A Conceptual Framework for Representation of Location-based Social Media Activities

Xuebin Wei
Department of Integrated Science and Technology, James Madison University, 701 Carrier Dr, Harrisonburg, VA 22807, USA
weixx@jmu.edu
https://orcid.org/0000-0003-2197-5184

Xiaobai Angela Yao
Department of Geography, University of Georgia, 210 Field St., Athens, GA 30602, USA
xyao@uga.edu
https://orcid.org/0000-0003-2719-2017

Abstract

This research develops a conceptual framework for the representation and analysis of location-based social media activities (LBSMA) in GIS. With increasing popularity of location-based social networking, social media platforms have become new channels to observe human activities in physical and virtual worlds. At the same time, there is a shift of some human interactions from the physical space to the virtual social space. Traditional geographical representation in GIS is not sufficient to handle the increased sophistication of human activities related to, or embedded in, location-based social media data. This research proposes an ontology for the location-based social media activity data and a conceptual framework for them to be modeled in a GIS environment so that interconnections of human activities in spatial-temporal-social dimensions can be represented, organized, retrieved, analyzed, and visualized in the system.

2012 ACM Subject Classification Information systems → Data management systems

Keywords and phrases GIS, Social Media, Ontology, Location-based Social Media Activity

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.62

Category Short Paper

Supplement Material Study Website: `www.lbsocial.net`

1 Introduction

Understanding human dynamics through human activities has been an important geographic inquiry in the literature. Researchers have studied human behavior from various perspectives. Behavior geography concerns the cognitive process of human behavior and draws on works in other fields such as psychology, physiology and economics. Another thread of research examines human activities through visualization, analysis, and modeling of human dynamics. Our research attempts to contribute to the latter, motivated by two fundamental issues. First, the growing popularity of computer network-based social media and the availability of data from these social media provide an unprecedented opportunity to study human activities in new lights. However, the new types of data require new conceptualization, new methodologies, and new tools to make the best out of them. Secondly, it has been well recognized that social connections play an important role in human behavior. However, social network has been ignored or oversimplified in current representations of human activities in current off-the-shelf GIS programs. Therefore, this research aims to develop a GIS conceptual

framework and associated logical models to represent space, time, and social connections from location-based social media activity (LBSMA) data in GIS.

By nature of the topic, studying human activities ideally requires data at the individual level, or so called disaggregated level, with fine spatial and temporal granularities. However, commonly available census data are usually aggregated. Thus the availability of location-based social media (LBSM) data provides an unprecedented opportunities for this type of research, as the LBSM data are inherently entered on individual basis and are of high granularities in space and time. In the information age, a message from social media is considered an extension of human mind [10]. Furthermore, details of human activities can now be extracted from the social media to reveal when and where people interact with others. Collections of such interactions can be used to develop social networks among people. This is particularly advantageous for research on human activities, as the context of social connection is particularly important to human activities. It has been argued that time, space and social differentiation should be coupled in the study of practices or phenomena [7]. From the relationalism-idealism perspective, the assumed existence of social networks sets the scope to which space and time should be conceptualized and analyzed in human activity analysis [12]. Different types of social media allow for different types of connections. For example, Twitter fosters an asymmetric network structure that people prefer to broadcast individual activities, while LinkedIn and Facebook capture pre-existing ties by focusing on social interactions among friends [9]. Previous studies have investigated the content and friendship structure on Twitter [9][5], Facebook [3] and Weibo [4]. The spatial distribution of location-based social activities from different social media has also been explored in recent studies.

