2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Can I Foresee the Success of My Meetup Group?

Soumajit Pramanik[1], Midhun Gundapuneni[1], Sayan Pathak[2] and Bivas Mitra[1]
[1]Department of Computer Science & Engineering, IIT Kharagpur, India
[2]Microsoft Corporation, Seattle, USA
soumajit.pramanik@cse.iitkgp.ernet.in, g.midhun95@gmail.com, sayanpa@microsoft.com, bivas@cse.iitkgp.ernet.in

***Abstract*—[1]Success of Meetup groups is of utmost importance for the members who organize them. Given a wide variety of such groups, a single metric may not be indicative of success for different groups; rather, success measure should be specific to the interest of a group. In this paper, accounting for the group diversity, we systematically define Meetup group success metrics and use them to generate labels for our machine learnt models. We crawl the Meetup dataset for three US cities namely New York, Chicago and San Francisco over a period of 8 months. The data study reveals the key players (such as core members, new members etc.) behind the success of the Meetup groups. This study leverages semantic, syntactic, temporal and location based features to discriminate between successful and unsuccessful groups. Finally, we present a model to predict success of the Meetup groups with high accuracy (0.81 with AUC = 0.86). Our approach generalizes well across groups, categories and cities. Additionally, the model performs reasonably well for new groups with little history (cold start problem), exhibiting high accuracy for the cross city validation.**

I. Introduction

Over the recent years, social networks have provided convenient online platforms for people to create, and organize social events. Meetup is such a popular event based social networking (EBSN) portal that facilitates hosting events in various localities around the world [1]. The platform has experienced a rapid growth in its population during recent times and currently, it has 25.58 million users spread over 180 countries, creating 580,960 social events every month[2].

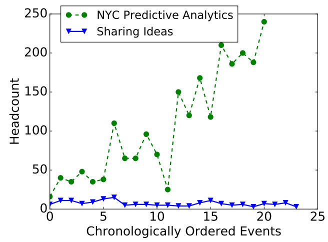
The popularity of EBSN leads to the problem of 'information overload'. Choosing suitable events to attend and selecting proper group to join require a lot of user deliberation. In order to mitigate this effort, different recommendation systems have been developed. The prior work stressed on the following two different recommendation systems. (a) Event recommendation - recommends suitable events to a single or a set of Meetup-users based on user's past preferences and current context [2], [1], [3], [4], [5], [6] etc. (b) Group recommendation - recommend groups to a newly joining member, considering both implicit and explicit factors that could influence users' decisions; this includes factors such as user's profile information, location and social features [7], [8]. However, most of these systems caters the need of the general Meetup event attendees and group members. Importantly, only a few work provide proper guidance to the group organizers and event hosts (jointly we refer as '*Meetup authorities*') in order to form a successful group. In [9], [10] She et al. proposed heuristics to resolve conflicts of different events, reducing the redundant arrangement and making event-participant arrangements in a global view. Nevertheless, as group organizers and event hosts put a lot of effort to make a group successful, suitable framework needs to be developed to cater their requirement.

Studies show that all the Meetup groups do not survive over a prolonged period of time [11]. Survival of a Meetup group is directly connected to its capability of attracting population. For example, Fig. 1(a) shows the attendance in the events organized by two Meetup groups in New York namely "Sharing Ideas" & "NYC Predictive Analytics", both created in July 2009, focusing on the 'Technology'. It is important to notice that, despite the similarity in the formation time, location and the interest, both the groups exhibit markedly different behavior in terms of attracting the people. One of the major objective of the of the *Meetup authorities* is to make their own groups successful. This immediately raises the question- 'Can we develop a framework which can predict the success of a Meetup group?'. Defining Meetup group success is open ended and ambiguous due to the wide variety of Meetup groups. Being an EBSN platform, organizing popular events, attracting many attendees, can work as a success measure for a set of Meetup groups. Nevertheless, for certain well established groups, maintaining a reasonably steady size could be attributed to the group success. The intrinsic objectives of these groups can be significantly different. For example, technical groups may prefer to organize small but productive events whereas travel groups may prefer to organize large scale events. Consequently, one single yardstick of success may not capture the objective of these two groups.

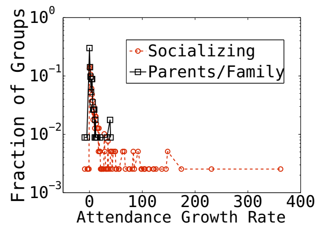
In this paper, we present a systematic approach to predict the success of a Meetup group. First we dissect the Meetup dataset, collected in three US cities, to identify different factors that make one group successful (section II). We discover that most of the groups exhibit a core-periphery structure where core members exhibit a strong topical alignment with the group. We observe that these core users have a strong impact on the event attendance; for instance event attendance

[1]An initial version of this work has been published as a poster in ICWSM 2016
[2]http://www.meetup.com/about/

IEEE/ACM ASONAM 2016, August 18-21, 2016, San Francisco, CA, USA
978-1-5090-2846-7/16/$31.00 © 2016 IEEE



(a) Headcount (event-attendance) for two groups from same category created during same month in 2009



(b) Distribution of Attendance growth rate for groups belonging to 'Socializing' and 'Parents/Family' categories of New York

Fig. 1.

