Predicting Group Stability in Online Social Networks

Akshay Patil
Stony Brook University
Stony Brook, NY
akshay@cs.stonybrook.edu

Juan Liu
Palo Alto Research Center
Palo Alto, CA
Juan.Liu@parc.com

Jie Gao
Stony Brook University
Stony Brook, NY
jgao@cs.stonybrook.edu

ABSTRACT

Social groups often exhibit a high degree of dynamism. Some groups thrive, while many others die over time. Modeling group stability dynamics and understanding whether/when a group will remain stable or shrink over time can be important in a number of social domains. In this paper, we study two different types of social networks as exemplar platforms for modeling and predicting group stability dynamics. We build models to predict if a group is going to remain stable or is likely to shrink over a period of time. We observe that both the level of member diversity and social activities are critical in maintaining the stability of groups. We also find that certain 'prolific' members play a more important role in maintaining the group stability. Our study shows that group stability can be predicted with high accuracy, and feature diversity is critical to prediction performance.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications—*Data Mining*; J.4 [**Computer Applications**]: Social and Behavioral Sciences

Keywords

Social Networks, Group Stability, Online Communities

1. INTRODUCTION

Understanding community structures has always been an interesting topic in social sciences. In many social network datasets, a social graph is presented in which nodes represent individuals and edges represent social ties. It is a common experience to observe community structure in such a graph, in the sense that a subset of vertices are well connected within them and less connected to the rest of the graph. For example, communities in a social network often represent social groupings, say by interest or background. Communities in a publication network may represent people who work on similar research problems. Communities of the web graph may suggest pages on related topics. As the complexity of online activity increases, formal group structures have come to play an increasingly important role in the experience and effectiveness of an individual's online life.

Online platforms have provided unprecedented opportunities to study large-scale behavior and dynamics of communities. A lot of studies have focussed on how to define and detect social communities in the network structure and how the groups evolve over time. For the later, the main thrust of such research has been to model the evolution of groups, from the standpoint of growth [6, 27, 19]. A community in these studies always grows. What to be examined is the rate of growth and when the community stops growing. There are two main reasons for this. First, in many online social network settings there is no restriction on the number of groups an individual belongs to. Also, in most cases individuals do *not* quit groups even though they may not be active participants in those particular groups. This often results in groups having a monotonically increasing membership curve throughout their lifetimes. In addition, a practical and commercial motivation for such studies has been to increase the 'stickiness' of an online community, i.e., the capability for it to attract new members. Therefore, a common model in modeling group growth is to consider it as a diffusion process. That is, the social ties that cross group boundaries may influence people not yet in the group to join the group. This observation has been one of the main philosophies in modeling and predicting the growth of online groups. Studies have been performed in examining how diffusion happens and what is the main factor in determining the speed of diffusion. It has been shown using Facebook data [2] that what attracts a new user depends not only on the number of friends on Facebook, but also on the diversity of these friends, as well as the network connectivity structure among them [24].

In this paper, we take a different perspective and study the complementary problem of group stability, i.e., why some groups fall apart and disappear while others thrive. The effectiveness of groups can be undermined when group members depart, taking with them, experience, resources and possibly other group members. The ability to predict the stability of groups is highly desirable, as it offers insights on factors that affect online group effectiveness. It also provides practical guidance to tasks such as risk management and customer retention.

In some online settings an individual can belong to *only* one group at any given point in time. In such settings the group serves as the main engagement platform for the individual. An individual who is not satisfied with his/her group will quit the group and join another one. The reasons for dissatisfaction can be plenty. In such cases the percentage increase/decrease in the number of group members over previous time periods is a good measure in determining whether a group is stable or shrinking. In the settings when users can join multiple groups and probably never quit these groups, the group size always grows. But the growth in group size does not necessarily capture the accurate picture. A group, though accumulating members over time, may still be a shrinking group, if most members do not participate in the group's activities. In addition, most previous studies treat all group members as equals when performing group evolution analysis. We know that groups are often led by a smaller set of leaders who have considerable influence over other group members. In our analysis we take both

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2013, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2035-1/13/05.

issues into consideration. For settings that allow multiple group memberships and do not have group quitting events we devise a membership score that will reflect participation level of individual members and prolificness/ranking of individual members.

We perform our analysis on two different types of social networks, a massive multiplayer online role-playing game (World of Warcraft [WoW]) and a large co-authorship network (DBLP). In the first dataset we tackle the scenario of an individual belonging to at most one group (a guild, in WoW terminology) at any given point in time and the second tackles the more general scenario of an individual belonging to multiple groups at any given point in time. Moreover for the second scenario, we also devise a membership score that we believe is more reflective of the stability/growth of a group (as compared to the number of members in the group). This membership score can be easily generalized for a host of social networks. Though the membership score was devised to encapsulate the growth (or lack of) of a group it can also be used to compare groups, as we will demonstrate later. We have built classifiers based on a diverse set of features to predict whether a group will have significant reduction or will remain stable over a period of time.

