Towards Geo-Social Intelligence: Mining, Analyzing, and Leveraging Geospatial Footprints in Social Media

James Caverlee, Zhiyuan Cheng
Texas A&M University
{caverlee,zcheng}@cse.tamu.edu

Daniel Z. Sui
The Ohio State University
sui.10@osu.edu

Krishna Y. Kamath
Twitter, Inc.
kkamath@twitter.com

Abstract

The widespread adoption of GPS-enabled tagging of social media content via smartphones and social media services (e.g., Facebook, Twitter, Foursquare) uncovers a new window into the spatio-temporal activities of millions of people. These "footprints" open new possibilities for understanding how ideas flow across the globe, how people can organize for societal impact, and lay the foundation for new crowd-powered geosocial systems. We describe recent efforts to mine, model, and analyze large-scale geospatial footprints toward enabling new intelligent geo-social systems that leverage these footprints.

1 Introduction

The exponential growth in social media over the past decade has recently been joined by the rise of *location* as a central organizing theme of how users engage with online information services and with each other. Enabled by the widespread adoption of GPS-enabled smartphones, users are now forming a comprehensive *geo-social* overlay of the physical environment of the planet. For example, the Foursquare location sharing service has enabled over 4.5 billion "check-ins" [13], whereby users can link their presence, notes, and photographs to a particular venue. The mobile image sharing service Instagram allows users to selectively attach their latitude-longitude coordinates to each photograph; similar geo-tagged image sharing services are provided by Flickr and a host of other services. And the popular Twitter service sees 500 million Tweets per day, of which around 5 million are tagged with latitude-longitude coordinates. Confirming this trend, a recent Pew Research Center report finds that location is now an increasingly central part of the social media experience [41].

In contrast to proprietary location-based data collected by many entities – e.g., search engine query logs with an associated IP address that can be resolved to a rough location, cell-phone call records that can pinpoint a user to a particular cell tower, and point-of-sale data collected by retailers – geo-social tags and check-ins are inherently voluntary and public. As a result, they provide a rich and growing body of geo-location evidence that can potentially support basic scientific inquiry into questions that heretofore were difficult for researchers to study. These difficulties were often due to the proprietary nature of traditional location data, the cost of acquiring new data through small lab-based studies (e.g., due to navigating university IRB protocols, overcoming resistance to personal tracking devices), and the difficulty of sharing such

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

sensitive data with other researchers. Not only do voluntarily shared geo-location cues provide an alternative basis for scientific inquiry, in addition, designers of information management systems (e.g., web search systems, social media discovery, personal information management) can integrate these new public location signals into more robust user models, intelligent "location-aware" services, and so forth. Indeed, we believe that the proliferation of these fine-grained (public) spatio-temporal footprints provides an unprecedented opportunity to gain new insights into:

- The dynamics of human behavior and rhythm/pulsation of social life from local to global levels;
- The dynamics of how ideas spread and how people can organize for societal impact; and
- The development of new geo-social information systems that leverage these global-scale geospatial footprints for real-world impact.

Already, we have witnessed compelling new studies along all three of these dimensions, spanning many research communities – including the data mining and machine learning [1, 4, 6, 11, 12, 28, 30], geographic information systems [9, 14, 27, 34, 38], web search and information retrieval [31, 33], and the emerging computational social science paradigm [15, 21, 23, 25, 29, 32, 36, 40]. For example, the dynamics of fundamental human mobility patterns have been modeled from check-ins mined from two location sharing services – Gowalla and Brightkite – and inherent constraints on these patterns by both geographic and social factors have been discovered [6]. Facebook researchers have provided a comprehensive analysis of the distance between Facebook users, leading to new insights into how social networks are impacted by geography [1]. The LiveHoods [10] project has shown how to identify "living neighborhoods" based on the revealed locations and movements of social media users. And new geo-social information systems have been proposed based on these location cues, including earthquake detection from Twitter information flows [31], a local search system that estimates a user's location utilizing the aggregate signals from the check-ins with real-time contextual information [33], and an event discovery system that organizes spatio-temporal footprints and corresponding media to allow consumers to travel through space and time to experience the world's stories [7].