City	Fast Crawl		Detail Crawl				
	Groups	Members	Groups	Events	Venues	RSVPs	Member Profiles
Chicago	5727	342773	5671	31719	435553	2749595	249652
New York	17180	1026901	17094	81786	1150394	7910224	814378
San Francisco	13381	753839	13297	65266	833045	6542690	606097

TABLE I
DATA COLLECTED BY BOTH CRAWLERS FOR ALL 3 CITIES

improves if core members stay at close proximity of each other. We also observe that a flush of newly joining users to a group, having strong inclination towards the group and to the event topic, results in high event attendance.

Next we propose a principled approach to measure the success of the Meetup groups. We identify a set of candidates metrics which may work as group success measure, such as (a) average event attendance organized by a group (b) growth rate in the attendance over a period of time (c) average size of the group (d) growth rate of the group. However, the success of a specific group gets determined by the objective of the participating members, attendees and the *Meetup authorities* since a single measure may not be able to capture the success motive of all these diverse category of groups. We classify the different Meetup groups into five categories and identify one success metric for each category (section III). Finally, we present a model for Meetup group success prediction. We perform experiments on the Meetup groups of three US cities namely New York, Chicago, San Francisco. We present different variations of the model; city specific, category specific and combined model. We also perform a cross city validation of the model to demonstrate its utility. The proposed model on average exhibits a high prediction accuracy of 0.81 (AUC = 0.86) (section V). Feature analysis shows that semantic feature plays an important role in the prediction model. The proposed framework produces decent results for all the classification & regression models; however Decision Tree and Linear Regression perform little better.

II. DATASET

A. Data Collection

We crawl the Meetup EBSN data using two crawlers for New York, Chicago and San Francisco during a period of 8 months (from August 2015 to March 2016). The two different crawlers are called Fast Crawler and Detail Crawler. These two crawlers gather the different Meetup network attributes detailed below.

(a) **Fast Crawl:** This is a fast crawler (cycle duration of 3 days) which collects only the members of all groups (no detail information) in each city and generates a member-to-group mapping along with timestamps. It does not crawl any event or venue related information. This crawler is designed to collect the member dynamics across the Meetup groups.

(b) **Detail Crawl:** This is a slow but detailed crawler (cycle duration $7-10$ days) which collects the event details of all the groups. Data crawled by this crawler includes group details including member profile, events hosted by them, event RSVPs, event venues etc.

Meetup group dynamics presents unique challenges to the data collection process, which leads us to develop the two aforementioned crawlers. Given our goal is to measure the temporal network evolution, we need to sample the individual groups' data with a reasonable periodicity to capture their temporal dynamics. While group memberships change frequently (less than a weekly granularity), the events hosted by different groups show less temporal variations with the details of the events being valid for a longer duration. While a single crawler is easier to design and maintain, to account for the aforementioned variations in data changes in Meetup group, we have developed two different crawlers (one faster than the other) that perform adequately to respect the Meetup crawling API constraints.

Table. I shows the statistics of the number of users, events and groups we crawled. In the following, we introduce the different actors and entities connected to the EBSN dataset.

B. Dataset: Major Components

1) Member and group profile: The profile of one member or a group gets specified by the set of Tags, which reflects their respective preferences. Whenever one member joins Meetup, she is asked to select some tags for describing her interests. Similarly, when a Meetup group gets formed by the group organizer, she is asked to select a set of tags which describes the group best.

2) Event attendance & attendees: In Meetup, for each event, there exists a field called "Headcount" which provides the actual attendance information of an event. However, this count does not provide the details of the individual attendees. On the other hand, details of the individual attendees can be obtained from the RSVP message {"Yes", "No", "Maybe"}. Event attendees for an event e_i are the participants who send "Yes" response to RSVPs corresponding to that event.

3) Group category: Category indicates the interest of a Meetup group. During formation, each group is assigned to one of the 33 'official' categories defined in Meetup. For examples, few popular Meetup categories are 'Career/Business', 'Tech', 'Health/Wellbeing', 'Socializing' etc.

III. Measuring Success of a Meetup group

In this section, we define the success metric of the Meetup groups. Given a wide varieties of Meetup groups, one universal metric may not be able to indicate the success of all the groups; success measure should be specific to the interest of a group. If we observe the Meetup groups closely, we discover different signatures of success (a) organizing a massive event attracting many attendees (b) large group size (c) steady growth in the group size & event attendance etc. In the following, we define a set of potential metrics to realize the success of different Meetup groups. Next, we judiciously form the metrics to feature success depending on the specific characteristics of the groups.

A. Candidate Metrics

In this paper, we mostly focus on the popularity centric metrics to feature group success; nevertheless other aspects, such as post event sentiment etc. can also be explored. Popularity of a group can be broadly measured from two perspectives - (a) size of the group - if it is able to attract new members to join the group (b) event attendance - if it is able to attract users to attend the events hosted by the group. For a group g organizing events $e_1, e_2, \ldots, e_k$ at times $t_1, t_2, \ldots, t_k$, the candidate metrics can be mathematically defined as,

(a) **Average group size** at t_k, $G_k = \frac{\sum_{i=1}^{k} |M_g^{t_i}|}{k}$ where $M_g^{t_i}$ is the set of group members of group g at time t_i.

(b) **Average event attendance** at t_k, $E_k = \frac{\sum_{i=1}^{k} H_{e_i}}{k}$ where H_{e_i} is the 'Headcount' of event e_i.