There is a long tradition that human activities are visually represented and analyzed, particularly in GIS. Starting from the space-time prism [1], trajectories of human activities are visually represented as a series of locations in space-time dimensions. Because human activities have innate spatial component, geographic information system (GIS) is naturally the most desirable environment for the visualization and analysis of it. Sui and Goodchild [8] suggest that GIS is a media for communicating and sharing knowledge and supporting location-based social networking. Meanwhile, the convergence of geographic information systems (GIS) and social media has resulted in a data avalanche that creates new challenges in GIScience [8]. Although location-based social media activities have been examined in many studies, a structured GIS representation is still absent for all three dimensions of space, time, and the context of social connections in which activities take place. GIS representation of space and time alone is already a critical research theme in the literature [11], adding more dimensions obviously is not a trivial issue. The goal of this paper is to fill the gap by developing an ontological framework and a conceptual model for the representation location-based social media activity (LBSMA) data in GIS, so that the space, time, and social connections associated with LBSM activities can be represented and analyzed further. In this research, as shown in Figure 1, the LBSMA refer to the subset of human activities of which locations can be georeferenced in the geographical space and contents are advertised in the networked social media. The scope of the study is limited to human activities that are recorded explicitly or implicitly in the LBSMA data, which is illustrated as the yellow-highlighted area in Figure 1.

The paper is organized as follows. The next section presents an ontological framework for LBSMA and a conceptual model for the representation of LBSMA in GIS. Section 3 introduces a pilot implementation based on the framework. A case study is conducted in the prototype. The paper is concluded with discussions in Section 4.

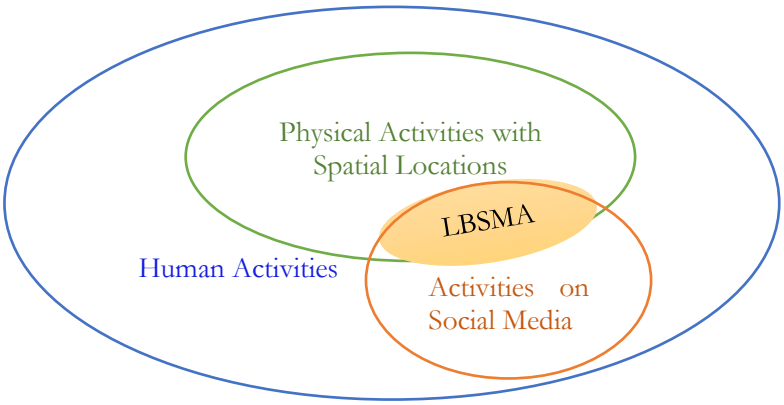

Figure 1 Location-based Social Media Activity.

2 LBSMA Ontological Framework and Representation in a GIS environment

Traditional GIS conceptual models use either object-based or field-based representations. The former distinguishes each spatial object with delineated spatial boundaries, while the latter enumerates all spatial locations systematically and stores attribute values for each location. However, none of them is able to directly account for social network (or social associations) or human activities in the context of such a network. Aiming to have a conceptual underpinning for later technical deployment to fill the gap, this paper first develops an ontological framework that identifies four primary categories for the LBSMA. They are Agents, Activities, Places, and Social Connections. Following the ontological framework, a conceptual data model is designed in the paradigm of object-oriented modeling. The model is illustrated in Figure 2. The purpose of conceptual model is to organize the data in a reasonable and retrievable way, so as to maximize the possibility to study hidden relationships and patterns in the spatio-temporal big data of growing size. The most important aspect is to allow information in the spatio-temporal-social dimensions, expressed either explicitly or implicitly, to be identified and represented in the system.

2.1 Agents

Here an agent refers to a person or a collective entity with a group of individuals as long as the entity has a unique ID in a social media platform. A person may have one or more active accounts in social media platforms and thus may be associated with multiple agents. An agent can participate in any number of activities, some of which show social connections with other agents. It is also possible that through analytical methods, multiple agents are found to be the same person in the real world. A collective entity, such as a restaurant, a university, or an association, which has an official account in a social media platform can be identified as an agent too. It is often more likely for such an agent to identify itself with its official name and thus can be readily associated with its real world identity.