In our findings regarding the two datasets, we have the following interesting observations:

- We find that the level of diversity has a strong correlation to the stability of the group. In order to keep a group alive, members of the community should vary in terms of expertise, seniority, responsibilities, etc. We also find that the level of activities has a strong predictive power of the group stability. Even when the size of the community stays the same (i.e., not attracting new members), as long as there is a lot of activities within the group the community survives.

- We find that in the case of WoW dataset the age of a community has a strong correlation with the stability of the group — if a guild can sustain itself for a long period of time, it is very likely that the guild does have the essential components necessary for a stable community. On the hand in the DBLP dataset the length of existence does *not* show any correlation with group stability. Whether a conference is old or new does not seem to play a significant role in determining whether it remains stable or shrinks. This observation can be attributed to fact that in WoW there is a lot of churn, whereas in DBLP (or other related social networks) we do not see as much churn.

- For DBLP dataset we observe that the 'average prolificness' feature is important. The correlation shows that groups with more prolific members are more likely to remain stable and groups with more dedicated authors (i.e. authors who continually contribute) are more likely to remain stable. Thus such members play an important role in maintaining the stability of the group.

The paper is organized as follows. We first briefly review literatures on detection of communities and studies of community evolution. We then provide an overview of the datasets, followed by definition of measures used to label groups as stable or shrinking. We then move on to define a range of features that we compute for both datasets. We will also analyze the best set of features that are useful for our prediction task. The later sections present predictive models to predict group stability. We conclude with a discussion of the important factors for predictions and an outline of future work.

2. RELATED WORK

Various methods have been used to detect communities. See the survey paper [21] for a thorough review. Earlier approaches define some measurement of importance of each edges and then define communities by either incrementally adding edges in the order of decreasing importance [25]; or removing edges in the order of increasing importance [14]. This leads to a hierarchical partitioning of the nodes, called a *dendrogram*. Classical clustering techniques are also used here, including k-means clustering, multidimensional scaling, principal component analysis, etc. Same for methods that identify clique-like components, or find min-cuts in graphs. In our datasets, community structures are formally defined and explicitly given, hence there is no need to detect them.

When time-stamped data is available it is natural to ask how the communities or groups evolve over time. In the literature the community evolution has been modeled as a diffusion process – ties spanning group boundaries can possibly influence individuals to join the community. Granovetter [15] pointed out that diffusion often benefits from 'weak ties' and indicated that the graph structure may be a critical factor in deciding whether and how fast a community grows. Recent studies, such as by Centola and Macy [10] and from Facebook datasets [24], revealed that one may require multiple contacts within the group to join the community and the diversity of these contacts actually matters. A couple of the analysis using real world datasets show conflicting and intriguing observations that high clustering property inside a community may at the same time both attracts new members and prevents overall growth. A very recent paper by Kairam *et al.* [18] pointed out that new members may join through diffusion (as in the case of being influenced by some friends), or may join the community without having any social ties inside the community, classified as non-diffusion growth. They further point out that in diffusion-based growth, the clustering does help. But groups that only grow through diffusion may not reach large size. Thus non-diffusion growth is important to create large communities.

Most of existing work on community evolution focused on the initial stage of community evolution, when growth is in the dominant form. Our work, on the other hand, mainly looks at the final stage of community evolution, i.e., how a community dies or falls apart. The closest work to ours is the research on group formation in large social networks [6]. They build classifiers based on a range of network-based features to firstly, predict whether an individual will join a community and secondly, to predict the growth of a community. They achieve reasonably good $(70-75\%)$ accuracy for both prediction tasks. The point to note is that members never quit communities in their model. Thus, for the second prediction task they are predicting from the standpoint of growth. Our work is complementary to their work. In our previous work [20] we have built models to predict if and when an individual is going to quit his/her group, and whether this quitting event will inflict substantial damage on the group. We quantify damage as influencing many of your friends to also quit the group after you do so, thereby leading to a large loss in group membership numbers. In [20] we analyzed quitting from an individual perspective, while this paper addresses the quitting behavior from a group perspective.

3. ANALYSIS ON WORLD OF WARCRAFT

3.1 Dataset

To explore group stability dynamics, we use data from a previous World of Warcraft (WoW) study [13]. WoW is a multiplayer online game in which users interact, collaborate with, or fight against each other. A web-based crawler was deployed to log in-game activities based on the API specified by Blizzard Entertainment, the producer of WoW. The crawler periodically issues "/who" requests every 5 to 15 minutes, depending on server load, to get a list of characters currently being played on a given server. We have data that spans six months, from November 2010 to May 2011. The data is sometimes referred to as the WoW census. Three types of servers are logged: player-vs-environment (PvE), player-vs-player (PvP), and role playing (RP). The servers may present players with different game tasks, but are otherwise identical in terms of game organization and support. In the game there is a social group setting named a 'guild'. Players of the same guild often organize to join battles, gain honors or even monetary returns. In WoW, one player can only join one guild and to join another he/she has to quit the former guild. Overall we observed more than 470,000 unique characters forming over 15000 guilds, scattered on three servers: Eitrigg (a PvE server), Cenarion Circle (a RP server), and Bleeding Hollow (a PvP server).

