RC25250 (D1112-002) December 23, 2011
Computer Science

IBM Research Report

Effective Event Discovery: Using Location and Social Information for Scoping Event Recommendations

Elizabeth M. Daly
IBM Research
Smarter Cities Technology Centre
Mulhuddart
Dublin 15, Ireland

Werner Geyer
IBM Research Division
One Rogers Street
Cambridge, MA 02142
USA

Effective Event Discovery: Using Location and Social Information for Scoping Event Recommendations

Elizabeth M. Daly and Werner Geyer
IBM T.J. Watson Research
Cambridge, MA 02116, USA
{dalyeliz, wgeyer}@us.ibm.com

ABSTRACT

The ever blurring line between online interactions and physical encounters presents an interesting challenge when recommending events. Events created on social networking sites may have ambiguous location scope. The location information provided may be fuzzy or non existent and additionally the reach and radius of interest in the event can vary greatly. In this work, we identify four categories of events: global, location dependent and socially independent, socially dependent and location independent, and location and socially dependent. We classify events from an organizations internal event management service where the location of the event is unknown, but the location of the attendees are known in order to improve scoping of event recommendations. Our results, investigate the impact of ignoring location properties when recommending events using classic collaborative filtering techniques. Additionally, once global and socially independent events are identified, they can be used to provide recommendations to new users, addressing the cold-start problem.

Categories and Subject Descriptors

H.5.3 [**Information Systems**]: Collaborative computing; J.2 [**Computer Applications**]: Social and Behavioral Sciences

General Terms

Algorithms, Human Factors

Keywords

Location-Based, Social-Networks, Recommender Systems

1. INTRODUCTION

Recommender systems have been used to provide recommendations of everything from consumer products, movies, restaurants and even friends. Location-based recommendations have gained increasing prevalence with the growing use of mobile phone location services. Generally, these recommendations are based on a user's proximity to a known location-based resource, such as a restaurant.

Events provide an interesting challenge where interest in an event may or may not be location-based. Additionally, even if there is a location-based element to the event, the radius of interest in the event may vary from event to event. For example, an event broadcasting a gig in a small local venue will have a dramatically different scope of interest when compared to an event such as a large scale music festival or an international conference. In this case, the dependency on the user location to infer relevance is less significant, however remains a factor. Moreover, events are also inherently social since they provide opportunities to meet old friends or make new friends, i.e., the interest in an event is socially dependent. For example, a local gallery opening might be more interesting if known friends are attending, or well-known artists. Considering these two factors, the radius of interest covers four types of events: local and socially dependent, local and socially independent, global and socially dependent, and global and socially independent.

Previous research has focused on matching users' interest in events using *content-based* approaches with an increasing body of location-based approaches. In this paper, we argue that events are both location dependent and socially dependent and we study the influence of these two factors on user interest for the purpose of recommendations. For our analysis, we focus on a large event data set from an event management service representing internal IBM events from all across the globe. The location of these events is often unknown and as a result, the locations of the attendees are used to infer the location information of the event. Additionally, it is assumed that not all events are location-based given they may be an online event or one that is of global interest to the user population. As opposed to content-based recommendation techniques, *collaborative-filtering* matches items that similar users have previously found interesting and in essence tries to match similar users assuming they will be interested in the same items [2]. In this work, we use collaborative-filtering because it has the advantage that it requires no knowledge about the content of the item. In the case of events, the title may only be available which may have limited use in providing recommendations. We present a mechanism for classifying events based on their social and location based properties, and investigate the weaknesses of collaborative based filtering for location dependent events. Additionally, we illustrate how these classifications may be harnessed to recommend events to new users.

2. RELATED WORK

Quercia et al. investigate recommending social events to cold start users based on their geographic location history where the location of the event is known [4]. Scellato et al. performed an analysis of the online location-based social networking applications Bright Kite, Foursquare and Gowalla [6]. Their analysis found social connections span large geographic distances, despite the location based element of these services. Minkov et al. present a collaborative approach to recommending events and found the collaborative approach performed better than a content based approach, however location of the event is not considered [3]. Takeuchi and Sugimoto explore recommending locations based on the locations a user frequently visits where the items they are recommending are at a fixed location [7].

