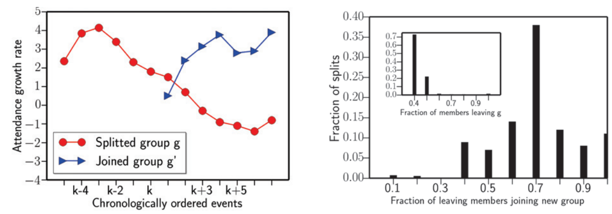
Proceedings of the Fourteenth International AAAI Conference on Web and Social Media (ICWSM 2020)

On the Splitting Dynamics of *Meetup* Social Groups

Ayan Kumar Bhowmick,[1] Soumajit Pramanik,[2] Sayan Pathak,[3] Bivas Mitra[1]
[1]Department of Computer Science & Engineering, IIT Kharagpur, India
[2]MPI for Informatics, Saarbrucken, Germany
[3]Microsoft Research, Redmond, WA, USA
ayankb@iitkgp.ac.in, pramanik@mpi-inf.mpg.de, sayanpa@microsoft.com, bivas@cse.iitkgp.ac.in

Abstract

Groups in online social networks witness continuous evolution by loss of existing members and gain of new members. In this paper, we present a study of *group split* in *Meetup*, where a major fraction of members leave the existing group together and join a newly formed group. We identify pivotal group members, called *splitters*, playing key roles in *group split* by influencing the existing members to leave the group. We provide an in-depth analysis of the empirical data to reveal key motivating factors leading to a *group split* and its subsequent impact. Finally, we develop a prediction model for early detection of *splitters*, as well as the group members likely to be influenced by the *splitter* to leave the group.



(a) Attendance growth rate for consecutive events of *splitted group g* and newly joined group g': Anecdote

(b) Fraction of *leaving members* joining new group; fraction of members leaving a group (inset)

Figure 1: *Meetup* group splitting dynamics

Introduction

Social networks are ubiquitous where a group of like-minded people can interact and come together to form a cohesive social community (Palla et al. 2009; Warner, Bowers, and Dixon 2012). Due to frequent changes in interactions between members, such social groups experience a natural evolution over time where they continuously loose existing members as well as gain new ones (Palla, Barabási, and Vicsek 2007). Members leaving one social group may lead to the formation of a new group from the existing one, which results in a *group split*. Such group splitting events are especially interesting since it may lead to the disappearance of existing social groups; on the other hand, formation of new groups encourages fresh ideas, concepts and activities. Hence, the social group splitting phenomenon raises multiple research questions such as (a) What leads to the splitting of an existing social group? (b) What are the (positive & negative) implications of *group split*? (c) Who plays critical roles behind *group split*? The objective of this paper is to shed some light on these questions associated with social group dynamics.

We consider *Meetup* (Liu et al. 2012), a popular event-based social network (EBSN) platform, to conduct the study. *Meetup* provides a convenient platform for the similar minded people to form *Meetup* groups, as well as host real-world events, where people get opportunities to participate in face-to-face interactions. Data study in Fig. 1(a) shows the incident where the popularity (number of participants) of the events hosted by the *Meetup* group "Women Entrepreneurs Secrets of Success" (designated as g) declines sharply after the k^{th} event, whereas a new group "PhotoMuse - Creative Photography Group" (designated as g'), formed after the k^{th} event of g exhibits an increasing trend in event popularity. Close inspection of Fig. 1(a) reveals the following two observations: (a) *Meetup* group g' is created due to a split in g after the k^{th} event, i.e. a significant population of g leaves the group after the k^{th} event and joins the new group g'. (b) The *group split* was initiated by a pivotal member, who leaves group g, creates the new group g' and influences a significant population of group g to leave g and join g'. We refer to this pivotal member as the *splitter* of the *splitted group* g. We observe that such split events occur for roughly 10% of the *Meetup* groups. Considering the broad impact of *splitters* on the survivability of *Meetup* groups, an in-depth investigation of this group splitting dynamics is an interesting research problem. Besides the academic interest, proper understanding of the splitting dynamics may facilitate the *Meetup* stakeholders (group organizers, event hosts etc.) to take necessary steps for retaining groups & events in the *Meetup* ecosystem.

