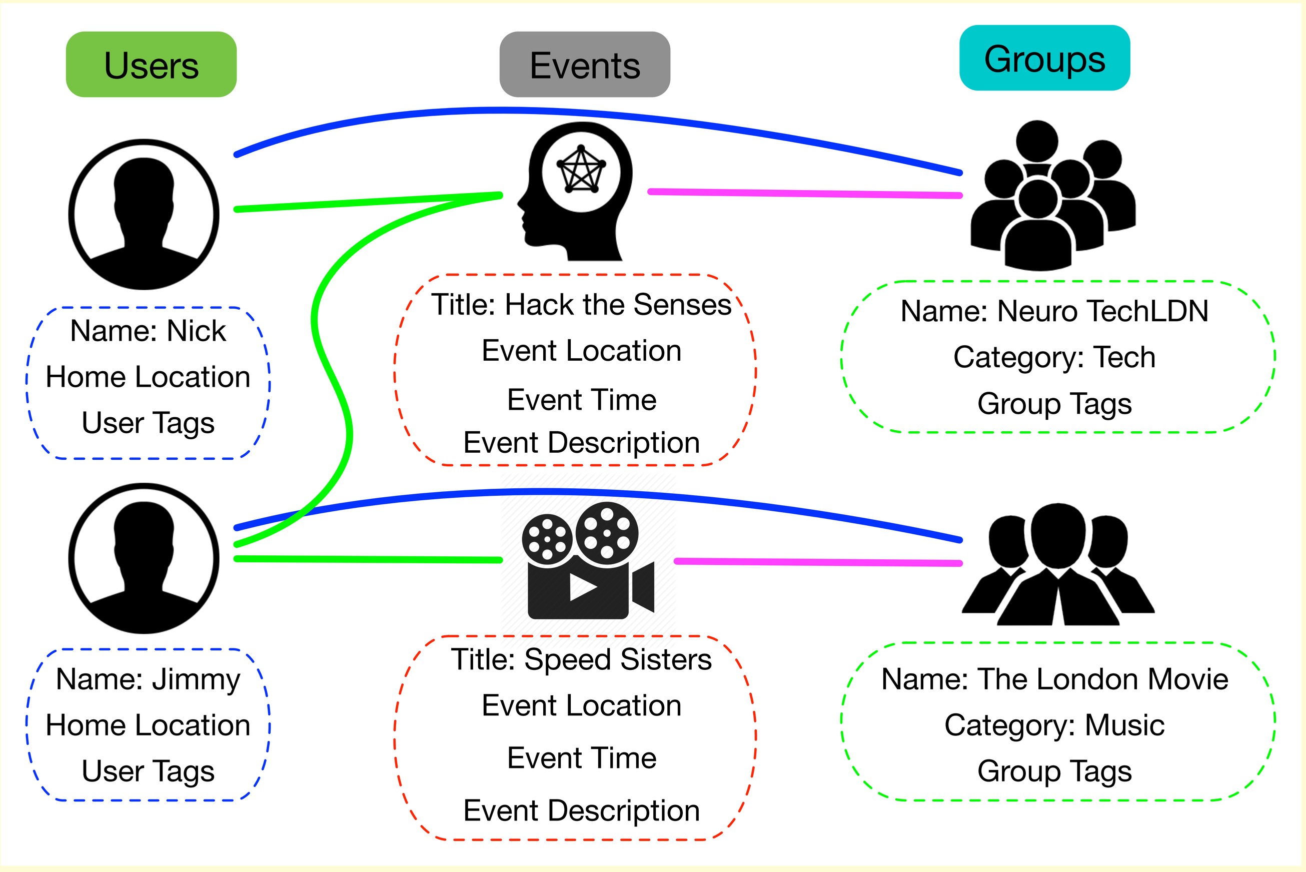
Event Organization 101: Understanding Latent Factors of Event Popularity

Shuo Zhang and Qin Lv,
Department of Computer Science, University of Colorado Boulder

1. Event-based Social Networks (EBSNs)

In the popular Meetup EBSN, the following figure illustrates three key elements. *Users* can join different Meetup *group*, which usually have specific themes such as hiking, writing or health. Each *group* can organize various types of real-world *events* and invite users to attend.



2. Meetup Datasets

We collected two years of Meetup data in three major cities, representing diverse cultures and group/event/user characteristics.