However, the aforesaid metrics fail to appreciate the newly created (small sized) groups, having potential to gain popularity in future. Hence, additionally we introduce the metrics which factor rate at which the group size and event attendance grow over time.

(c) **Event attendance growth rate** at t_k,

$$E_g = \frac{\sum_{i=2}^{k} \frac{H_{e_i} - H_{e_{i-1}}}{H_{e_{i-1}}}}{k-1} \tag{1}$$

(d) **Group size growth rate** at t_k,

$$G_g = \frac{\sum_{i=2}^{k} \frac{|M_g^{t_i}| - |M_g^{t_{i-1}}|}{|M_g^{t_{i-1}}|}}{k-1} \tag{2}$$

In summary, we use a suite of 4 candidate metrics $\langle$ G_k, E_k, E_g and G_g $\rangle$ to quantify the success of a group where each of the metric can be computed based on past k events organized by the group.

B. Category Specific Success Metrics

1) Key Idea: We aim to assign one (or more) success metrics for each Meetup category. The key observation is that for each Meetup category, the distribution of the candidate metrics for all the groups widely varies; very concentrated for few metrics whereas widely spread for others. We check this hypothesis in Fig. 1(b) for two categories 'Socializing' and 'Parents/Family' considering 'Event attendance growth rate' E_g as a candidate metric. This becomes clearly evident that E_g exhibits well distributed behavior for the groups in 'Socializing' category discriminating successful and unsuccessful groups. On the other hand, high concentration of E_g for the 'Parents/Family' category makes it unable to mark any distinction between the groups. Hence E_g may be considered as a suitable success metric for 'Socializing' category.

Category Group	Meetup official categories
Activity	dancing, fitness, sports/recreation, health/wellbeing, games etc.
Hobby	fine arts/culture, fashion/beauty, hobbies/crafts etc.
Social	movements/politics, socializing , singles, parents/family etc.
Entertainment	food/drink, movies/film , music, sci-fi/fantasy etc.
Technical	career/business, tech, education/learning etc.

TABLE II
MEETUP CATEGORIES DIVIDED INTO GROUPS

2) Methodology: The objective is to identify the most discriminating success metrics for each Meetup category. The following steps are followed.

Preprocessing: In order to address the data sparsity issue, we classify the official Meetup categories into the following five classes - (a) Activity (b) Hobby (c) Social (d) Entertainment and (e) Technical (see Table. II). We perform a small in house survey with 20 participants and 17 of them completely agreed with this grouping whereas three of them suggested overall 5% of change. Henceforth, we use the aforementioned five categories even though our methodology can be extended to any number of categories provided adequate crawled data is available.

Metric selection: We use 'Entropy' to characterize the discriminative property of each candidate metric M for Meetup groups in category C. One pleasing property of this measure is that it does not make any distributional assumptions. We propose 'Binned Entropy' where we apply standard entropy metric after optimally binning the data points. Here the data points of size N_M within the range $[L_M, H_M]$ represents the quality of the Meetup groups in category C following the metric M. Once we fix the correct bin size, we split this range $[L_M, H_M]$ into a number of bins (with a possibility of empty bins). Finally we compute the fraction of data points p_i in bin i and calculate the entropy of the distribution p_i of metric M for category C. Higher entropy indicates better suitability of metric M for the Meetup groups in category C.

(a) Computing Bin Size: The first challenge is to determine the proper bin size. Bin size regulates the two competing factors (i) accumulated error for binning (ii) number of bins. The objective is to fix a bin size which optimizes both of these factors.

We vary the bin size from 1 to $H_M - L_M + 1$. For each bin size, we calculate an error as follows, which estimates the goodness of the current binning. For bin size S, if the i^{th} bin has N_i data points - $i_1, i_2, \ldots i_{N_i}$ with mean μ_i, then the overall error can be calculated as $E_S = \frac{1}{N_M}\sum_{i=1}^{B_S}\sum_{j=1}^{N_i}(i_j - \mu_i)^2$ where B_S is the total number of bins of size S. On the other side, we want to keep the number of bins B_S small (significantly lower than the number of data points), which may increase the error. In order to find the optimal bin size, we propose a penalty function for bin size S which is calculated as $P_S = \frac{E_S}{\max_{\forall S}(E_S)} + \frac{B_S}{\max_{\forall S}(B_S)}$ We vary the bin size and using 5 fold cross validation we obtain the optimal bin size for which the penalty is minimum.

(b) Entropy Calculation: Once we fix the bin size, we calculate the fraction of data points p_i in bin i for the range $[L_M, H_M]$ & compute the overall entropy of p_i distribution. For example, we measure the entropy of metric M for Meetup groups in category C as $\sum_{i=1}^{B_O} -p_i \log p_i$ where B_O is the number of bins. We use a 66.67^{th} percentile (highest one-third) as threshold on this entropy values and select the metrics crossing this threshold as the group success metric of that category. In Table. III, we present the entropy of each candidate metric (selected metrics are marked in bold) for every category in Chicago, New York and San Francisco. In case, none of the metrics are above the threshold, we choose the one with maximum entropy. We observe that, irrespective of cities, either one of E_g & G_k or both got selected as metrics for all category of groups. Even the chosen metrics remain exactly same for three of the categories across the cities, showing robustness of our metric selection procedure.