Challenges. Yet there are key challenges to delivering on this promise. First, many users in social media reveal broad, imprecise locations (e.g., at the city or state level), while others provide fine-grained latitude-longitude information. In particular, users are less likely to post precise locations such as street addresses on Twitter and related services. How can these multiple location granularities be integrated to account for uncertainty at different levels? Second, models based on users who do willingly share fine-grained location information will necessarily be biased away from the general population of social media users (and more generally, from the underlying population). How can we model and assess the impact of this bias (and its ultimate impact on applications like local information access or expert finding)? Third, personal location-revealing information may be interspersed in an inherently noisy stream of updates reflecting many daily interests (e.g., food, sports, daily chatting with friends). Are there clear location signals embedded in this mix of topics and interests that can be accurately extracted? Fourth, not all relationships in social media are the same. Some ties are stronger than others, and presumably some ties are more impactful on the development of a location-oriented user profile. How does this variable tie strength nature impact these profiles?

Even the design space of geo-social information systems is not clearly understood. For example, the ecosystem is driven both by the many users of current location sharing technologies who explicitly share location across multiple platforms (e.g., via Twitter geo-tagged posts, by Foursquare check-in) as well as the many "non-sharing" users who still leave implicit clues based on their non-geo-tagged content posts, via location-revealing images left on their social networks, and other implicit signals. How do new systems integrate these disparate signals? Do users perceive a difference in ownership of their "location" in scenarios where they explicitly reveal it versus it being inferred from large-scale data-driven approaches (e.g., by applying machine learning approaches)? And how does information access in geo-social information systems differ from traditional web search and friend finding in social networks? These and related questions lead us to believe that there is a compelling need for new techniques for mining, analyzing, and leveraging geospatial footprints in social media.

Roadmap. The rest of this paper highlights two of our parallel efforts towards the goal of enabling new geo-social intelligence. The first focuses on *the dynamics of ideas* via an exploration of fine-grained Twitter based geospatial footprints, coupled with a *predictive analytics application* that seeks to estimate the popularity of future ideas (approximated by Twitter hashtags) in particular locations. The second focuses on the *dynamics of people* via an exploration of location sharing through services like Foursquare, coupled with a prototype *location-based search* system that takes advantage of these new signals. We conclude with final thoughts on the future of geo-spatial intelligence research at the intersection of the emerging spatial computing and computational social science.

2 GeoSocial Footprints: Modeling and Predicting Spatial Diffusion

Geospatial footprints provide a new perspective on how information is shared at a global scale. Understanding how ideas are adopted, how new communities form and evolve, and how these activities shape opinions and actions are all long-standing questions, e.g. [19, 20, 22, 26, 39]. With access to fine-grained geo-spatial footprints, we face new opportunities to investigate the spatio-temporal dynamics of ideas. In this section, we highlight our research on analyzing the spatial diffusion of one type of social media – Twitter hashtags – and in building predictive models of what hashtags will ultimately be popular where. In addition to impacting these long-standing foundational questions about information diffusion, such investigations have the potential to impact the design of a variety of systems and applications, including targeted advertising, location-based services, social media search, and content delivery networks.

2.1 Dynamics of Ideas: Spatial Diffusion

We begin by describing a study of the spatio-temporal properties of social media spread through an examination of the fine-grained sharing of one type of global-scale social media – a sample of 2 billion geo-tagged Tweets with precise latitude-longitude coordinates collected over the course of 18 months. The study itself (reported more fully in [18]) focuses on the propagation of hashtags across Twitter, where a hashtag is a simple user-generated annotation prefixed with a # and serves as a simplified "semantic" marker that we can track across space and time. Questions we have explored include: (i) What role does distance play in the adoption of hashtags? Does distance between two locations influence both what users in different locations adopt and when they do so? (ii) While social media is widely reported in terms of viral and global phenomenon, to what degree are hashtags truly a global phenomenon? (iii) What are the geo-spatial properties of hashtag spread? How do local and global hashtags differ? (iv) How fast do hashtags peak after being introduced? And what are the geo-spatial factors impacting the timing of this peak? (v) How can the spatio-temporal characteristics of hashtags describe locations? Are some locations more "impactful" in terms of the hashtags that originate there?