State-of-the-art literature have primarily focused on the study of group evolution in social networks (Bródka, Saganowski, and Kazienko 2013; Doreian and Stokman 2013; Hessel, Tan, and Lee 2016). For instance, (Palla, Barabási, and Vicsek 2007) have studied the formation of social groups and revealed the type of groups that are likely to split into smaller sub-groups, while (Dakiche et al. 2019) have studied the evolution of dynamic communities in online social networks. In case of *Meetup*, existing attempts have primarily concentrated on predicting the survivability of the *Meetup* groups (Lai 2014a; 2014b). For instance, (Ribeiro 2014) have revealed the factors that may eventually lead to the death of a *Meetup* group, while (Pramanik et al. 2016; Li et al. 2017) have predicted the success of *Meetup* groups using feature-based models. However, the existing literature have mostly overlooked the interplay of *Meetup group split* vis-a-vis *Meetup* group formation and the role of *splitters* in this context.

The major contribution of this paper is to explain the dynamics of social group splitting phenomenon in the context of *Meetup*. First, we mine the *Meetup* dataset in order to detect *group split* events and corresponding *splitters*. Moreover, we also identify the *leaving members* of a group who are influenced by *splitters*. Once we recognize these key players, we perform a detailed study revealing the major motivations and impact of splitting behavior in *Meetup*. We show that *group split* in *Meetup* mostly occurs when events hosted by a group do not align with the interests of a major population of the group; this results in a steady decline of event attendance. The formation of a new group by the influential *splitter* causes an avalanche effect in this dynamics. On the other hand, events hosted by the newly formed group observe a continuous rise in popularity over time. Finally, leveraging these insights, we develop prediction models for early detection of *splitters*, as well as detection of the influenced population leaving the *Meetup* group. In these models, we have relied on both interest-based and influence-based features; we observe that our models exhibit decent performance with traditional classifiers achieving more than 90% *F1-score*, depicting the robustness of our discovered features.

Dynamics of *Group Split* in *Meetup*

Dataset & Notations

We have crawled detailed information about *Meetup* groups, members and events across three US cities from August 2015 to August 2019. For each *Meetup* group and group member, we collect information about the set of tags describing their interests. For every *Meetup* event, we collect basic information in the form of its location, time and textual description. In addition, we obtain RSVP responses from participants attending the events. Detailed statistics are summarized in Table 1.

Let $\mathcal{U}$, $\mathcal{E}$ & $\mathcal{G}$ denote set of all members, events and groups respectively in *Meetup*. The *popularity* of a *Meetup* event $e \in \mathcal{E}$ is measured as the volume of participating members (event attendance a_e), which can be approximated from the number of 'yes' RSVP responses. Detailed information of event e can be obtained from its textual description denoted by $\mathcal{Q}_e$. We rely on *fastText* (Bojanowski et al. 2017) to learn a latent vector representation of $\mathcal{Q}_e$ denoted as $\mathbf{v}_e$.

Table 1: Dataset overview

City	Groups	Members	Events
Chicago	7718	427613	458087
New York	23270	1192431	1008317
San Francisco	17647	848032	713967

Definition and detection of *group split* and *splitters*

Consider a group $g \in \mathcal{G}$ such that a member u leaves g at time t and creates a new group $g' \in \mathcal{G}$ at time $t' > t$. We define this phenomenon as a *group split* and u is called as a *splitter* of source group g. Suppose L_g denotes the set of all members of group g who leave[1] g at time t (termed as the *leaving members*). If $u_j \in L_g$ joins the new group g' at time $t'_j > t'$ following the *splitter* u, then u_j is said to be influenced by the *splitter* u to join g'. We designate the set of all such *leaving members* as $L_g^u \subseteq L_g$, denoting those population of g influenced by the *splitter* u.

First glimpse: Role of *splitters*

Splitters result in a sharp decline in group size for a *splitted group* g. Fig. 1(b) (inset) shows that once the *splitter* leaves a group g, a large fraction of people denoted by L_g leave g and follow the *splitter* to join the newly formed group g'. In fact, we find that around $40 - 50\%$ of the group members exit the group for a majority (90%) of *group split* events. Fraction of group members exiting the group is even higher (more than 70%) for remaining 10% of cases. Moreover, our data study also reveals that 38% of groups disappear after a *group split* as they stop hosting events within 5 events after the occurrence of *group split*.