3) Labeling groups: Once the success metrics are chosen for each category, we label the corresponding Meetup groups as 'successful' or 'unsuccessful'. For the categories with multiple success metrics, we use a 'Veto' strategy. We label a group as successful if it has more than 66.67^{th} percentile value for at least one of the chosen metrics. On the other hand, if no metric labels in one group is 'successful' and additionally if it has less than 33.33^{th} percentile value for at least one of the chosen metrics, then that group is labeled as 'unsuccessful'. The number of 'successful' and 'unsuccessful' groups labeled for each category is shown in Table. III.

City	Category	E_k	E_g	G_k	G_g	Successful / Unsuccessful
	Activity	0.32	**2.60**	2.42	1.53	24 / 25
	Hobby	0.08	**2.29**	1.93	1.92	20 / 20
CH	Social	0.66	**2.84**	**2.71**	1.82	207 / 187
	Entertainment	0.36	2.43	**2.65**	1.69	61 / 59
	Technical	1.00	**2.63**	2.32	2.18	31 / 30
	Activity	0.28	**2.54**	2.39	1.45	50 / 48
	Hobby	0.46	**2.64**	2.25	1.87	59 / 60
NY	Social	1.34	**3.00**	**3.00**	1.68	629 / 588
	Entertainment	1.04	**2.73**	**2.83**	1.91	198 / 185
	Technical	1.47	**2.88**	**2.98**	2.39	539 / 492
	Activity	0.58	**2.85**	2.30	1.68	58 / 57
	Hobby	0.30	**2.76**	1.93	1.89	44 / 45
SF	Social	1.02	**3.09**	**2.59**	1.88	408 / 378
	Entertainment	0.41	2.49	**2.58**	1.91	124 / 128
	Technical	1.57	**3.12**	**3.15**	2.15	592 / 551

TABLE III
ENTROPY VALUES FOR DIFFERENT SUCCESS METRICS FOR DIFFERENT CATEGORY OF GROUPS IN CHICAGO (CH), NEW YORK (NY) AND SAN FRANCISCO (SF) (CONSIDERING ONLY GROUPS ORGANIZING MORE THAN 10 EVENTS). SELECTED METRICS ARE MARKED IN BOLD.

IV. KEY PLAYERS REGULATING SUCCESS OF A MEETUP GROUP

In this section, we dissect the dataset and highlight the factors that contribute to the success of a Meetup group. In the previous section, we show that two major dimensions of measuring the success of a Meetup group are - (a) Group size and (b) Event attendance. The investigation reveals that two kinds of group members (a) New members and (b) Core members play important roles in regulating Meetup group success. We start with defining these two kinds of members.

A. New members:

In Meetup, people search for events and if they intend to participate in one event, they need to join the organizing group first. For a group g organizing events $e_1, e_2, \ldots, e_k$ at times $t_1, t_2, \ldots, t_k$, we define the new members at time t_i as a set of users who join the organizing group g just before the event e_i (i.e. in between t_i and t_{i-1}). This has been observed that a significant fraction of event attendees join the group just before the event.

B. Core members:

Informally, the core members of a group are the dedicated set of members who have a strong interest overlap with the group. In order to identify the core, we propose a similarity metric between tags of the members and the groups.

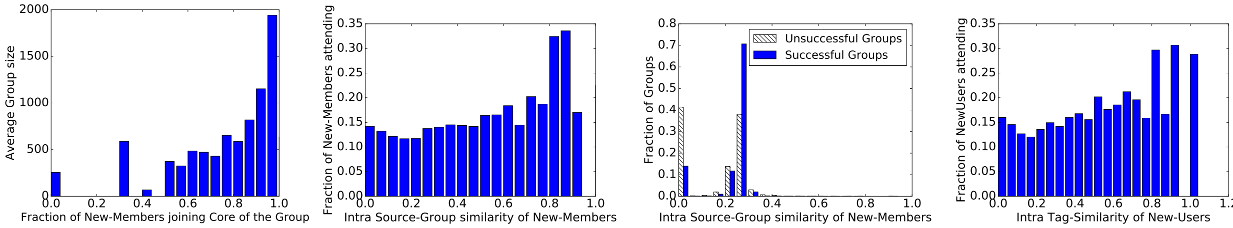
Tag similarity ($TagSim$)**:** We represent tags of an individual member/group as a tag vector TV of length N_T where N_T is the number of all possible tags. The coefficient of each component (tag) of this vector is the normalized usage frequency of the corresponding tag. The similarity between two tag vectors $\mathbf{TV_i}$ and $\mathbf{TV_j}$ is calculated as the cosine similarity of these two vectors i.e. $TagSim(\mathbf{TV_i}, \mathbf{TV_j}) = \frac{\mathbf{TV_i}.\mathbf{TV_j}}{\|\mathbf{TV_i}\|\|\mathbf{TV_j}\|}$

We define core members as the Meetup group members exhibiting a threshold 66.67^{th} percentile (0.2 in our case) of tag similarity between groups and their corresponding members. So, for a group g with tag-vector $\mathbf{TV_g}$ at time t_i, we denote the set of core-users as $C_i = \{u : u \in M_g^{t_i}$ & $TagSim(\mathbf{TV_u}, \mathbf{TV_g}) > 0.2\}$.

In the following, we demonstrate the importance of new members and core members on the group's success.

1) How do the new members affect group success?: If newly joined members are highly aligned to the profile of the group (i.e. if they join the core of the group), it facilitates the group to grow. In Fig. 2(a), we show that joining of new members at the core of the group helps to increase the group size (with correlation coefficient 0.62) In fact, it is found that for 'successful' groups the average fraction of new members joining the core is 35% more than that of the 'unsuccessful' groups.