Social interaction may be an important influencing factor in guild-quitting events. First, we define a friendship network among guild members, where nodes are characters, and edges indicate co-occurrence within gaming zones — if two characters were observed at the same game location (zone in WoW), an edge is added between the corresponding nodes. A gaming zone is a predefined area in the WoW map. A zone can be small or large in size & can contain varying number of characters at any given point in time, depending on several parameters in the game. The underlying assumption is that if characters co-occur in a gaming zone, it is highly likely that the characters are collaborating on a gaming activity. Two possible limitations are noted: (1) there are some gaming zones that are not necessarily associated with any gaming activity, for instance, characters are often left "AFK" (Away from keyboard) in the game's main cities before or at the end of a play session. In this case, the geographic proximity does not necessarily reflect any kind of joint activity. In our data logger, we remove such ambiguous zones from the co-occurrence criteria. (2) Characters may co-occur by chance. This is treated as noise in the social network graph. The basic assumption is that with a large amount of accumulated gaming data, the ties between characters driven by real social interaction will dominate.

Secondly, we add a membership network to indicate the affiliation between characters and guilds. An important point to remember is that a character can belong to only one guild at given point in time. Thus in order to join another guild he/she is required to quit his/her current guild. Nodes fall into two categories: (1) guild nodes, and (2) character nodes. If a character is observed appearing in a guild, an affiliation edge is added. The overall network is the super-imposition of the friendship and the affiliation networks. It is an undirected multi-graph i.e. it allows for multiple edges between any two nodes in the network.

Table 1 lists some statistics in the raw social networks for the three servers. Guild quitting events are fairly common — around $20 - 32\%$ of characters quit from a guild at least once in our observation period. In constructing our social network using the co-occurence heuristic, we eliminate characters that do not join any guilds or collaborate with any other characters (i.e. characters that do not have any social component). Such characters are generally played by new players at initial levels of the game. These characters are uninteresting from the point of view of the problem we seek to tackle and thus can be ignored from our analysis. Similarly we ignore degenerate guilds (i.e. guilds which are observed to have no members over our entire 6 month time-period) from our analysis.

3.2 Gauging Group Stability

It is often of general interest to understand the stability issues of social groups, for instance, the stability of a company, an informal organization, or a user group. In WoW and other MMORGs, as mentioned earlier, guilds have high turn-out rates. In our observation data of over 6 months, some guilds live throughout (188 days), but many other do not survive very long. The average guild lifespan is 82.57 days, with a large standard deviation of 71.25 days. This begs the question, "Why are some guilds more stable than others? In other words, what constitutes a stable guild?"

Guild stability may be related to a variety of factors, some of which have been identified from social psychology studies. In this section, we take a data driven approach — "Can we identify stability or instability patterns from the data?"

It turns out that since a character can only belong to one guild at any given point in time, computing the number of guild members at regular intervals should give us a good idea of how the guild evolves. Futhermore computing the percentage of increase/decrease in the number of members over the previous interval would then give us an accurate idea of whether, (a) the guild is stable (i.e. there is a minimal percentage decrease or a percentage increase in the number of members), or (b) the guild is shrinking (i.e. there is a substantial percentage decrease in the number of members).

To put things more formally, given that a guild G has m_1 members at time snapshot t_1 and m_2 members at time snapshot t_2 ($t_1 < t_2$) the percentage change in membership is defined as $\delta = \frac{m_2}{m_1} - 1$. We could then label a guild as being in the stable or shrinking phase at time t_2 as follows,

$$label = \begin{cases} stable, & \text{if } \delta > -0.15 \\ shrinking, & \text{if } \delta \leq -0.15 \end{cases} \quad (1)$$

Thus we label a guild as shrinking if it loses 15% or more of its members as compared to the previous interval, otherwise the guild is labeled as being stable. In our experiments, we have experimented with several values of δ. Results were comparable when δ ranged between $[0.10, 0.20]$. For $0 < \delta < 0.10$, accuracy was reduced due to addition of noisy/fringe samples to the "shrinking" set. For $\delta > 0.20$, we again observe a drop in accuracy due to vastly fewer "shrinking" samples. Thus, in our experiments, we use a threshold of $\delta = 0.15$ to label whether a guild is stable or shrinking at any given point in time.

3.3 Guild-Level Features

In order to be able to model group stability dynamics, we consider a range of features that span different categories. Almost all of the features are efficient (with linear running time) to compute thereby allowing to compute them at regular intervals & also making our approach scalable to large networks.