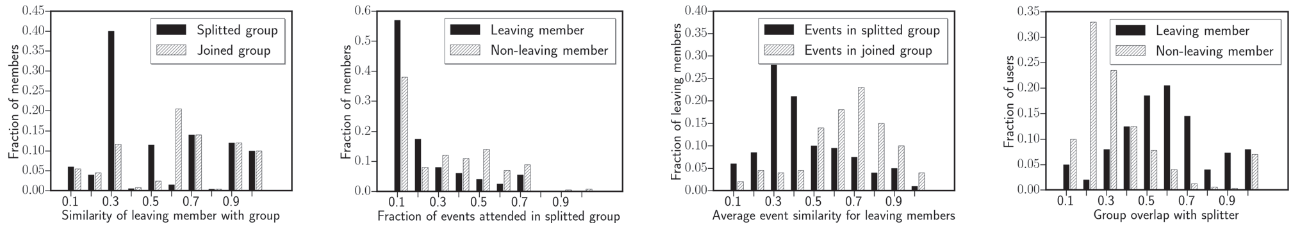
Further, we investigate what happens to the *leaving members* of g after the *group split*. Fig. 1(b) shows that a high fraction (> 0.7) of *leaving members* L_g join a single group g', formed by the *splitter*, after they exit from the *splitted group* g, for nearly 70% of *group split* events. This implies that a majority of *leaving members* of g, denoted as L_g^u, join a single group after they exit from g, instead of getting dispersed across multiple different groups.

Motivation behind *Group Split*

In this section, we delve deep and identify the motivating factors leading to a split of the source group.

(a) Role of topical interests of the source group: Lack of interest match of the *leaving members* L_g with the *splitted group* g plays a major role in *group split*. We compute the interest overlap in terms of Jaccard coefficient between the sets of tags of a *leaving member* $u_j \in L_g$ and the splitted group g, denoted as T_{u_j} and T_g respectively. Fig. 2(a) shows that a significant fraction of *leaving members* (50%) exhibit low similarity (< 0.3) with group g from which it splits. On the other hand, the joining group g' of this population exhibits significant interest overlap. Fig. 2(a) shows that a majority (60%) of *leaving members* of g that join g', denoted as L_g^u, exhibit high tag similarity (greater than 0.6) with g'.

[1]We assume that a member left a group if she does not attend any further event of that group.



(a) Similarity of *leaving members* with groups

(b) Fraction of events attended by *leaving members*

(c) Event similarity of *leaving members* with joined group

(d) Similarity of group members with corresponding *splitters*

Figure 2: Motivation of *group split*

This indicates that group members tend to leave a group in pursuit of joining a new group with which their interests are more likely to match.

(b) Drop in interest in *Meetup* events hosted by source group: Lack of interest in the events hosted by the *splitted group* may lead to drop in event attendance for *leaving members*. We compute the fraction of events that a *leaving member* $u_j \in L_g$ attends in source group g before leaving g. From Fig. 2(b), we observe that 75% of *leaving members* attend less than 20% events hosted by the group before leaving, while only 46% of non-leaving members attend less than 20% of a group's events. This implies that albeit non-leaving members maintain their activity at a lower level in *splitted group*, unlike *leaving members*, they continue to attend the group's events at regular intervals without leaving the group.

This drop in event attendance for *leaving members* can be explained in terms of their interest overlap with the *Meetup* events hosted by group g. Let $\mathcal{E}_{u_j}$ denotes the set of all *Meetup* events attended by a *leaving member* $u_j \in L_g^u$ and $\mathcal{E}_g$ denotes the set of last k events hosted by *splitted group* g before *group split*. Similarly, we denote $\mathcal{E}_{g'}$ as the set of first k events hosted by the newly created group g' where the *leaving member* u_j joins after *group split*. Now, we compute the average similarity of $\mathcal{E}_{u_j}$ with the individual events in $\mathcal{E}_g$ & $\mathcal{E}_{g'}$ and denote them as $s_g^{u_j}$ and $s_{g'}^{u_j}$ respectively. In both cases, the similarity is measured in terms of cosine similarity between feature vectors $\mathbf{v}_e$[2] of respective events. From the distribution of $s_g^{u_j}$ and $s_{g'}^{u_j}$ values plotted in Fig. 2(c), we observe that only 17% of the values in case of $s_g^{u_j}$ are greater than 0.7 while 53% of values for $s_{g'}^{u_j}$ are greater than 0.7. This shows that type of events that *leaving members* of g prefer to attend are closely aligned with events hosted by newly created group g' compared to events hosted by g. Hence, *leaving members* prefer to join g' due to interest match.