Here we focus on two findings: the geo-spatial constraints of information sharing and peak analysis.

Geo-spatial constraints. Let's begin by modeling a location by the set of hashtags that have been observed there over a particular time window. We could define a "location" using a simple grid technique overlaid on the globe of equal-area locations, or equal-population locations, or encode some other semantic cues (e.g., by metropolitan statistical area). What then is the impact of distance between two locations on the adoption of an idea? Tobler's first law of geography [37] states that locations that are closer to each other are more alike; hence, it is reasonable to assume that nearby locations will share similar ideas (here, represented by

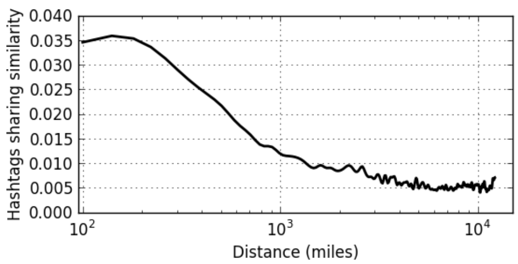


Figure 1: Hashtag sharing similarity vs Distance.

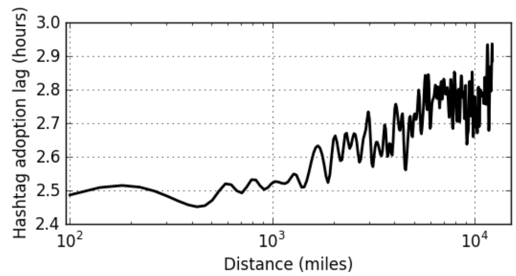


Figure 2: Hashtag adoption lag vs Distance

hashtag adoption). Alternatively, there is growing evidence of the "flattening" of the world as information spreads through Internet-enabled (geographically-neutral) media.

Given two locations, we can measure their hashtag "similarity" using the Jaccard coefficient between the sets of hashtags observed at each location: Hashtag Similarity$(l_i, l_j) = \frac{H_{l_i} \cap H_{l_j}}{H_{l_i} \cup H_{l_j}}$, where H_l is the set of unique hashtags observed in a location l. Locations that have all hashtags in common have a similarity score of 1.0, while those that share no hashtags have a score of 0.0.

Similarly, we can measure the adoption lag between two locations, to capture the time of when a new hashtag made its first appearance. Locations that adopt a common hashtag at the same time can be considered as more temporally similar than are two locations that are farther apart in time (with a greater lag). Letting t_l^h be the first time when hashtag h was observed in location l, we can define the hashtag adoption lag of two locations as: Adoption Lag$(l_i, l_j) = \frac{1}{|H_{l_i} \cap H_{l_j}|} \sum_{h \in H_{l_i} \cap H_{l_j}} |t_{l_i}^h - t_{l_j}^h|$, where the adoption lag measures the mean temporal lag between two locations for hashtags that occur in both the locations. A lower value indicates that common hashtags reach both the locations around the same time.

The relationship between hashtag similarity and distance is plotted in Figure 1. We see a strong correlation, suggesting that the closer two locations are, the more likely they are to adopt the same hashtags. As distance increases, the hashtag sharing similarity drops accordingly. Similarly, we see in Figure 2 a relatively flat temporal adoption relationship up to ~500 miles, then a generally positive correlation, suggesting that locations that are close in spatial distance tend also to be close in time (e.g., they adopt hashtags at approximately the same time). Locations that are more spatially distant tend to adopt hashtags at greater lags with respect to each other. These findings suggest the strong impact of geographic constraints (representing language commonalities, culture sharing) on meme spreading.