Several types of guild-level features may be important in modeling guild stability. We loosely categorize them into three categories: (1) guild composition, (2) game activities aggregated over the guild population, and (3) the structure of social network graph.

Guild composition features reflect diversity or homogeneity of guild members. In WoW, guilds need to have a certain span in skills and roles. For instance, a healthy blend of experts and novices may be important to a guild's long-term survival. Novice players can mature in the game and take over if an expert leaves the guild. In addition, WoW activities are designed to encourage collaboration across roles. Characters are categorized into 10 classes (warriors, paladins, hunters, priests, death knights, etc) with different capabilities (DPSs to cause damage, tanks to contain damage, and healers to heal damage). Coordination among the classes and capabilities is

Statistic	Eitrigg	Bleeding Hollow	Cenarion Circle
Number of Characters	51,224	72,108	47,499
Number of Guilds	2906	3425	2911
Number of Friendship Edges	577,250	937,989	673,502
Number of Membership Edges	1,870,327	2,775,401	2,154,287
Average Collaboration Time (hrs.)	1.73 ± 1.09	1.70 ± 1.10	1.79 ± 1.08
% Characters changing Guild	26.53	32.28	20.69

Table 1: Overall Network Statistics for World of Warcraft

essential to the guild's success. There is a "sweet spot" of diversity, where, a certain degree is desirable, while excessive diversity may be a sign of lack of management and may imply poor guild performance. To investigate the effect on guild stability, we compute the following guild composition features:

- Number of guild members: The size of the guild at time t.
- Length of existence (in days): This feature calculates the number of days since the guild came into existence.
- Average level of guild members: This measures the average level of characters who are members of this guild at time t.
- Standard deviation of character levels of guild members: This feature measures how consistent good/bad characters are across the guild. A smaller standard deviation indicates a bunch of members with equal skill sets and a higher standard deviation indicates a bunch of members with varying skill sets.
- Percentage of character classes present: A character can belong to any one of around 10 character classes. This metric measures whether all character classes are represented amongst its members.
- Entropy of character class distribution: This feature calculates the entropy of the class distribution for the given guild.
- Entropy of character category distribution: There are 3 categories of characters in the game, DPS, Healers and Tanks. Each character class can perform one or more of these category roles. This feature calculates the entropy for category distribution within a guild.

Aggregated game activity features measures the overall game engagement across guild members. We contrast the game activities within the guild and prior to joining the guild. The former is indicative of the devotedness to the guild, while the latter measures the overall engagement in the entire WoW game.

- Average playing time within guild & prior to joining the guild: The features calculate average playing time of all guild members in the current guild and prior to joining the guild respectively.
- Average collaboration time within guild & prior to joining the guild: The features calculate the average collaboration time of all guild members in the current guild and prior to joining the guild respectively.
- Average collaboration coefficient within the guild & prior to joining the guild: The collaboration coefficient for a guild member is defined as the ratio of his/her collaboration time to playing time. These two features compute the average collaboration coefficient across all guild members by considering collaborations within the guild and prior to joining the guild respectively.

We suspect that guild topological structure may have implications on guild stability. Guilds exhibit remarkable diversity in topology, some guilds have a hierarchical structure, where some nodes (typically guild leaders) are of central importance, while other guilds are formed of closely knitted friendship circles, where nodes are more evenly connected. One may speculate that a star topology may be less stable since the removal of the center node may cause the whole graph to fall apart. In the topological features, we measure the average clustering coefficient, which is a metric of degree to which nodes in a graph tends to cluster together. Based on the concept, we compute the following topological features:

- Average clustering coefficient of guild members: We measure the clustering coefficient at each guild member node and then calculate the average of this clustering coefficient across all guild members. The clustering coefficient at each node is also known as the local clustering coefficient [26] and quantifies how close its neighbors are to being a clique.
- Average clustering coefficient of guild members within the guild: The clustering coefficient calculated in this case only takes into consideration the graph induced by all members of the guild.
- Entropy of degree distribution: This feature is a good measure of diversity in node connectivity.

3.3.1 Feature Importance & Correlation

The observation data is organized into temporal snapshots, sampled every 4-day interval. Overall there are about 63000 guild-snapshots. Guild features (composition, game activity, and structural) are computed for each guild-snapshots. Furthermore, we label the data samples as shrinking or stable guilds. If a guild will lose more than 15% of its membership in 4 weeks, the guild at the current time will be labeled as "shrinking", otherwise the guild is labeled as "stable". This simplifies the guild stability problem into binary classification. Thus we are trying to predict 4-weeks into the future as to whether a guild will remain stable or will shrivel.

Random sampling is used for drawing training samples. There are more shrinking guilds in the data than stable guilds, hence an uncontrolled random sampling may cause the classifier to overfit shrinking guilds. We control sampling to produced balanced classes, 2000 samples from each class. We would like to understand which features are important in modeling guild stability. Table 2 reports the correlation coefficient between each feature and the class labels (1 for shrinking, 0 for stable) for Eitrigg. Results for Bleeding Hollow & Cenarion Circle qualitatively agree with these results and hence we omit them for the sake of brevity.