Table: Statistics of Meetup Datasets

City	#groups	#users	#events	#rsvps
New York	2,802	248,211	270,321	1,613,634
London	1,534	155,883	117,862	945,669
Sydney	706	55,768	55,295	353,149

5. Temporal Preference

Temporal preference refers to how well event start time matches group members' temporal references. We measure the overall satisfaction for event e by adding up the Jaccard similarity between event's start time and each member's temporal preference.

$$\hat{S1}_{temporal}(e) = \sum_{u \in E_u} Jaccard(\vec{e_t}, \vec{u_t})$$

$$Jaccard(\vec{e_t}, \vec{u_t}) = \frac{|\vec{e_t} \cap \vec{u_t}|}{|\vec{e_t} \cup \vec{u_t}|}$$

References

[1]. Jensen, P. 2009. Analyzing the localization of retail stores with complex systems tools. In Proc. of the 8th Intl. Symposium on Intelligent Data Analysis , 10–20.

[2]. Hutto, C. J., and Gilbert, E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In ICWSM , 216–225.

3. Spatial Features

Location Quality refers to the co-location frequency of venues in different group categories.
According to Jensen's [1] attractiveness value between different group categories, we can define the quality of location for each event e.

Table: Top 4 most (and least) attractive group categories for category "Women", "Movies/Films", and "Sports/Recreation" based on Jensen's attractiveness value (New York dataset).

Women		Movies/Film		Sports/Recreation	
Support	4.49	Literature/Write	7.84	Photography	2.93
Fashion/Beauty	3.04	Sci-fi/Fantasy	2.62	Music	1.66
Parents/Family	2.65	Tech	2.12	Food/Drink	1.62
Environment	2.48	food/drink	2.01	Paranormal	1.60
LGBT	0.53	Support	0.44	Career	0.48
Games	0.43	Paranormal	0.42	New age	0.39
Sports/Rec	0.38	Cars/Motor	0.39	Support	0.28
Cars/Motor	0.27	Fitness	0.36	Fashion/Beauty	0.16

Location Competitiveness refers to the frequency that groups with similar topics choose to meet in the same area, such events compete with each other to attract a shared pool of users.

6. Semantic Features

Sentiment Analysis: To capture the sentiment of event content, we implemented Vader [2], a lexicon and rule-based sentiment analysis tool.

Part-of-Speech Features: The POS features we used are: adjective, adposition, adverb, conjuction, determiner, noun, numeral, particle, pronoun, verb and punctuation marks.

Text Novelty: We use Jaccard similarity to Identify the novelty of event titles by comparing with previous event titles.

7. Group-based Social Influence

To utilize group-specific information in EBSNs, we propose a new social propagation network to model people's social influences on event popularity that are specific to event's group organizers.

We define user v's direct influence credit on user u as follows:

$$w_{v,u}(e) = \sum_{e'} \frac{infl(u))}{|N(u,e')|}[\delta(G(e) = G(e')) \cdot \lambda_g \cdot decay_{v,u}(e'))] + \delta(G(e) \neq G(e)) \cdot \lambda_g \cdot decay_{v,u}(e)]$$

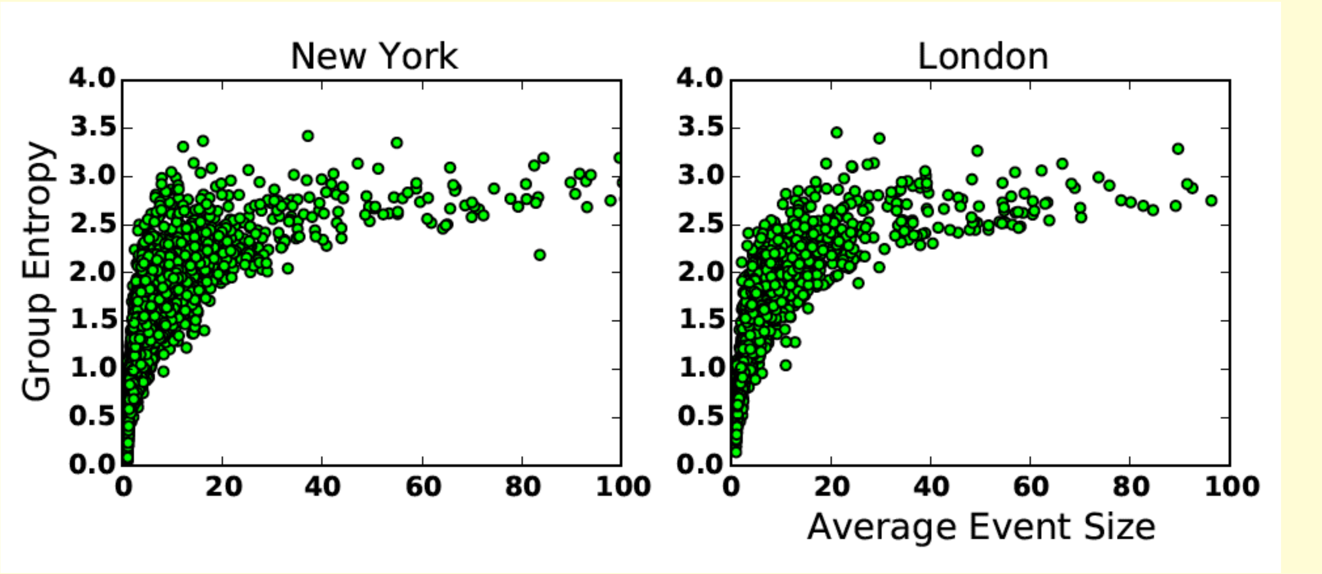
Using the social proporgation graph we can compute the total influence of user v on user u on event e:

$$\Omega_{v,u}(e) = \sum_{z \in N(u,e)} \Omega_{v,z}(e) w_{z,u}(e)$$

4. Group Features

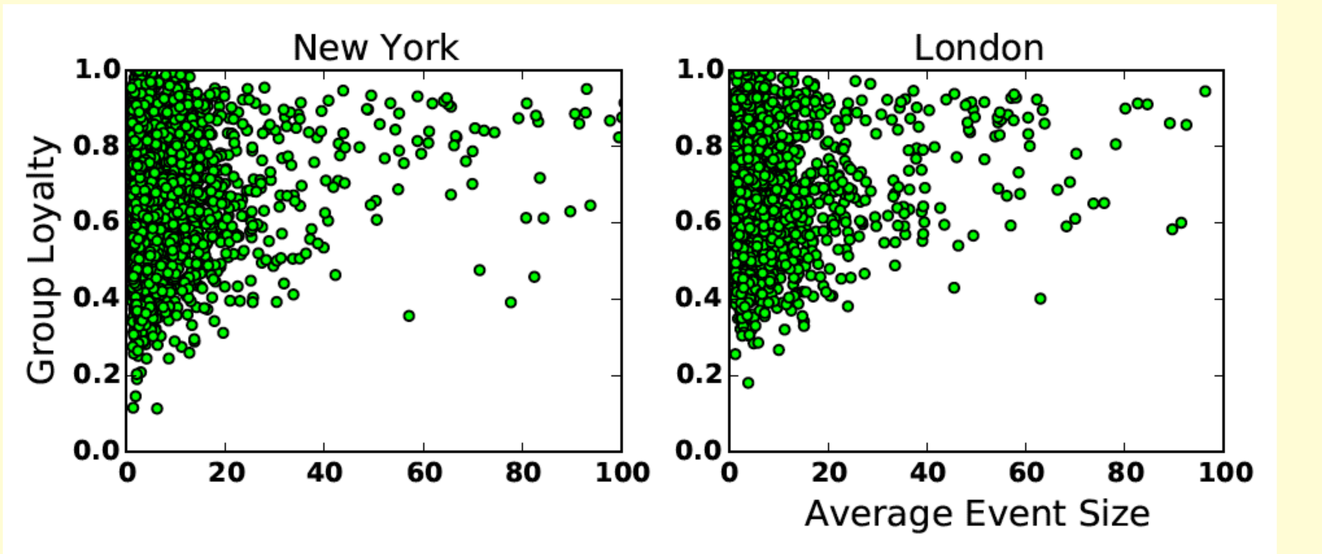
Group Member Entropy refers to the diversity of group members' interests. Given a group g, its member diversity is based on the probability of a single user u attending its offline events.

Figure: group entropy vs. group's average event size



Group Loyalty Entropy refers to what extent are the users focused on attending events within the same category. The following figure shows

Figure: group loyalty vs. group's average event size



8. Results

To integrate all context features that we have discovered, we fit them into Classification and Regression Tree model. We apply R square as the evaluation metric.

Table: Performance comparisons of different Models (R square)

	NM	SVD-MFN	Cont	CASINO(-)	CASINO
NYC	0.240	0.319	0.730	0.744	**0.758**
LON	0.140	0.305	0.672	0.692	**0.723**
SYD	0.117	0.289	0.653	0.685	**0.718**

9. Conclusions

1. Our combined CASINO framework achieves high prediction accuracy for real world Meetup event popularity.

2. Our study offers initial new insights for event organizers as well as targeted advertising strategies for EBSN service providers.

Email: {jasonzhang, qin.lv}@colorado.edu