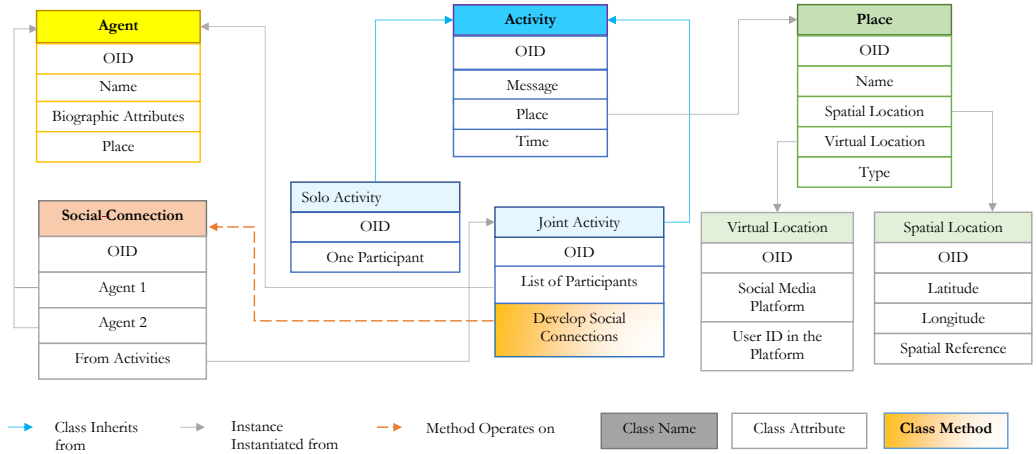

Figure 2 Conceptual Model of LBSMA.

2.2 Places

An agent may exist in a physical geographic space, a virtual social media space, or in both spaces. Uniquely identifiable locations in either type of frame of references are called places in this ontology. Therefore two types of locations are distinguished in the framework: the geographic location (or spatial location) and the virtual location. For example, just like an address can refer to the geographic location of a person's home, a uniform resource locator (URL) of a user's profile page refers to the person's virtual location on social media. Sometimes, information of both types of locations may be available for an agent. For instance, a restaurant can have its footprints in the geographical world, while its URL is a location in the virtual world where its menu, public reviews or other types of information can be retrieved. Virtual locations are equally important in the framework because they not only facilitate the organization of human activities in the virtual world, but also provide the source of rich information about people, activities and the context of environment.

2.3 Activities

An activity in this ontology refers to any action of an agent in either the physical space or the virtual space. For example, visiting, commuting, reading, participating in a party are examples of activities in the physical world, while posting a message and following another account on Facebook or Twitter are examples of activities in the virtual space. Activities can be further classified into solo activities and joint activities based on the number of participants. Because the scope of this study is the intersection area of human activities in the physical space and the social media space, the activities of concern are those reflected in a social media regardless which space the actual activity took place.

2.4 Social Connections

In this framework, social connections are personal relationships expressed via social interactions. It is a mind-dependent construct that can be reflected by mind-independent human activities. Social connections can be explicitly expressed or identified through their self-reported relationships such as kinship, workplace connection, friendship, and so on, which can be explicitly indicated in the profiles or implicitly revealed via connections between

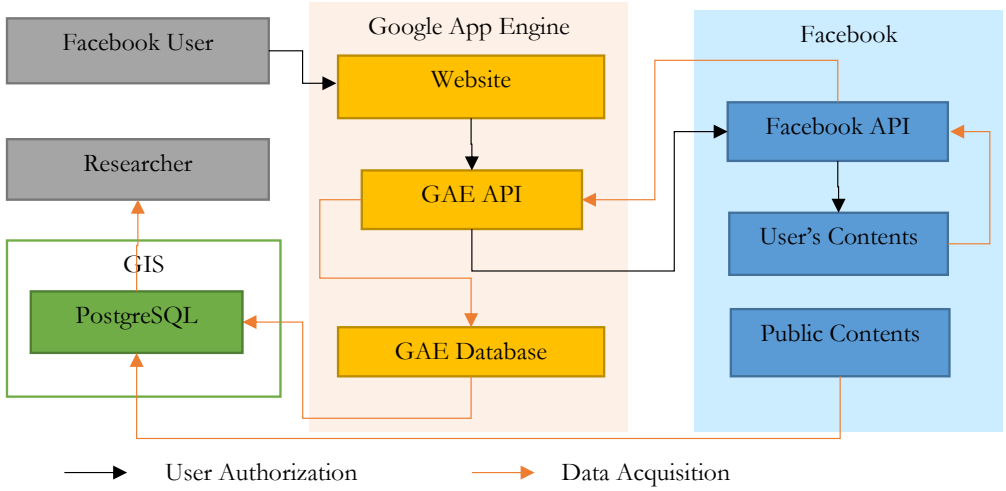

Figure 3 LBSMA Data Collection.

