAL, 199–208.
11. Shimrat, M. (1962). Algorithm 112: Position of point relative to polygon. Communications of the ACM, 5(8), 434–434.
12. Hormann, K., and Agathos, A. (2001). The point in polygon problem for arbitrary polygons. Computational Geometry, 20(3), 131–144.
13. Kupper, A., Bareth, U., and Freese, B. (2011). Geofencing and background tracking - the next features in LBS. In INFORMATIK11.
14. Yu, Y., Tang, S., and Zimmermann, R. (2013). Edge-Based Locality Sensitive Hashing for Efficient Geo-Fencing Application. In ACM SIGSPATIAL, 576–579.
15. Qu, Y., and Zhang, J. (2013). Trade Area Analysis using User Generated Mobile Location Data. In WWW, 1053–1064.
16. Quercia, D., Di Lorenzo, G., Calabrese, F., and Ratti, C. (2011). Mobile Phones and Outdoor Advertising: Measurable Advertising. IEEE Pervasive Computing, 10(2), 28–36.
17. Scellato, S., Noulas, A., Lambiotte, R., and Mascolo, C. (2011). Socio-spatial Properties of Online Location-based Social Networks. In ICWSM.
18. Scellato, S., Mascolo, C., Musolesi, M., and Crowcroft, J. (2011). Track Globally, Deliver Locally: Improving Content Delivery Networks by Tracking Geographic Social Cascades. In WWW, 457–466.
19. Leung, D. and Newsam, S. (2012). Exploring Geotagged Images for Land-Use Classification. In GeoMM, 3–8.
20. Hauger, D. and Schedl, M. (2012). Exploring Geospatial Music Listening Patterns in Microblog Data. In 10th International Workshop on Adaptive Multimedia Retrieval.
21. Ikawa, Y., Vukovic, M., Rogstadius, J., and Murakami, A. (2013). Location-Based Insights from the Social Web. In WWW'13 Companion, 1013–1016.
22. Eisenstein, J., Ahmed, A. and Xing, E. (2011). Sparse additive generative models of text. In ICML, 1041-1048.
23. Hong, L., Ahmed, A., Gurumurthy, S., Smola, A. and Tsioutsiouliklis, K. (2012). Discovering Geographical Topics in the Twitter Stream. In WWW, 769–778.
24. Joachims, T., Finley, T., and Yu, C.-N. (2009). Cutting-plane training of structural SVMs. Machine Learning, 77(1), 27-59.
25. Hofmann, T. (1999). Probabilistic Latent Semantic Indexing. In ACM SIGIR, 50–57.
26. Kamahara, J., Nagamatsu, T., and Tanaka, N. (2012). Conjunctive Ranking Function using Geographic Distance and Image Distance for Geotagged Image Retrieval. In GeoMM, 9–14.
27. Liu, Y., Shi, Z., Wang, G., and Guan, H. (2012). Find You Wherever You Are: Geographic Location and Environment Context-Based Pedestrian Detection. In GeoMM, 27–32.
28. Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. Journal of Documentation, 60(5), 503-520.
29. Zheng, Y., Zhang, L., Xie, X., and Ma, W.Y. (2009). Mining interesting locations and travel sequences from gps trajectories. In WWW, 791-800.
30. Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, **3**, 993–1022.
31. Kamath, K. Y., Caverlee, J., Lee, K., and Cheng, Z. (2013). Spatio-Temporal Dynamics of Online Memes: A Study of Geo-Tagged Tweets. In WWW, 667–678.
32. Yamasaki, T., Gallagher, A., and Chen, T. (2013). Personalized Intra- and Inter-City Travel Recommendation Using Large-Scale Geotags. In GeoMM'13, 25–30.
33. Yu, Y., Aizawa, K., Yamasaki, T., and Zimmermann, R. (2014). Emerging Topics on Personalized and Localized Multimedia Information Systems. Accepted by ACM MM.
34. Wang, J. (1999). A Survey of Web Caching Schemes for the Internet. ACM SIGCOMM Computer Communication Review, 29(5), 36–46.