(c) Influence of *splitters*: *Leaving members* of a group exhibit similar group membership behavior as the *splitter* of the group. First, we compare the similarity of group members (both leaving and non-leaving) of a *splitted group* g with the *splitter* u in terms of their group membership behavior and plot the distribution in Fig. 2(d). From this figure, it can be observed that the group membership overlap between a *leaving member* and a *splitter* is greater than 0.3 for 75% of such cases. On the other hand, the group overlap between a non-leaving member and a *splitter* is < 0.3 for 66% of the cases. This demonstrates that the *splitters* have high interest overlap with the *leaving members* compared to the non-leaving ones, indicating the potential role of *splitters* in motivating *leaving members* to leave the *splitted group* g and join g' created by *splitters*.

Effect of *Group Split*

In this section, we analyze the impact of *group split* on the popularity of the *splitted group* as well as the newly created group in the long run.

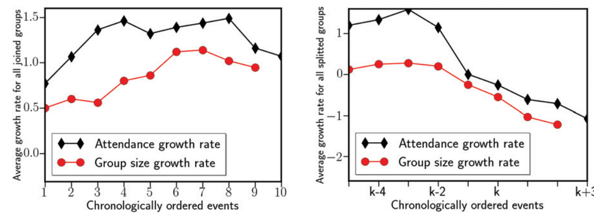
(a) Observe elegance in newly created group: Given the attendance of two consecutive events e_i and e_{i+1} denoted as a^{e_i} and $a^{e_{i+1}}$ respectively, we compute the attendance growth rate of event e_{i+1} as $\frac{a^{e_{i+1}} - a^{e_i}}{a^{e_i}}$. We compute the attendance growth rate for sequence of events hosted by a newly joined group g' created by the *splitter* of *splitted group* g, starting from the first event of g'. In a similar way, we compute the sequence of group size growth rates. In Fig. 3(a), we plot the average growth rates of group size and event attendance over sequence of events hosted by newly joined groups in case of all *group split* events. It is observed that the average event attendance growth rate over consecutive events of g' rises at a fixed rate for first few events and then remains steady at a high positive value. This implies that attendance of consecutive events of g' continues to increase over time. Similar trend is observed for the group size growth rate. This implies that newly joined groups continue to gain new members from the *splitted group* and from other groups, leading to a healthy growth in its popularity and hence such groups are likely to sustain over a longer period of time.

(b) Observe decline in *splitted group*: We now plot the average growth rates of event attendance and group size over the sequence of events hosted by all *splitted group*s, taking 5 events before and after *group split* occurring at event k in Fig. 3(b). Here we observe that, for few events hosted much before the *group split* occurs, event attendance growth rate increases steadily. However, just before the *group split* at event k, this growth rate drops (though it remains positive). For events following *group split*, the attendance drops at a faster rate, with the growth rate becoming negative over subsequent events. A similar observation can be seen for the average group size growth rate (computed similarly to event

[2] Generated using *fastText* (Bojanowski et al. 2017)

Table 2: Classification performance for *splitter* detection and prediction of *split pairs* in *Chicago* city

	Model A		Model B	
Classifier	**Accuracy**	**F1-score**	**Accuracy**	**F1-score**
Decision Tree	0.935	0.937	0.913	0.909
SVM	0.827	0.834	0.819	0.824
GradientBoost	0.875	0.884	0.893	0.895
Logistic Regression	0.802	0.817	0.807	0.806



(a) Joined groups after split (b) *Splitted groups* (before & after split)

Figure 3: Average growth rate of group size and event attendance

attendance growth rate) over 5 events hosted before and after the *group split* from Fig. 3(b). This happens because most of the group members attend the group's earlier events hosted before the split, as the group continues to gain members which explains the steady increase. Just before *group split* occurs, a section of existing group members stop attending events that leads to fall in the growth rate. Finally, several group members start leaving the group in large numbers at subsequent events after *group split*. Hence, the growth rate drops to below zero (negative) since majority of the group members have left the group.