There is another aspect of the new joining members. In general, the new members are also member of different other groups, which we call as 'source groups' of the new member. The impact of a new member on the success of a group she is joining, also depends on her source group(s). Intuitively, if



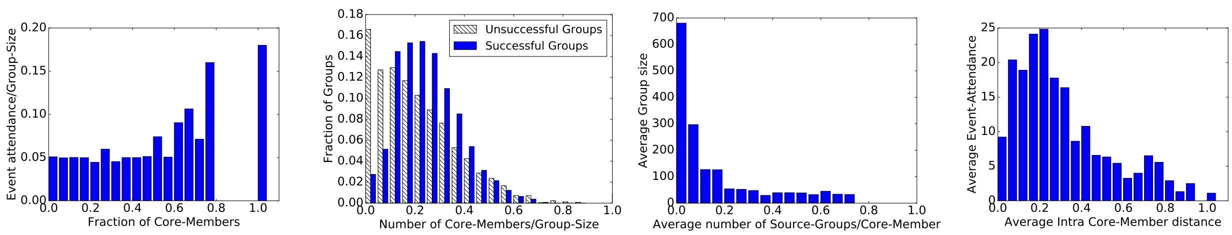
(a) Fraction of new members joining the core of organizing group vs. group size

(b) Intra source group tag similarity of new members vs. fractional attendance of new members

(c) Distribution of intra source group similarity of new members for successful and unsuccessful groups

(d) Intra tag similarity of new members vs. fractional attendance of new members

Fig. 2. Impact of new members on group success in Chicago



(a) Fraction of core members vs. fractional attendance of group members

(b) Distribution of fraction of core members for successful & unsuccessful groups

(c) Average number of source groups of core members vs. group size

(d) Intra core members distance vs. fractional attendance of group members

Fig. 3. Impact of core members on group success in San Francisco

new members join from similar source groups, this increases the group cohesivity and as a result, they start attending events together. Fig. 2(b) depicts that the fraction of new members attending an event highly correlates (with correlation coefficient 0.71) with the intra source group similarity among themselves. Moreover, the distributional disparity between 'successful' & 'unsuccessful' groups in Fig. 2(c) emphasizes that the intra source group similarity of new members plays an important role in making the groups 'successful'. For the same reason, in case of attending events, we also find high correlation (with correlation coefficient 0.77) between tag similarity among the new users and their probability of attending event (see Fig. 2(d)).

2) How do the core members influence group success?: The fraction of users belonging to core of a group directly correlates (with correlation coefficient 0.80) with the event attendee fraction of the group (see Fig. 3(a)). Furthermore, in Fig. 3(b), we observe that 'successful' groups have significantly large fraction of members in their core compared to 'unsuccessful' groups.

We observe an interesting property regarding the source group(s) of the core members. We find an anti-correlation (with correlation coefficient -0.60) between the average number of source groups of core members and the group size (see Fig. 3(c)). Meetup group with core members having less source groups makes the group more focussed and involved. This eventually inspires other users to join the group and increases the group size.

Understandably, core users are also found to be event hosts for more than 51% of events. Intuitively, it should always help them to organize events & attend events together (which directly influence event attendance), if they are closely located. Fig. 3(d), completely supports this hypothesis, depicting high anticorrelation (with correlation coefficient -0.83) between average pairwise distance among core members and event popularity.

V. Group Success Prediction

In this section, we describe the features used in our proposed machine learnt models to predict success of a group. We have trained several classical machine-learning models using those features and evaluated their ability to predict group success.

A. Experimental Setup

1) Labeling Groups: First we prepare the 'ground truth' labeling of the Meetup group as 'successful' and 'unsuccessful' using the chosen metric corresponding to its category (as described in Section. III). We use the statistics of the past k events of a group for labeling. For instance, if we label a group at the i^{th} event, we use the statistics of the sequence of events starting from e_{i-k+1} to e_i. We have attempted different k = 5, 10, 15 and 20 (results not shown here). However, we notice that $k > 10$ significantly reduces the number of groups for evaluation, as Meetup groups do not sustain for a long time. On the other hand, very low values of $k < 5$ doesn't provide sufficient insights into the track record of a group since the initial events are often exploratory. We found $k = 10$ provides reasonable trade off between the number of groups with labeled data and sufficient history for track record. We deem any group which has less than 10 events to be failure if they have stopped having any event in the recent past 2 out of the 3 months since our crawler started.

2) Extracting Features: The model features chosen can be broadly divided into four classes -

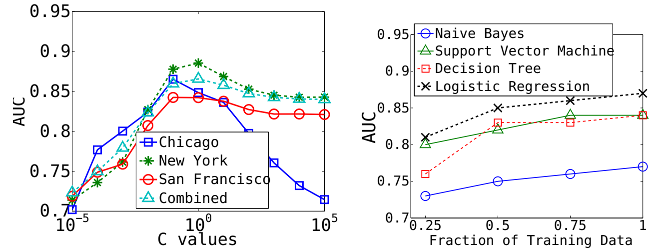
- **Semantic or tag related features** - Average tag vector similarity between the organizing group & the group members, average intra member tag vector similarity, average tag vector similarity between source groups of members and the organizing group, average intra source group tag vector similarity etc.
- **Syntactic or count based features** - Average pairwise count of common past events between group members, fraction of group members sending 'Yes' RSVP, fraction of group members joining the core of the group, average pairwise overlap of source groups of group members, average number of source groups per group member etc.
- **Time related features** - Day of week on which the event occurred, duration of the event, time zone of the event, time gap between announcement and occurrence of the event.
- **Location related features** - Average pairwise distance between venues of group members, average distance between the venue of the event& the group member etc.