profiles on social media platforms. However, many additional social connections can be identified through spatial-temporal reasoning. For instance, two agents may already have, or are potentially developing, a social connection if they participate in joint-activities. Frequent joint-activities at the same home address suggests close family ties or friendship.

3 Pilot Prototype Implementation and Case Study

Based on the proposed conceptual model, a pilot prototype has been implemented and tested. A case study is performed to validate the prototype and most importantly to evaluate the usefulness of the proposed framework of LBSMA representation in GIS. This research has widely recruited students in the University of Georgia to collect their Facebook data. The extracted Facebook data are organized in the implemented LBSMA data model. About 500 unique Facebook accounts and 2,500 posts have been collected in this case study.

This website obtained the Institutional Review Board (IRB) and Facebook App approval, and has gathered participants' information through the Facebook Application Programming Interface (API) with explicit authorizations of Facebook users. The website is running on the Google Application Engine (GAE). The collected LBSMA data is then organized and maintained in a PostgreSQL database with the PostGIS plugin to provide GIS functions. Figure 3 shows the process. When a post is received, the name and user account ID are recorded in an Activity table. Other people who are tagged in the same post are also kept in the Participants field. Since the number of the tagged people is not predictable, this data filed utilize the JSON format to record all the participants.

The prototype has developed a set of visualization and analysis tools for the LBSMA in ArcGIS, including visualize activities and places, query people-based social network, create location-based social network and identify spatial-temporal interactions of activities. The demo of the developed tools can be found at `https://www.youtube.com/watch?v=aJnm0GTqV5w&list=PLHutrxqbP1BxvYmC0GX5fDQkLYUrbqsS5`

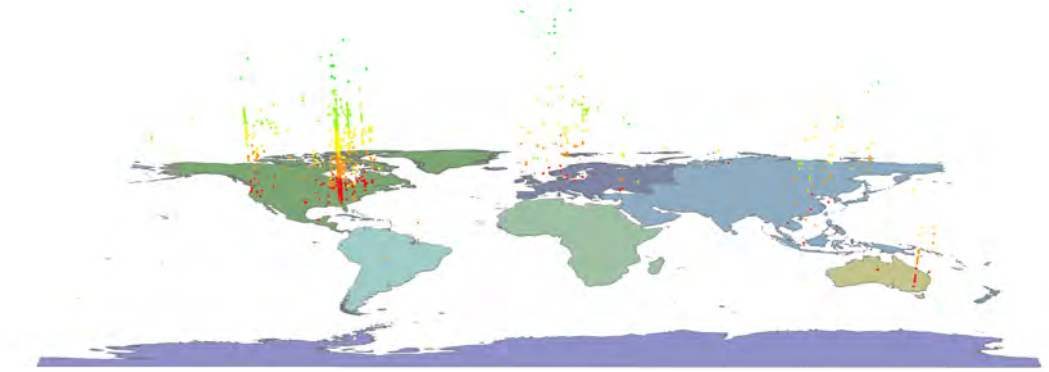

Figure 4 Visualization of Activities in Space and Time Dimensions.

3.1 Visualize Activities and Places in GIS

This tool reads the activity table and place table from PostgreSQL in ArcGIS, and displays the places and activities on a map. In addition, since the activity records have the time stamps, the activities can also be visualized on a 3D map in which the z coordination represents the time an activity took place. The map in Figure 4 is such a map showing the data collected from the case study.