3. EVENT CLASSIFICATION

In order to infer the social and location scope of an event where the location is unknown, the location of the users attending the event and the social ties between them may be used. Taking into account these two factors we identify four categories of events shown in figure 1:

- **Local Not Social:** These events are highly clustered around a given location, however there is little existing historical overlap between the attendees.
- **Social And Local:** These events have a high amount of social overlap between the attendees and are focused around a small number of geographically close locations.
- **Social Not Local:** Given the large number of online related events, attendees are not necessarily geographically co-located and but do have a history of attending the same events suggesting social ties amongst the attendees.
- **Global:** These events are those where the attendees are geographically distributed and have little overlap in attending similar events. These could be online web meetings that require all employees to attend.

By taking into account the locations based properties of the attending users, we can infer the spatial scope of the event even, when the location of the event is unknown. The social based properties are derived using overlapping attendance at previous events.

3.1 Social Scoping

The social scoping of the event is determined based on the previous history between two users attending overlapping events which may be used to derive the social network. The assumption is that if a large proportion of the attendees have attended events together in the past, then the event may have a social aspect to it. A social edge E_{soc} is created between each attendee that has previously attended the same event to form graph G_{soc}. The social element of the event is quantified by using the density of the graph $G_{soc} = (V, E_{soc})$ measuring how many edges exist in set E_{soc} compared to the maximum number of possible edges between vertices V [1].

$$SocialDensity = \frac{2|E_{soc}|}{|V||V-1|} \tag{1}$$

Table 1: w3Inviter Event Data

Users	Events	Active Users	Active Events
81191	6168	8934	3314

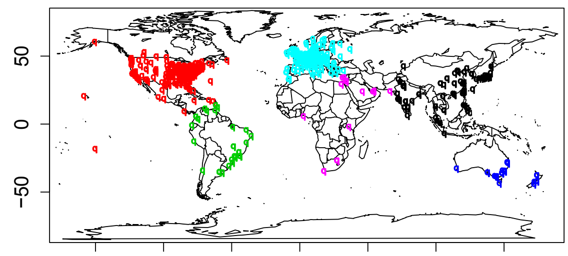


Figure 2: w3Inviter User Locations

3.2 Location Scoping

In order to infer the location scoping of the event we retrieve the location information in the form of geographic coordinates for users who attended the event. An event where attendees are from various, distant locations are more likely to represent online events where the location of the attendee is less relevant. A spatial social edge E_{loc} is created between each attendee to form a graph G_{loc}. The length associated with each edge is calculated based on the distance D_{ij} between two nodes i and j given their geographic coordinates. The radius of interest of an event may be determined based on the geographic distances between all edges in the graph G_{loc}. For our purposes the distance scoping is taken as the 90-percentile of all edge distances in order to exclude outliers. Additionally, the centroid of the event may be determined taking into account the geographic location of the attendees.

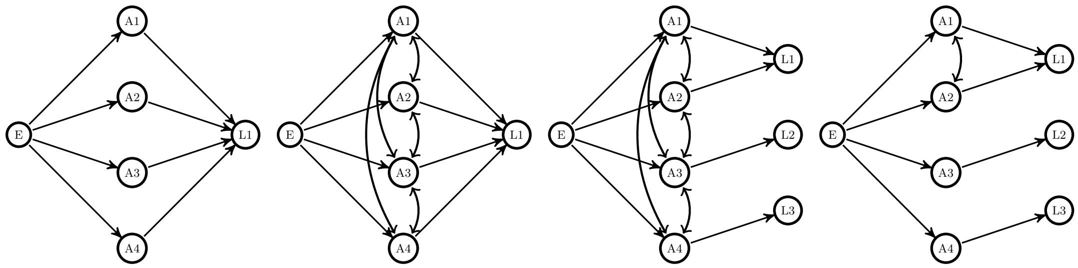
4. ANALYSIS

We used an IBM internal event management service to evaluate the impact of understanding social and location scope on recommendations.

4.1 Dataset

w3Inviter data: The w3Inviter service provides event management within the organization. Events include face-to-face meetings and talks along with online events hosted through web meetings and conferencing services. Users register interest in attending an event by requesting an invite to be added to their business calendar. The reported work location of each user is collected using the company profile directory and these locations are mapped to geographic coordinate locations. Users are located across the world in 716 unique locations shown in figure 2. Out of the 6168 events we focus on events where there were at least 5 participants, leaving 3314 events for analysis.

The social and location properties of each event are measured as described in subsections 3.1 and 3.2. We performed k-means clustering on the events where $k = 4$ based on the metrics of the 90th geographic distance of the attendees, and the social density coefficient as shown in figure 3. As can be seen, four clusters have been identified highlighting the four categories described in section 3.