The features are deliberately chosen in a way such that the metrics used in labeling each group get excluded from being features in our ML models. We calculate these features separately for three most important types of group members - event attendees, core members and new members.

In order to calculate these features, we use the information of the first $k/2$ (say 5) past events. We denote this sequence of events as the 'feature window'. For instance, to predict the group label at the event e_i, we use features of the sequence of events starting from the event e_{i-9} to event e_{i-5} event. This prediction enables the group organizes to foresee the performance of the group much early, from the feature window.

3) Prediction Models: We develop three different versions of the prediction model - (a) City specific model where we consider groups in a specific city (b) Category specific model where we consider groups of individual categories and (c) 'Combined' universal model considering all the groups of all the different cities. The 'Combined' model becomes specially useful to mitigate the cold start problem for a new group. We demonstrate the prediction results using four standard models - Naive Bayes, Support Vector Machine (SVM), Decision Tree & Logistic Regression. We use these models as classifiers as well as regressors. The prediction results are evaluated based on the classification accuracy and the area under the Precision-Recall curve (AUC) metrics.

Setting Model Parameters: Different model parameters such as maximum depth for Decision Tree (DT), regularizer weightage for Logistic Regression (LR), Kernel type for SVM etc are determined using 10-fold cross-validation. We use 80% of the dataset for training and remaining 20% as the hold-out set (data set not used in training). Fig. 4(a) shows the change in AUC with varying the regularization parameter in LR. For choice of parameters C between 0.1 and 1.0 we find there to be an optimal trade off between model complexity and classification error (bias - variance trade-off). Similarly a depth of 4 in decision tree classifier ensures the appropriate bias-variance trade-off.



(a) AUC for Logistic Regression with different C values (C = inverse of weightage of regularizer)

(b) AUC for Logistic Regression with different amount of training data for the 'combined' model.

Fig. 4. Avoiding overfitting for Logistic Regression and showing impact of amount of training data present for different models

City	Naive Bayes		SVM		Decision Tree		Logistic Regression	
	ACC.	AUC	ACC.	AUC	ACC.	AUC	ACC.	AUC
Chicago	0.77	0.77	0.74	0.82	**0.82**	**0.86**	0.79	**0.86**
New York	0.76	0.78	0.76	0.84	**0.82**	0.82	0.80	**0.89**
San Francisco	0.73	0.77	0.73	0.80	**0.79**	**0.84**	0.77	**0.84**
Combined	0.75	0.77	0.75	0.84	**0.80**	0.84	0.78	**0.87**

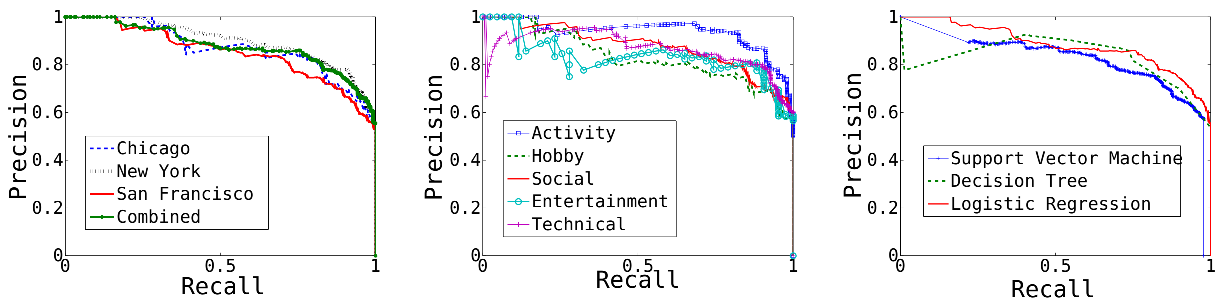
TABLE IV
CLASSIFICATION ACCURACY VALUES (ACC.) AND REGRESSION AUC VALUES FOR ALL 3 CITIES

B. Evaluation Results

1) City specific models: The classification accuracy and AUC values for city specific models & the 'Combined' model are available in Table. IV. On average, we get around 0.81 accuracy and around 0.86 AUC for city specific models. Decision Tree produces the best classification results and Logistic Regression gives the best AUC values. In Fig. 5(a), we plot the Precision-Recall curves for city specific models using Logistic Regression. 'New York' city exhibits the best performance. Interestingly, if we look at our dataset (Table. I) carefully, we observe that for 'New York' city the crawled data size is maximum. This leads to an obvious question - *Do the performances of different models depend on the size of training data?*

To address this question, we train the prediction models with 25%, 50% and 75% of the training dataset. Fig. 4(b) demonstrates the performance improvement of all the classification models with training data size; the Decision Tree model enjoys the maximum performance gain.

2) Category specific models: The classification accuracy and AUC values for category specific models are available in Table. V. Here, on average we get around 0.82 accuracy and 0.87 AUC. Especially, for the 'Activity' category, all the models perform exceptionally well. Like city specific models, here



(a) Logistic regression performance across different cities.