Peak Analysis. We next zoom in on the spatial properties of hashtag propagation during the minutes pre- and post- peak. When hashtags peak, do they peak suddenly in different locations simultaneously or do they slowly accumulate a larger spatial footprint? What are the dynamics of their spatial properties as they become popular? For every hashtag ($h \in H$) and location ($l \in L$) pair, if we let O_l^h be the set of all occurrences of h in l, then the probability of observing hashtag h in location l is defined as $P_l^h = \frac{O_l^h}{\sum_{l \in L}\{O_l^h\}}$ and the *hashtag entropy* is defined as $\mathcal{E}^h = -\sum_{l \in L} P_l^h \log_2 P_l^h$, which measures the randomness in spatial distribution of a hashtag and determines the minimum number of bits required to represent the spread. A hashtag that occurs in only a single location will have an entropy of 0.0. As a hashtag spreads to more locations, its entropy will increase, reflecting the greater randomness in the distribution.

Here we divide each hashtag's lifecycle into equal length time intervals of 10 minutes. For each time interval, we compute the hashtag entropy ($\mathcal{E}^h(t)$) over just that interval. We plot this interval-specific entropy in Figure 3. We observe that hashtags reach their lowest interval focus and highest interval entropy about 10-20 minutes after their peak. Rather than peaking with their most "global" footprint, hashtags instead reach this state *after* their peak. In effect, this single location is "championing" a hashtag. In the 10-20

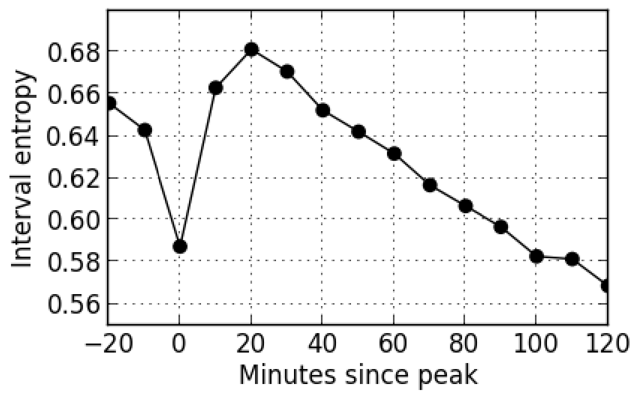


Figure 3: Hashtags peak with their most "global" footprint 10-20 minutes after their peak

minutes after this peak period, other locations adopt the hashtag, resulting in a decrease in interval focus and an increase in entropy as the hashtags becomes more global. About 30 minutes after reaching peak, focus and entropy reverse, with focus increasing and entropy decreasing as the hashtag withdraws back to its original focus location. In essence, hashtags are spread via a single location "championing" a hashtag initially, spreading it to other locations and then continuing to propagate it after it has become popular. In [2], the authors observed a similar pattern for YouTube videos which they called the "spray-and-diffuse" pattern. Our observations over hashtags suggest that this pattern may be a fundamental property of social media spread.

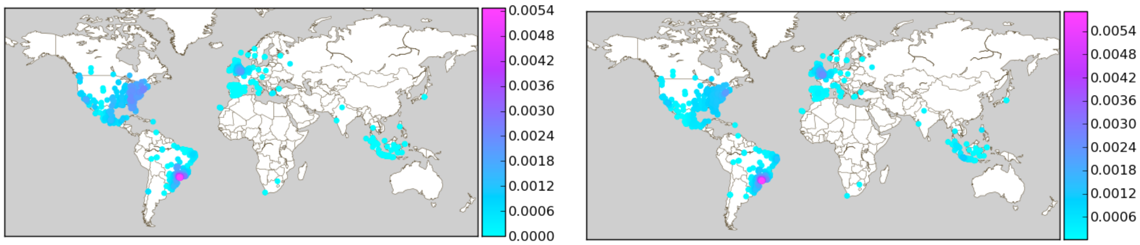
2.2 Dynamics of Ideas: Predictive Analytics

Given these and related insights into the spatio-temporal spread of ideas, we can directly inform the design of distributed content delivery networks and search infrastructure for real-time Twitter-like content. For example, caching decisions to improve fast delivery of social media content to users and to support applications like real-time search can build upon the results presented here. Insights into the role distance plays and the impact locations have on hashtag spread could inform new algorithms for geo-targeted advertising. This work can also complement efforts to model network structures that support (or impede) the "viralness" of social media, measure the contagion factors that impact how users influence their neighbors, develop models of future social media adoption, and so forth.