- Among the guild composition features — Large guilds tend to be more stable. Guilds which have survived for longer periods tend to continue to survive. These agree with empirical observations and intuition. Average member level does not seem to matter much, however, diversity seems to play an

Category	Feature	Correlation coefficient
Composition	number of guild members	**-0.1213**
	length of existence	**-0.1501**
	average level of members	0.0037
	standard deviation of member levels	-0.0649
	percentage of character classes present	**-0.1153**
	entropy of character class distribution	-0.0707
	entropy of character category distribution	0.0087
Game stats	average playing time in guild	**-0.1038**
	average playing time prior to guild	0.0053
	average collaboration time in guild	**-0.1025**
	average collaboration time prior to guild	-0.0019
	average collaboration coefficient in guild	**-0.1026**
	average collaboration coefficient prior to guild	0.0067
Structural	average clustering coefficient	**-0.1021**
	average clustering coefficient in guild	**-0.1103**
	entropy of degree distribution	**-0.1288**

Table 2: Correlation coefficient between class labels and feature values. Correlation coefficients with absolute value exceeding 0.10 are marked in bold-face fonts.

important role. For instance, standard deviation of member levels are negatively correlated with guild shrinkage, indicating that diversity seems to help guild stability. Likewise, diversity in character classes is important. Guilds with more number of character classes present survive better.

- Among the game activity features — In-guild activity matters a lot. The more guild members collaborate and play, the more stable the guild is. The activity of guild members prior to joining the guild does not seem to matter at all.

- Among the structural features — All features seems to be very strong features. Balance and diversity in topology helps to improve overall guild stability.

Another common method for assessing the relative importance of features is the mutual information between a feature and the class label. This indicates how informative a feature is. We use Weka [16], a machine learning toolbox. It provides information gain computation and rank the features. In descending order of information gain, the top ten features are listed in Table 3. Compared to correlation coefficient analysis, the information gain ranking is more precise. It does not rely on single-mode distribution, which is an inherent limitation of correlation coefficient. However, the information gain does not reveal the insight that positive/negative correlation reveals. Qualitatively, Tables 2 and 3 are in rough agreement.

Rank	Feature	Category
1	number of members	C
2	entropy of class distribution	C
3	percentage of classes present	C
4	average collaboration time within guild	G
5	length of existence	C
6	average playing time within guild	G
7	entropy of degree distribution	S
8	standard deviation of member levels	C
9	average clustering coefficient within guild	S
10	average clustering coefficient	S

Table 3: Top ten features, ranked in descending order of information gain. In the category column, C stands for composition, G stands for game activity, and S stands for structural.

3.4 Predicting Guild Stability

Given the feature set and the class labels (stable or shrinking), we want to predict whether a group or community is likely to remain stable or will start shrinking over a period of time. We experiment with a range of supervised learning methods to achieve this.

With the feature set, we are able to predict guild stability with good accuracy. For instance, using guild size feature alone (number of members) and simple classification such as Naive Bayes, we can predict shrinking or stable labels with about 59% accuracy. Using two features for prediction, the number of members and length of existence, Naive Bayes produces a prediction accuracy of roughly 62%.

Table 4 summarizes the result of guild stability prediction using a variety of classification methods. The testing set is balanced, where an equal amount of testing samples from each class are randomly drawn from the observation data. The results are reported after a 10-fold cross validation process.

Classification methods include the following, with the first three methods serving as benchmarks.

- ZeroR baseline: It is a degenerated classifier, always predicting a shrinking guild regardless of the features.

- Naive Bayes: It assumes that all features are independent given the class label, and constructs a probabilistic model for each feature separately. The classifier computes likelihood from all features and chooses the maximum likelihood class label as the classification result.

- Decision stump: It is an one-level decision tree making a prediction with just a single input feature. In our training data, the single feature is average collaboration time within guild. If it is more than 5.002 hours, the guild is predicted to be stable. Despite its simplicity, the prediction accuracy is decent, in the $60 - 70\%$ range.

Method	Classification Accuracy	Precision	Recall	F-measure
ZeroR baseline	50%	0.25	0.5	0.333
Naive Bayes	63.74%	0.682	0.637	0.614
Decision Stump	62.10%	0.651	0.621	0.601
J48 decision tree	78.86%	0.790	0.789	0.789
Bagging	81.98%	0.822	0.821	0.820
Random Forest	84.78%	0.848	0.848	0.848

Table 4: Guild Stability Prediction Results

- J48 tree [22]: It progressively grows a decision tree by identifying the attribute that discriminates the training set most clearly according to an information gain criterion. The tree branch terminates if the training samples at the leave are homogeneous. The prediction accuracy of J48 tree is close to 79%.