(b) Logistic regression performance across different categories.

(c) Different ML techniques' performance using 'combined' model.

Fig. 5. Precision-Recall curves across different cities and categories

Category	Naive Bayes		SVM		Decision Tree		Logistic Regression	
	ACC.	AUC	ACC.	AUC	ACC.	AUC	ACC.	AUC
Activity	0.87	0.88	0.82	0.91	**0.93**	**0.94**	0.89	0.93
Hobby	0.75	0.73	0.73	0.76	**0.81**	**0.87**	0.77	0.81
Social	0.74	0.72	0.73	0.81	**0.80**	0.83	0.76	**0.87**
Entertainment	0.76	0.75	0.70	0.78	**0.79**	0.70	0.76	**0.83**
Technical	0.72	0.77	0.71	0.83	**0.77**	0.82	0.76	**0.86**

TABLE V
CLASSIFICATION ACCURACY VALUES (ACC.) AND REGRESSION AUC VALUES FOR ALL 5 CATEGORIES

City	Only 1^{st}	Upto 3^{rd}	Upto 5^{th}	Upto 7^{th}
Chicago	0.835	0.856	0.859	0.873
New York	0.852	0.879	0.889	0.891
San Francisco	0.800	0.837	0.842	0.852
Combined	0.823	0.858	0.866	0.876

TABLE VI
AUC VALUES FOR ALL 3 CITIES USING LOGISTIC REGRESSION WITH VARIOUS SIZES OF THE FEATURE WINDOW

Decision Tree model produces the best classification results and Logistic Regression gives the best AUC. In Fig. 5(b), we plot the Precision-Recall curves for different category specific models using Logistic Regression.

In Fig. 5(c), we focus on the 'Combined' model and show the Precision-Recall curves for different classifiers. Here Logistic Regression performs specially well at the very low and very high recall regions whereas Decision Tree performs better in the middle regions. Overall, we get around 0.81 classification accuracy and 0.86 AUC values across all city specific and category specific models. We observe Logistic Regression and Decision Tree to be the most suitable regression & classification models.

In the following, we pose various research questions regarding the proposed success prediction models.

3) What is the importance of feature window?: In all the aforesaid experiments, we consider a fixed feature window of five past events starting from the event e_{i-9} to e_{i-5}, to predict the group success at the event e_i. In this subsection, we show the impact of the feature window on the model performance. In Table. VI, we show the AUC values for different sizes of the feature window, all starting at the event e_{i-9}. We observe a AUC improvement of 4%-5% if the feature window size increases from 1 to 7. Interestingly, we note that even with just the first event e_{i-9} information, we can predict the group success for the event e_i with more than 0.80 AUC for all the cities. This further confirms the robustness of our model against the cold start problem.

4) How generic are our models?: In order to show the generality of the model, we perform a cross city validation in two different ways. (a) Firstly, we train our models with two of the three cities - 'Chicago', 'New York' & 'San Francisco' and test on the third city. This gives an accuracy of 0.89 for 'Chicago', 0.89 for 'New York' and 0.83 for 'San Francisco' with Logistic Regression. This points to the fact that Meetup groups behave uniformly across the cities. This is important to note that the AUC value of 'Chicago' in this case is even higher than its corresponding city specific model. This happens primarily because of the lack of enough labeled groups in 'Chicago' (only 1575 'Chicago' groups got selected for labeling whereas for each of 'New York' and 'San Francisco, there are more than 4000 labeled groups). (b) Secondly, we crawl the Meetup data for a brand new city 'Las Vegas' for the past two months (February-March 2016) and tested it using the 'combined model' of 'Chicago', 'New York' & 'San Francisco'. This gives 0.85 AUC value with Logistic Regression.

5) Feature importance: Most of the Machine-Learning models (say Decision-Tree, Logistic Regression & Support Vector Machine) assign the relative weightages to individual features. This helps us in understanding the key features influencing success of a Meetup group. The top and bottom set of features selected by these three models do not vary significantly. In Table. VII, we consolidate the top and bottom set of features according to the weights (absolute values) assigned by Logistic Regression for different city specific models. It is clearly evident that the semantic features heavily dominate the set of top features. Tag based features like 'average intra source group similarity' or 'average source group organizing group similarity' and count-based features like 'average source group overlap' are important for almost all different city specific models. Additionally, we perform the statistical significance test of the individual features to

	Chicago		New York		San Francisco	
Top Features	W_{LR}	p	W_{LR}	p	W_{LR}	p
Avg. source group overlap	1.00	0.00	1.47	0.00	1.24	0.00
Avg. source group organizing group similarity	1.33	0.00	1.24	0.00	1.28	0.00
Avg. intra source group similarity	1.33	0.00	1.02	0.00	1.28	0.00
Avg. intra user distance	0.68	0.00	1.55	0.00	1.03	0.00
Timegap between announcement & event	-0.92	0.04	-0.99	0.00	-1.93	0.00
Bottom Features	W_{LR}	p	W_{LR}	p	W_{LR}	p
Fraction of new members sending 'Yes' RSVP	-0.62	0.27	-0.86	0.00	-0.33	0.59
Number of source groups per user	-0.40	0.01	-0.27	0.00	-0.71	0.01
Time-zone of the event	-0.29	0.08	-0.01	0.66	-0.23	0.98
Avg. user venue distance	0.09	0.26	0.05	0.02	-0.09	0.05
Avg. pairwise count of common past events	0.00	0.33	-0.12	0.10	-0.23	0.12

TABLE VII

LOGISTIC REGRESSION WEIGHTS(W_{LR}) AND P-VALUES FOR TOP & BOTTOM FEATURES CORRESPONDING TO ALL THREE CITY SPECIFIC MODELS

distinguish between the 'successful' and 'unsuccessful' groups and show the corresponding p-values in Table. VII[3]. We observe that most of the features which are assigned high weights by the models, also obtain low p-values (and vice versa). Exceptions also exist. For example, 'number of source groups per user' is a feature which has low p-value; however, it exibits low weightages in Logistic Regression.