3.2 Create a Place-Based Social Network

This tool allows users to interactively select the places in ArcGIS, and create a social network of the visitors from those places. Participants in the same activity are connected to each other in the social network. In addition to visualize the location-based social network, some network measures, e.g., number of nodes, number of cliques, average clustering coefficient and etc. are also reported in the output.

3.3 Query People-Based Social Network

This tool allows users to query the people-based social network based on a user-defined query which will be translated into an SQL sentence. The user can also visualize and analyze the social networks for the selected people.

3.4 Identify Spatial-Temporal Interactions

The spatial-temporal interactions are identified with the Knox test [2] by using the Pysal python library [6]. This tool reports the identified spatial-temporal interactions based on the user-defined spatial (delta) and temporal (tau) intervals.

4 Conclusion

Current GIS environment is not suitable for the representation and analysis of rich information embedded in location-based social media data due to its lack of capability to represent some of the key components of the data. This research aims to fill the blank by developing an ontological framework and a conceptual data model for multiple-dimensional representation

of geography, time, and social connections. The prototype of the conceptual model is implemented and a case study is carried out. The effectiveness of the LBSMA model is evidenced by a case study which collects and analyzes Facebook data.

The findings of this research yield new insights regarding human activities in virtual and physical space, and will enhance technical capabilities for social media analysis in GIS. The developed methods can help identify place-based or people-based strategies, e.g., urban planning, traffic planning, commercial advertising or energy communicating. The proposed framework paves new avenues for future research, such as public health, transportation, urban geography and social science. Based on the proposed model and prototype, we believe there are many more potential ways to mine the organized datasets. This study has only provided a case study with a few application examples, both of which asked questions that are only related to two dimensions of space, time, and social network. Starting from here, many exciting future research avenues should be explored. Examples include development of new analytical methods and explorations of new application studies, particularly those involve all three dimensions of the LBSMA data.

References

1. Torsten Hägerstrand. What About People in Regional Science? *Papers in Regional Science*, 24(1):7, 1970.
2. E. G. Knox and M. S. Bartlett. The Detection of Space-Time Interactions. *Applied Statistics*, 13(1):25, 1964.
3. Kevin Lewis, Jason Kaufman, Gonzalez A. Marco, Wimmer B. Andreas, and Christakis A. Nicholas. Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks*, 30:330–342, 2008.
4. Yuan Li, Haoyu Gao, Mingmin Yang, Wanqiu Guan, Haixin Ma, Weining Qian, Zhigang Cao, and Xiaoguang Yang. What are Chinese Talking about in Hot Weibos? *arXiv preprint arXiv:1304.4682*, 2013.
5. Mor Naaman, Jeffrey Boase, and Chih-Hui Lai. Is it really about me?: message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pages 189–192. ACM, 2010.
6. Sergio J. Rey and Luc Anselin. PySAL: A Python Library of Spatial Analytical Methods. *The Review of Regional Studies*, 37(1):5–27, 2007.
7. Tim Schwanen and Mei-Po Kwan. Critical Space-Time Geographies: Guest Editorial. *Environment and Planning A*, 44(9):2043–2048, 2012.
8. Daniel Sui and Michael Goodchild. The convergence of GIS and social media: challenges for GIScience. *International Journal of Geographical Information Science*, 25(11):1737, 2011.
9. Yuri Takhteyev, Anatoliy Gruzd, and Barry Wellman. Geography of Twitter networks. *Social Networks*, 34(1):73–81, 2012.
10. Ming-Hsiang Tsou and Michael Leitner. Visualization of social media: seeing a mirage or a message? *Cartography and Geographic Information Science*, 40(2):55, 2013.
11. Xiaobai Yao. Modeling and analyzing cities as spatio-temporal places. In Bin Jiang and Xiaobai Yao, editors, *Geospatial Analysis and Modelling of Urban Structure and Dynamics*, GeoJournal Library, pages 311–328. Springer, Dordrecht, the Netherlands, 2010.
12. May Yuan, Atsushi Nara, and James Bothwell. Space–time representation and analytics. *Annals of GIS*, 20(1):1–9, 2014.