Towards putting these results into practice, we have developed new predictive models to estimate what hashtags will eventually be popular where [17]. For example, can we accurately predict which hashtags will be popular in San Francisco over the next two hours? Can the same model also predict which hashtags will be popular in a small town like College Station, Texas? Can we identify which hashtags that have been popular in New York in the past two hours but will drop in interest? Toward answering these questions, we have develop a reinforcement learning-based approach that builds upon two competing hypotheses of information spread over geo-spatial networks we first discussed in the previous section.

- Spatial Affinity: The first hypothesis, based on Tobler's first law of geography, states that the information spread between two locations is impacted by the distance between two locations. For example, according to this hypothesis hashtags spread faster between San Francisco and Mountain View, since they are closer to each other; but slower between San Francisco and Austin.

- Community Affinity: The second hypothesis is that the "world is flat" and information spreads based on virtual communities enabled by the prevalence of the Internet. In this hypothesis, distance is less important than are the strength of these virtual ties between locations; e.g., under this hypothesis



(a) Predicted (estimated using spatial influence model after 5 minutes)

(b) Actual (real distribution after 2 hours of propagation),

Figure 4: Example of using spatial influence model for the hashtag #ripstevejobs

San Francisco and Austin may be considered closer in terms of common interest (and hence, hashtags should flow more rapidly between the two), rather than Austin and its more proximate neighbor Houston.

We have investigated a series of features inspired by these two hypotheses for predicting which hashtags will be popular in a specific location at a specific time. An example of modeling propagations using the spatial influence model for the hashtag #ripstevejobs is shown in Figure 4. We predicted the future distribution of this hashtag using the spatial influence model based solely on its initial (first 5 minutes) distribution. The comparison between the predicted and actual distribution is shown in Figure 4(a) and Figure 4(b) respectively. We observe that the relative distribution (indicated by color) and its values (indicated by scale) are very close to each other. In our experimental evaluation over 755 million geo-tagged tweets, we find the best approach is able to predict close to 70% of future hashtags occurrences accurately. Interestingly, both the spatial affinity and community affinity models are valid for *particular hashtags* and for *particular locations*, suggesting the interplay of both geographic constraints and community homophily across large distances impacting what ideas flow where.

3 GeoSocial Footprints: Activity-Based Information Access

While geosocial footprints provide evidence of idea spread, they also can provide an insight into actual human movements. This human mobility and the interaction of human movement with space provides a fascinating and unique opportunity. What do large-scale voluntarily contributed human mobility data reveal? And how can these insights be incorporated into the design of new mobile+location-based services, traffic forecasting, urban planning, and models of disease spread?

3.1 From Dynamics of People to Location-Based Information Search

Toward better understanding the spatial, temporal, and social characteristics of how people use these services, we have engaged in a large-scale study of location sharing services in [5] that focuses on the wheres and whens of over 22 million check-ins across the globe. Specifically, we have studied human mobility patterns revealed by these check-ins and explored factors that influence this mobility, including social status, sentiment, and geographic constraints. We have found (i) that locations can be modeled by the activity patterns of people; and (ii) that people follow simple, reproducible patterns – motivating us to explore whether activity patterns (i.e., the temporal dynamics of check-ins) can be used to augment traditional information access (see, e.g., [3]) through a prototype location-based search system. In many ways analogous to how clickstreams have been successfully incorporated into traditional search systems based on content similarity and link analysis by connecting real-world user actions (clicks) to relevance, this prototype framework mines the spatio-temporal activity patterns of location-based crowds for augmenting traditional location-based search and for supporting enhanced location recommendations. In particular, we have developed an activity pattern-driven approach for supervised location categorization, wherein activity patterns can be used to accurately predict the semantic category of uncategorized locations. We augment this approach with an activity-driven location clustering algorithm that can group semantically related locations, and we show how activity-driven semantic organization of locations may be naturally incorporated into location-based web search. We conclude with an activity-driven location recommendation system that incorporates multiple factors, including physical distance, semantic correlation, social desires, interest-based reputation, and temporal dynamics.