VI. CONCLUSION

In this paper, we have developed a framework to predict the success of Meetup groups while considering the diverse objectives of different '*Meetup authorities*'. We have proposed a principled approach to fix the success yardstick of a Meetup group. This approach has proved to be robust enough to detect category specific success metrics which remain valid across different cities. It also allows us to generate pseudo-ground truth from the data that is subsequently used for an inexpensive and scalable way to generate reliable label data. We have presented several machine learnt models to predict group success by leveraging semantic, syntactic features, temporal as well as location information. The most efficient model achieves an average accuracy of 0.81 with 0.86 AUC. Moreover, we present the performance of individual city specific, category specific as well as a 'combined' model. Our 'combined' model has a low generalization error and more than 0.83 AUC even for groups belonging to some 'unknown' city. For most of the cases, the semantic and syntactic features of the core & new members are found to be performing better than distance related and time related features. Overall, we observe that Decision Tree & Logistic Regression with L2 regularization appears as the most suitable models for our experiments.

VII. ACKNOWLEDGEMENTS

This work has been partially supported by the SAP Labs India Doctoral Fellowship program, Microsoft Corporation and DST - CNRS funded Indo - French collaborative project titled "Evolving Communities and Information Spreading".

[3]The p values & weights assigned by Logistic Regression for all the features are available in http://bit.ly/1Wq2FRb

REFERENCES

[1] A. Q. Macedo, L. B. Marinho, and R. L. Santos, "Context-aware event recommendation in event-based social networks," in *Proceedings of the 9th ACM Conference on Recommender Systems*, ser. RecSys '15. New York, NY, USA: ACM, 2015, pp. 123–130. [Online]. Available: http://doi.acm.org/10.1145/2792838.2800187

[2] C. Luo, W. Pang, Z. Wang, and C. Lin, "Hete-cf: Social-based collaborative filtering recommendation using heterogeneous relations," in *2014 IEEE International Conference on Data Mining, ICDM 2014, Shenzhen, China, December 14-17, 2014*, 2014, pp. 917–922. [Online]. Available: http://dx.doi.org/10.1109/ICDM.2014.64

[3] Z. Qiao, P. Zhang, Y. Cao, C. Zhou, L. Guo, and B. Fang, "Combining heterogenous social and geographical information for event recommendation," in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, 2014, pp. 145–151.

[4] I. Boutsis, S. Karanikolaou, and V. Kalogeraki, "Personalized event recommendations using social networks," in *16th IEEE International Conference on Mobile Data Management, MDM 2015, Pittsburgh, PA, USA, June 15-18, 2015 - Volume 1*, 2015, pp. 84–93. [Online]. Available: http://dx.doi.org/10.1109/MDM.2015.62

[5] T. N. Pham, X. Li, G. Cong, and Z. Zhang, "A general graph-based model for recommendation in event-based social networks," in *31st IEEE International Conference on Data Engineering, ICDE 2015, Seoul, South Korea, April 13-17, 2015*, 2015, pp. 567–578.

[6] Q. Yuan, G. Cong, and C.-Y. Lin, "COM: A Generative Model for Group Recommendation," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '14. New York, NY, USA: ACM, 2014, pp. 163–172. [Online]. Available: http://doi.acm.org/10.1145/2623330.2623616

[7] W. Zhang, J. Wang, and W. Feng, "Combining latent factor model with location features for event-based group recommendation," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '13. New York, NY, USA: ACM, 2013, pp. 910–918. [Online]. Available: http://doi.acm.org/10.1145/2487575.2487646

[8] X. Liu, Y. Tian, M. Ye, and W.-C. Lee, "Exploring personal impact for group recommendation," in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, ser. CIKM '12. New York, NY, USA: ACM, 2012, pp. 674–683. [Online]. Available: http://doi.acm.org/10.1145/2396761.2396848

[9] J. She, Y. Tong, L. Chen, and C. C. Cao, "Conflict-aware event-participant arrangement," in *31st IEEE International Conference on Data Engineering, ICDE 2015, Seoul, South Korea, April 13-17, 2015*, 2015, pp. 735–746. [Online]. Available: http://dx.doi.org/10.1109/ICDE.2015.7113329

[10] J. She, Y. Tong, and L. Chen, "Utility-aware social event-participant planning," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '15. New York, NY, USA: ACM, 2015, pp. 1629–1643. [Online]. Available: http://doi.acm.org/10.1145/2723372.2749446

[11] C.-H. Lai, "Can our group survive? an investigation of the evolution of mixed-mode groups," *Journal of Computer-Mediated Communication*, vol. 19, no. 4, pp. 839–854, 2014. [Online]. Available: http://dx.doi.org/10.1111/jcc4.12075