- Bagging [7]: Bagging is an ensemble method which improves the classification accuracy through sampling and model averaging. We get an accuracy of close to 82% using Bagging.

- Random Forest [8]: Similar to bagging, random forest is also an ensemble method. It builds a library of decision trees from a set of random samples. Each decision tree is grown by randomly choosing the variables to split data upon. The classifier predicts class label by average voting from the decision trees. This method works well when there is sufficient training data. The accuracy is around 85%.

One hypothesis regarding guild stability is the continuity — if a guild has been shrinking recently, it is anticipated to continue the loosing streak. This hypothesis has been raised in the literature of social network analysis. To validate this hypothesis, we added an additional feature to capture the temporal aspect, i.e., the difference between the guild size in the current snapshot and the previous one. Positive value indicates a growing guild, while negative value indicates a shrinking guild. We have computed this temporal feature for all guild-snapshots. The correlation coefficient with the class label is -0.0382. The negative correlation is expected. Correlation appears very mild, indicating that past history is not a strong indicator of future trend. Furthermore, including this feature in the feature set for classification does not improve accuracy either. Essentially guild stability can be predicted from the guild features listed above (composition, activity, and structure), and temporal continuity seems to provide little additional information. Table 4 gives the detailed prediction results; it can clearly seen that we are able to achieve high accuracy (85%) in predicting guild stability on the Eitrigg server. Similar analysis is performed on Bleeding Hollow, a player-vs-player (PvP) server, and Cenarion Circle, a role playing (RP) server. We achieve qualitatively similar results for the other two servers; for instance, the random forest classifier produces a prediction accuracy of 81% on Bleeding Hollow, and 84.3% on Cenarion Circle.

4. ANALYSIS ON DBLP DATA

4.1 Dataset

DBLP [1], our second dataset provides bibliographic information on major computer science journals and conferences. Each publication is accompanied by its title, list of authors and conference/journal of publication. For the purposes of our study, we view DBLP as a social network of researchers who co-author papers at different conferences or in different journals. Thus the data resembles the social structure of our WoW dataset; the friendship network is defined by linking people that have co-authored a paper and the conferences/journals serve as groups where these friendships are formed. Table 5 show the size of our data and the network that we construct from the raw data. An important distinction between the two datasets is the group membership requirement; in the DBLP network an author can often be a member of multiple groups at any given point in time though with varying commitments.

Statistic for DBLP	
Number of Publications	1,607,524
Number of Authors	1,105,457
Number of Conferences/Journals	7073
Number of Friendship Edges	7,367,343
Number of Membership Edges	5,084,657

Table 5: Overall Network Statistics for DBLP

4.2 Gauging Group Stability

As mentioned earlier there is no concept of an author quitting a group in the DBLP dataset; on the contrary an author is a typically a member of several groups. This is a more commonly occuring scenario in most online social networks as compared to the group membership dynamics of World of Warcraft. The DBLP dataset is also used in [6] to study the formation and evolution of groups. Due to the lack of an explicit quitting action most studies have focussed on evolution from the standpoint of growth. We take a different approach when tackling datasets such as DBLP. Even though an author can belong to multiple groups his activities in individual groups can vary significantly over the course of time. Thus we need a measure that can quantify the involvement of a person in a community. We believe such a measure should encapsulate the following properties,

- A person that contributes frequently to a group should have a higher involvement score as opposed to a person that contributes rarely.

- Recent involvement/activities in a community should be weighted higher than past activities in the community.

- The prolificness of the author should be reflected in the measure of involvement.

Since we have timestamped (year of publication) data detailing the activities (publications) of a person (author) in a community (conference/journal), we can define a measure that reflects all of the above properties by adapting the exponential summarization kernel described in [23]. Let $N_1, N_2 \ldots N_t$ denote the number of publications of an author in a given group at discrete time intervals $t_1, t_2 \ldots t_t$ & $P_{A,t}$ denote the standing/prolificness of the author at time t, the "Involvement Score" of the author A in the given group

G at time t is defined as,

$$I_{A,G,t} = \begin{cases} (1-\theta)I_{A,G,t-1} + \theta N_t P_{A,t} & \text{if } t > t_0 \\ \theta N_t P_{A,t} & \text{if } t = t_0 \end{cases} \quad (2)$$

where t_0 is defined as the initial time and θ controls the rate of decay. The prolificness of an author can be defined in several ways; we define it as the ratio of the total number of publications the author has at time t to the total number of publications the most prolific author has at time t. Prolificness ranges between $[0, 1]$ and serves as a way of determining standing of the author. As mentioned before, the standing can be computed in different ways, for example, total citation count being another effective measure of calculating prolificness. However, the DBLP dataset has no way of determining citation counts, experimenting with other prolificness measures is out of scope of this paper. Table 6 demonstrates the use of the "Involvement Score" measure that we define to uncover trends in publications for prominent authors.