Predicting *Splitters* in *Meetup* Groups

In the previous section, we have explored various factors motivating the group splitting phenomenon and its effect on the associated stakeholders. Leveraging on these insights, we propose two machine learning based models in this section to - (a) detect *splitters* early and (b) identify *leaving members* who leave the *splitted group* under the influence of *splitters*. We evaluate the models only for the *Chicago* city for the interest of space and readability.

Model A: Detecting *splitters*

In this model, we predict whether u, a member of source group g, would potentially split g after next K events.

Feature selection In order to detect *splitters*, we first define two sets of features which distinguish the *splitters* from the rest of the group members.

(a) Interest based features: We identify the following features based on member interest:

(1) **Tag-based features:** Every member $u \in \mathcal{U}$ in *Meetup* is characterized by a set of tags T_u specifying her interests. We use three tag-based features for each member u of group g - (i) number of tags of u ($|T_u|$), (ii) tag similarity of u with group g's tags and (iii) average tag similarity[3] of u with other members of g.

(2) **Group activity:** It measures the fraction of events hosted by a group g, in which a member u participates.

(3) **Multiple group membership:** It measures the number of groups in $\mathcal{G}$ in which u is a group member.

(4) **Activity level:** It measures the total number of *Meetup* events in which u has participated.

(b) Influence based features: We identify two features measuring the capability of a group member u to influence other members of g:

(1) **Group joining influence:** This measures the influence power of u in convincing members of other groups to join her group. Let u be a member of a group g while u_j is not a member of g at time t_1. Suppose u and u_j participate in a common event $e \in \mathcal{E}$ at time $t_2 > t_1$ and subsequently u_j joins g at time $t_3 > t_2$. In such a scenario, we say that u has influenced u_j to join group g. For the member u, this feature is computed as the number of members u has influenced to join her group.

(2) **Event participation influence:** This measures the influence power of u in convincing other members of her own group to participate in events. Given an event e hosted by g, suppose u and u_j are i^{th} and j^{th} members to send 'yes' RSVP to attend e at times t_1 and t_2 respectively such that $t_2 > t_1$. We say that u has influenced u_j to attend e if $t_2 - t_1 \leq p$, $j - i \leq q$ and this incident occurs for more than r times (where p, q and r are some threshold values empirically chosen to be 6 hours, 3 and 5 respectively for our experiments). This feature is computed as the number of members u has influenced to participate in an event.

Experimental setup For ground truth labeling, we label a member u as a *splitter* of g if she leaves g at time t and creates (& joins) a new group g' at time $t' > t$. Otherwise, we label her as a *non-splitter* of g. In order to detect the *splitters* K events early, we compute the aforementioned features leaving out the most recent K events before the split. We use $K = 5$ in our experiments and apply standard classifiers for the prediction task.

Results We demonstrate the results for classifying *splitters* using standard machine learning classifiers in Table 2, by applying $10-$fold cross validation. As we can observe, the best classification accuracy is 93.5% while the best *F1-score* is 94% with *Decision Tree* performing the best.

Model B: Detecting *leaving members*

In this model, given a member u splitting a group g and creating group g', we predict whether u_j (another member of g) would follow him i.e. leave the group g and join group g'.

Feature selection Similar to the previous model, here also we define two sets of features which might influence a group member's decision of following the *splitter*.

[3]The similarity between two sets of tags T_1 and T_2 is computed using Jaccard coefficient (Niwattanakul et al. 2013) as $\frac{|T_1 \cap T_2|}{|T_1 \cup T_2|}$.

(a) Interest based features: These are based on similarity of interests between a pair of members u and a u_j:

(1) **Tag similarity:** It is measured as the Jaccard similarity between the sets of tags T_u and T_{u_j} for u and $u_j \in L_g^u$ respectively.

(2) **Event overlap:** It is measured as the Jaccard similarity between sets of events attended by u and u_j.

(3) **Co-group membership:** It is measured as the Jaccard similarity between sets of groups that u and u_j have joined.

(b) Influence based features: We use the same influence-based features used in Model A. However, here we measure the influence of the *splitter* u specifically on the group joining and event participation behavior of $u_j \in L_g^u$, computed as the number of times u influences u_j to join a group or attend an event.

Experimental setup We label a pair of group members u, u_j of g as a *split-pair* if u is a *splitter* of group g and influences the *leaving member* u_j to join new group g' at time t' created by u; otherwise, we label the pair u, u_j as a *non-split-pair*. Here also, we compute the features leaving out the most recent K events before the split. We take $K = 5$ in our experiments and apply standard classifiers for the prediction task.