(a) Local Not Social (b) Social And Local (c) Social Not Local (d) Global

Figure 1: Event Classification

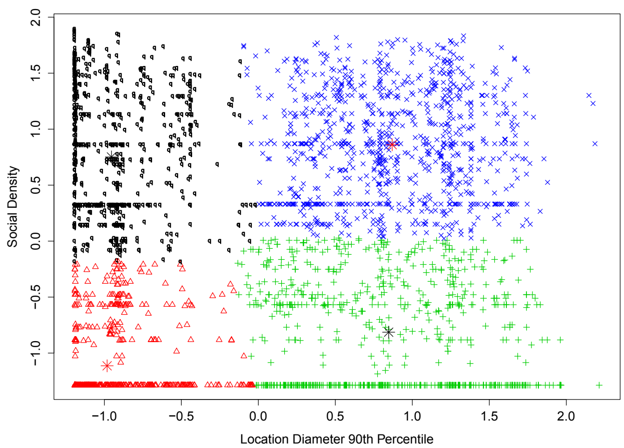


Figure 3: w3Inviter Event Classification

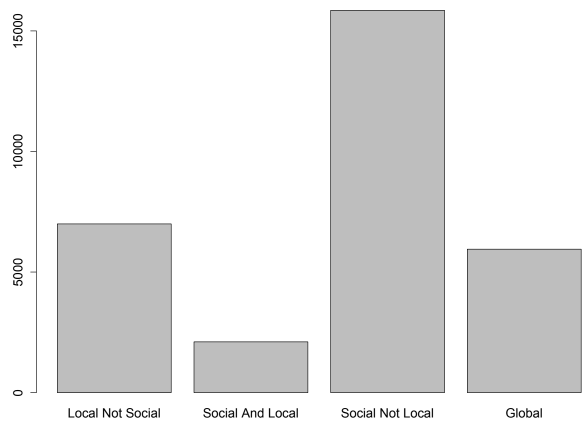


Figure 4: Event Recommendation Distribution

4.2 Location-based Impact on Recommendations

To understand the impact of scoping events taking into account the location and social based categorization, we analyze the recommendations computed using a k-Nearest Neighbor collaborative filtering recommender strategy using the LogLikelihood Similarity model [5]. Figure 4 shows the recommendation frequency across the different categories. For the purpose of this evaluation, recommendations are only considered for users who attended at least 5 events. As can be seen, events that have a high social interaction and are not locally based are strongly represented in the recommendation set, even though the cluster sizes of each category are approximately equal. The second highest number of recommendations comes from local events with low social interaction. We evaluated these recommendations on the basis of information retrieval metrics of the top-N where N=10. The precision and recall are 40.2% and 34.5% respectively. The low precision is a result of the fact that many of the users attend a relatively small number of events.

When examining the recommended events that have been classified as location dependent, 11% of the recommended events were outside of the scope of the users location suggesting location scope should be incorporated into the recommendation scoring.

4.3 New User Recommendations

Providing recommendations for new users with little history presents a huge challenge for recommender systems. A large portion of the users in the dataset have attended less than 5 events as shown in figure 5. In order to evaluate if the classification may aid in providing recommendations to cold-start users, we randomly select 10% of the user population who have attended less than 5 events. Events from the individual classifications are retrieved and the top-N results are scored based on popularity, $Rank_{pop}$ and location proximity $Rank_{loc}$ which is based on the user distance from the calculated centroid of the event location based on the attendees geographic coordinates.

Table 2 shows the percentage of new users where an event they attended was successfully returned in the top-N results. As can be seen recommendation of events that are locally scoped @10, and @20 produce better recommendations when ranked based on proximity data. Recommendations for events that are not location dependent achieve better performance when ranking based on popularity. Interestingly, recommendations for events that are either local and not social or global produce the two strongest results, supporting our assumption that new users may benefit from being recommended events that do not contain a strong social element. The improvement of event ranking based on location and popularity evens out for local non-social events

Table 2: Recommendations for New Users Based on Classification

Cluster	$Rank_{pop}$@10	$Rank_{loc}$@10	$Rank_{pop}$@20	$Rank_{loc}$@20	$Rank_{pop}$@50	$Rank_{loc}$@50
Local Not Social	4.31	**7.29**	7.47	**9.12**	**10.66**	10.55
Social And Local	4.26	**5.37**	5.89	**8.7**	10.00	**11.66**
Social Not Local	**4.12**	0.55	**7.05**	1.63	**11.20**	3.65
Global	**11.09**	1.14	**14.21**	2.04	**16.87**	3.87