Year	Top-3 Conferences		
1996	STOC	FOCS	SODA
1998	STOC	DM&KD	VLDB
2000	STOC	FOCS	JComputing
2002	JCSS	STOC	JACM
2004	FOCS	JACM	STOC
2006	FOCS	KDD	IPSN
2008	JComputing	KDD	EC
2011	ICWSM	WWW	FOCS
2012	WWW	ICWSM	WSDM

Table 6: Top-3 conferences based on Involvement Score for Jon Kleinberg. One can clearly see a change in the trend, from publishing in Theory conferences (yellow) to publishing in Data Mining conferences (green).

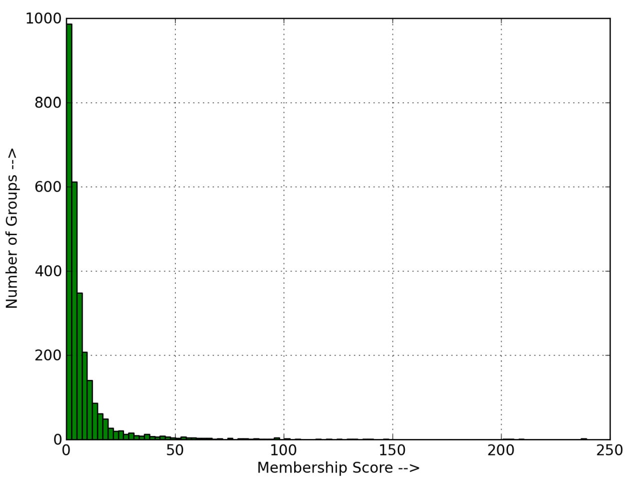

Figure 1: Histogram of membership scores for groups in 2011.

Now that we have defined a measure in Equation 2 to quantify the involvement of a person in a community, we proceed to define "Membership Score" of a group G at time t as the sum of involvement scores for all it's members. Formally,

$$MS_{G,t} = \sum_{A \in G} I_{A,G,t} \quad (3)$$

The membership score that we define has the following desirable properties,

- A group with more number of regularly contributing members has a higher membership score as compared to a group with large number of infrequently active members.
- A group that has more members of repute/standing will have a higher membership score as compared to a group with fewer prolific members.

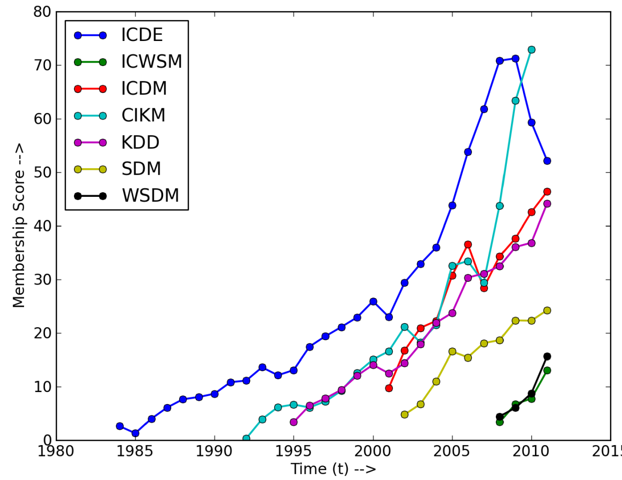

Figure 2: Membership score across time for well-known Data Mining Conferences.

Conference	Publications	H-index	Membership Score
ICDE	1303	35	52.15
KDD	670	30	44.20
CIKM	1348	26	95.81
ICDM	1197	18	46.39
SDM	338	18	24.28
ICWSM	221	18	13.06
WSDM	199	18	15.69

Table 7: Top-7 conferences in Data Mining with their Membership Scores

Figure 2 plots the membership score for 7 well-known conferences in the Data Mining area across the length of their existence. In order to test the efficacy of the membership score, we compute the membership score in 2011 for top-7 conferences in the Data Mining area in the last 5 years (as ranked by H-index [17] using Microsoft Academic Search [4]). Table 7 shows that the two measures are in rough agreement with each other. It is important to point out that we do NOT intend to advocate the membership score as a replacement for H-index (and such related measures). The membership score is able to capture group dynamics and hence can be used to gauge group stability. We compute the membership score for a group at regular time intervals. Thus computing the percentage increase/decrease in the membership score over the previous interval would then give us an accurate idea of whether a group is stable or shrinking. Given that a group G has membership score of ms_1 at time snapshot t_1 and membership score of ms_2 at time snapshot t_2 ($t_1 < t_2$) the percentage change in membership score is defined as $\delta = \frac{ms_2}{ms_1} - 1$. We could then label a group as being in the stable or