Results We demonstrate the result of classifying pairs of group members using standard classifiers in Table 2 where we observe that we obtain good classification performance with *Decision Tree* performing the best in terms of *Accuracy* and *F1-score* values of 0.908 and 0.905 respectively.

Overall, we can conclude that the proposed features are highly robust in early prediction of *splitters* in a group and predicting group members influenced by *splitters* to leave the group, even using simple non-neural classifiers. Due to lack of sufficient data points for *group split*, we have refrained from using deep neural classifiers. However, with availability of abundant data, deep neural models are expected to further improve model performance using proposed features.

Conclusion

This paper puts forward the importance of *splitters* in splitting dynamics of *Meetup* social groups. In general, failure of a *Meetup* group may be contributed by several factors such as gradual departure of members, lack of interest in attending events, infrequent event hosting etc. We define *group split*, a prime indicator of group failure, where a significant fraction of members leave their current group together, and (majority of them) join a new group formed by a pivotal *leaving member*, called *splitter*. Our study revealed that *group split* becomes inevitable when the interest profile of a major population in the group deviates from the declared interest of the group or its hosted events. Moreover, most of the *leaving members* get highly influenced by the *splitter* to join her newly created group. Unlike the *splitted group*, the events hosted by the newly formed group gains high popularity by hosting events of popular choice among its group members. Finally, we leverage on the aforesaid insights to develop a simple prediction model for early detection of the *splitters* and the influenced population. We believe that the analysis and the model presented in this paper will benefit the *Meetup* stakeholders in maintaining the sustainability of *Meetup* groups.

References

Bojanowski, P.; Grave, E.; Joulin, A.; and Mikolov, T. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics* 5:135–146.

Bródka, P.; Saganowski, S.; and Kazienko, P. 2013. Ged: the method for group evolution discovery in social networks. *Social Network Analysis and Mining* 3(1):1–14.

Dakiche, N.; Tayeb, F. B.-S.; Slimani, Y.; and Benatchba, K. 2019. Tracking community evolution in social networks: A survey. *Information Processing & Management* 56(3):1084–1102.

Doreian, P., and Stokman, F. N. 2013. The dynamics and evolution of social networks. In *Evolution of social networks*. Routledge. 9–26.

Hessel, J.; Tan, C.; and Lee, L. 2016. Science, askscience, and badscience: On the coexistence of highly related communities. In *Tenth International AAAI Conference on Web and Social Media*.

Lai, C.-H. 2014a. Can our group survive? an investigation of the evolution of mixed-mode groups. *Journal of Computer-Mediated Communication* 19(4):839–854.

Lai, C.-H. 2014b. Understanding the evolution of bona fide mixed-mode groups: An example of meetup groups. *First Monday* 19(1).

Li, G.; Liu, Y.; Ribeiro, B.; and Ding, H. 2017. On group popularity prediction in event-based social networks. In *12th International AAAI Conference on Web and Social Media, ICWSM 2018*, 644–647. AAAI press.

Liu, X.; He, Q.; Tian, Y.; Lee, W.-C.; McPherson, J.; and Han, J. 2012. Event-based social networks: linking the online and offline social worlds. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1032–1040. ACM.

Niwattanakul, S.; Singthongchai, J.; Naenudorn, E.; and Wanapu, S. 2013. Using of jaccard coefficient for keywords similarity. In *Proceedings of the international multiconference of engineers and computer scientists*, volume 1, 380–384.

Palla, G.; Barabási, A.-L.; and Vicsek, T. 2007. Quantifying social group evolution. *Nature* 446(7136):664.

Palla, G.; Pollner, P.; Barabási, A.-L.; and Vicsek, T. 2009. Social group dynamics in networks. In *Adaptive Networks*. Springer. 11–38.

Pramanik, S.; Gundapuneni, M.; Pathak, S.; and Mitra, B. 2016. Predicting group success in meetup. In *Tenth International AAAI Conference on Web and Social Media*.

Ribeiro, B. 2014. Modeling and predicting the growth and death of membership-based websites. In *Proceedings of the 23rd international conference on World wide web*, 653–664. ACM.

Warner, S.; Bowers, M. T.; and Dixon, M. A. 2012. Team dynamics: A social network perspective. *Journal of Sport Management* 26(1):53–66.