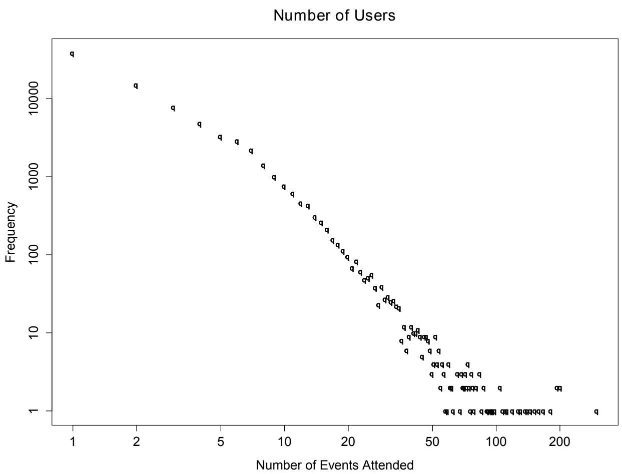


Figure 5: Number of Events Attended Distribution

@50. Though these success values are low, history based recommendations are not available for such users, and as a result recommending events that are geographically scoped to the user provides some means to support new user recommendations. Whereas if the event is location independent, popularity may be used.

5. CONCLUSION

This paper extends existing work on collaborative filtering by analyzing the role of social and location properties for event recommendations. Although proximity has become a popular recommendation mechanism for many mobile event applications, our data indicates that location independent events can be highly relevant when a strong social component exists. We have identified four categories of events and presented a mechanism for event classification based on social and location-based properties where the location and radius of interest of the event is unknown. Social classification takes into account the social density of the participants, assuming that events with a strong social element may be less relevant to new users than those that are of more general interest. Location classification is achieved by inferring the radius of interest in an event based on the distance between attendees and the centroid of these locations. Once the different categories are identified, appropriate ranking may be used to recommend items to new users, addressing the cold-start problem. Experimentation demonstrates that if the event is location independent, popularity is a reasonable metric for recommending events to new users, when the event is classified as location dependent, taking into account the scope of the event and the proximity of the user is more appropriate.

Our analysis showed that by understanding the different classification of the events, an appropriate ranking measure based on user proximity and the popularity of an event may be applied in order to make event recommendations more effective. Future work will explore incorporating event classification into the recommendation scoring mechanism.

6. ACKNOWLEDGMENTS

We wish to thank Joshua Scribner for sharing the w3Inviter data with us.

7. REFERENCES

[1] Estimation of Sparse Jacobian Matrices and Graph Coloring. In *SIAM Journal on Numerical Analysis*, number 20(1), pages 187–209.

[2] John S. Breese, David Heckerman, and Carl Kadie. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In *Proceedings of the14th Annual Conference on Uncertainty in Artificial Intelligence (UAI98)*, pages 43–52, 1998.

[3] Einat Minkov, Ben Charrow, Jonathan Ledlie, Seth Teller, and Tommi Jaakkola. Collaborative future event recommendation. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, CIKM '10, pages 819–828, New York, NY, USA, 2010. ACM.

[4] Daniele Quercia, Neal Lathia, Francesco Calabrese, Giusy Di Lorenzo, and Jon Crowcroft. Recommending Social Events from Mobile Phone Location Data. In *2010 IEEE 10th International Conference on Data Mining (ICDM)*, pages 971–976, December 2010.

[5] Badrul Sarwar, George Karypis, Joseph Konstan, and John Reidl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, WWW '01, pages 285–295, New York, NY, USA, 2001. ACM.

[6] Salvatore Scellato, Anastasios Noulas, Ranaud Lambiotte, and Cecilia Mascolo. Socio-spatial Properties of Online Location-based Social Networks. In *In Proceedings of Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, July 2011.

[7] Yuichiro Takeuchi and Masanori Sugimoto. CityVoyager: An Outdoor Recommendation System Based on User Location History. In Jianhua Ma, Hai Jin, Laurence Yang, and Jeffrey Tsai, editors, *Ubiquitous Intelligence and Computing*, volume 4159 of *Lecture Notes in Computer Science*, chapter 64, pages 625–636. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2006.