Enhancing Local Search. Activity patterns and category information for venues can be easily incorporated into traditional location-based search to answer the information need for traffic. One scenario for answering the activity-driven query is: Karen is searching for a restaurant which is off-peak during dinner time between 5 - 7 PM, so that she can enjoy the quiet environment talking with her friends. Knowing the activity patterns and category for venues, the system could easily retrieve the venues nearby in the category of food, and rank the results by the descending order of busyness.

Location Recommendation. Another use case is recommendation of venues having similar activity patterns. For example, Jerry plays a lot of basketball, and tennis. He usually goes to the Williams Park during Wednesday early evening, and Saturday afternoon, which are both free time for him and peak times for guys to get-together and play basketball and tennis. Recently, he moves to a new neighborhood, and wants to find places nearby that have similar activity patterns, so that he can meet new friends there and play some basketball or tennis. A activity-driven location-based search can also easily handle this kind of queries. Given the name of the venue, the system calculates temporal similarity between activity patterns of the venue and other venues in the same category (or in other categories as well), and return the locations with the highest temporal similarities.

4 Conclusion

We believe that the increasing ubiquity of location-based social media has the potential to fundamentally disrupt basic scientific inquiry into questions that heretofore were difficult to study and to provide the basis for new "intelligent" geo-social information systems. Accomplishing this will require new methods, new algorithms, and new frameworks for mining and analyzing vast fine-grained (public) spatio-temporal footprints, as well as new systems and techniques to leverage these footprints. We have outlined some of the challenges facing this opportunity and highlighted two of our related efforts toward informing this emerging research area. Moving forward, we believe that geo-social intelligence research is poised to make major breakthroughs in the years to come due to the growing interests of social scientists in computational/data-intensive approaches and the 4th paradigm [16, 24] and computer scientists in spatial computing [8]. We also believe that transformative research in geo-social system can be accelerated along multiple fronts if we continue to embrace and fine-tune the emerging open science paradigm [35] to promote interdisciplinary collaboration and improve the infrastructure for geo-social intelligence research.

References

[1] L. Backstrom, E. Sun, and C. Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. In *WWW*, 2010.

[2] A. Brodersen, S. Scellato, and M. Wattenhofer. Youtube around the world: geographic popularity of videos. In *WWW*, 2012.

[3] Z. Cheng, J. Caverlee, K. Y. Kamath, and K. Lee. Toward traffic-driven location-based web search. In *CIKM*, pages 805–814, 2011.

[4] Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet: a content-based approach to geo-locating twitter users. In *CIKM*, 2010.

[5] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui. Exploring millions of footprints in location sharing services. In *ICWSM*, 2011.

[6] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.
[7] T. Coletta. Seen, by mahaya: A tool to sort the mess of social media. http://www.popularmechanics.com/technology/gadgets/news/seen-by-mahaya-a-tool-to-sort-the-mess-of-social-media-15510890/, 2013. [Online; accessed Sep 24, 2013].
[8] C. C. Consortium. From gps and virtual globes to spatial computing - 2020: The next transformative technology. Technical report, 2013.
[9] J. W. Crampton, M. Graham, A. Poorthuis, T. Shelton, M. Stephens, M. W. Wilson, and M. Zook. Beyond the geotag: situating Ôbig dataÕ and leveraging the potential of the geoweb. *Cartography and Geographic Information Science*, 40(2):130–139, 2013.
[10] J. Cranshaw, R. Schwartz, J. I. Hong, and N. M. Sadeh. The livehoods project: Utilizing social media to understand the dynamics of a city. In J. G. Breslin, N. B. Ellison, J. G. Shanahan, and Z. Tufekci, editors, *ICWSM*. The AAAI Press, 2012.
[11] C. Davis, G. Pappa, D. de Oliveira, and F. de L Arcanjo. Inferring the location of twitter messages based on user relationships. In *Transactions in GIS*, 2011.
[12] J. Eisenstein, B. O'Connor, N. Smith, and E. Xing. A latent variable model for geographic lexical variation. In *EMNLP*, 2010.
[13] Foursquare. About foursquare, April 2012.
[14] D. D. Ghosh and R. Guha. What are we ÔtweetingÕ about obesity? mapping tweets with topic modeling and geographic information system. *Cartography and Geographic Information Science*, 40(2):90–102, 2013.
[15] B. Hecht, L. Hong, B. Suh, and E. Chi. Tweets from justin bieber's heart: the dynamics of the location field in user profiles. In *SIGCHI*, 2011.
[16] T. Hey, S. Tansley, and K. M. Tolle, editors. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, 2009.
[17] K. Y. Kamath and J. Caverlee. Spatio-temporal meme prediction: Learning what hashtags will be popular where. In *CIKM*, 2013.
[18] K. Y. Kamath, J. Caverlee, K. Lee, and Z. Cheng. Spatio-temporal dynamics of online memes: A study of geo-tagged tweets. In *WWW*, 2013.
[19] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *KDD*, 2003.
[20] D. Kempe, J. Kleinberg, and É. Tardos. Influential nodes in a diffusion model for social networks. In *IN ICALP*, 2005.
[21] G. King. Ensuring the data-rich future of the social sciences. *Science*, 331(6018):719–721, 2011.
[22] G. Kossinets, J. Kleinberg, and D. Watts. The structure of information pathways in a social communication network. In *KDD*, 2008.
[23] J. Kulshrestha, F. Kooti, A. Nikravesh, and K. Gummadi. Geographic dissection of the twitter network. In *International AAAI Conference on Weblogs and Social Media*, 2012.
[24] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. BarabĞsi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009.
[25] K. Leetaru, S. Wang, G. Cao, A. Padmanabhan, and E. Shook. Mapping the global twitter heartbeat: The geography of twitter. *First Monday*, 18(5), 2013.
[26] K. Lerman and R. Ghosh. Information contagion: an empirical study of the spread of news on digg and twitter social networks. *CoRR*, 2010.
[27] L. Li, M. F. Goodchild, and B. Xu. Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*, 40(2):61–77, 2013.
[28] R. Li, S. Wang, H. Deng, R. Wang, and K. Chang. Towards social user profiling: unified and discriminative influence model for inferring home locations. In *KDD*, 2012.
[29] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman. I'm the mayor of my house: Examining why people use foursquare - a social-driven location sharing application. In *SIGCHI*, 2011.
[30] A. Sadilek, H. Kautz, and J. Bigham. Finding your friends and following them to where you are. In *WSDM*, 2012.
[31] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social

sensors. In *WWW*, 2010.
[32] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo. Socio-spatial properties of online location-based social networks. In *ICWSM*, 2011.
[33] B. Shaw, J. Shea, S. Sinha, and A. Hogue. Learning to rank for spatiotemporal search. In *WSDM*, 2013.
[34] A. Stefanidis, A. Cotnoir, A. Croitoru, A. Crooks, M. Rice, and J. Radzikowski. Demarcating new boundaries: mapping virtual polycentric communities through social media content. *Cartography and Geographic Information Science*, 40(2):116–129, 2013.
[35] D. Z. Sui. Opportunities and impediments of open gis. *Transactions in GIS (in press)*, 2014.
[36] Y. Takhteyev, A. Gruzd, and B. Wellman. Geography of twitter networks. *Social Networks*, 34(1):73 – 81, 2012.
[37] W. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46(2), 1970.
[38] C. Xu, D. W. Wong, and C. Yang. Evaluating the Ògeographical awarenessÓ of individuals: an exploratory analysis of twitter data. *Cartography and Geographic Information Science*, 40(2):103–115, 2013.
[39] J. Yang and J. Leskovec. Modeling information diffusion in implicit networks. In *ICDM*, 2010.
[40] S. Yardi and D. Boyd. Tweeting from the town square: Measuring geographic local networks. In *ICWSM*, 2010.
[41] K. Zickuhr. Location-based services. *Pew Research Center*, September 2013.