Category	Feature	Correlation coefficient
Group	number of group members	**-0.3008**
	length of existence	0.0499
	membership score	**-0.1948**
	average prolificness	**-0.5443**
Activities	average number of collaborations within group	**-0.6202**
	average number of collaborations outside group	**-0.4874**
	total number of collaborations within group	**-0.2293**
	total number of collaborations outside group	**-0.2687**
	average number of publications within group	**-0.7245**
	average number of publications outside group	**-0.5216**
	total number of publications within group	**-0.2670**
	total number of publications outside group	**-0.2732**
	average member loyalty coefficient	**-0.6114**
Structural	average clustering coefficient	**-0.6477**
	average clustering coefficient in group	**-0.6705**
	entropy of degree distribution	**-0.7144**

Table 8: Correlation coefficient between class labels and feature values. Correlation coefficients with absolute value exceeding 0.10 are marked in bold-face fonts.

shrinking phase at time t_2 as defined in equation 1. In our experiments we compute the membership scores at yearly intervals (i.e. $t_2 - t_1 = 1yr$).

4.3 Conference-Level Features

Rank	Feature	Category
1	total number of publications within group	A
2	number of members	G
3	total number of collaborations within group	A
4	average prolificness	G
5	average number of publications within group	A
6	total number of publications outside group	A
7	average number of collaborations within group	A
8	total number of collaborations outside group	A
9	average member loyalty coefficient	A
10	entropy of degree distribution	S

Table 9: Top ten features, ranked in descending order of information gain. In the category column, G stands for group-specific, A stands for activity features, and S stands for structural.

In order to model group stability for the DBLP dataset we consider a range of features that can be broadly classified into three categories: (1) conference/group-specific, (2) publications/activities-specific, and (3) structural features. The following is a list of group specific features,

- Number of members: The size of the group at time t.
- Length of existence (in years): This feature calculates the number of years since the conference/journal came into existence.
- Membership Score: Membership Score of group at time t as defined in 3.
- Average Prolificness: Average Prolificness of group at time t, where prolificness is a measure of standing/repute for an individual. It ranges between $[0, 1]$.

We compute the following list of features to capture the activities of members in a particular group,

- Total & Average Number of Collaborations Within & Outside Group: These features capture the number of collaborations involving the group members. These collaborations can be within the given group or in some other groups.
- Total & Average Number of Publications Within & Outside Group: These features capture the number of publications for group members. Again an individual may have publications within and outside the given group.
- Average Member Loyalty Coefficient: Loyalty Coefficient for a group member is defined as the ratio of the number of publications that member has in the given group to the overall number of publications of the member. It ranges between $[0, 1]$ and is a measure of the loyalty of the member towards a particular group he/she is a member of.

Following is a list of features intended to capture the connectivity information of a group,

- Average clustering coefficient of group members: We measure the clustering coefficient at each group member node and then calculate the average of this clustering coefficient across all group members.
- Average clustering coefficient of group members within the group: The clustering coefficient calculated in this case only takes into consideration the graph induced by all members of the guild.
- Entropy of degree distribution: This feature is a good measure of diversity in node connectivity.

4.3.1 *Feature Importance & Correlation*

We perform similar analysis as performed on the WoW dataset. We compute features along with class labels (1 for shrinking and 0 for stable) at yearly intervals since most conferences/journals have

Class	Classification Accuracy	Precision	Recall	F-measure
Stable		0.878	0.942	0.909
Shrinking	90.55%	0.937	0.869	0.902
Weighted Average		0.908	0.906	0.905

Table 10: Group Stability Prediction Results using Bagging

an yearly cycle of publication. Thus, we are trying to predict one year into the future as to whether a group will remain stable or not. This results in around $40,000$ feature samples; 22.51% of these samples have "shrinking" class labels & 77.49% of the samples have "stable" class labels. In order to avoid overfitting we draw equal number of samples from both classes. Table 8 reports the correlation coefficient between each feature and the class labels (1 for shrinking, 0 for stable) for the DBLP dataset.

Assessing the importance of features by computing the Information Gain, the top ten features are listed in table 9. Tables 8 and 9 are in general agreement about the important features required for the predictiont task.

4.4 Predicting Group Stability

Again, we will use supervised learning techniques and apply them to our feature set to see if we can predict group stability. Due to the unbalanced class problem, we randomly draw equal number of samples from both the classes (≈ 9000 samples per class). Table 10 shows the accuracy achieved by using Bagging (we achieve similar accuracy levels by using Decision Trees and Random Forests) . Bagging achieves the best accuracy of 90.55% with a MAE of 0.1402 and a MSE of 0.2601; proof of the fact that our feature based approach produces significantly high accuracy in predicting group stability.

5. INTERNAL CONNECTEDNESS OF FRIENDS

The study of Backstrom et al [6] is amongst the first to comprehensively analyse evolution of groups using real-world social networking data. They demonstrated that the probability of joining increases as the density of linkage increases among the individual's friends in the community. These results are supported by arguments based on social capital [11, 12] that suggest that there is a trust advantage to having friends in a community who know each other. An individual joining such a community is assured of the fact that such a community is a close-knit family of members who know most of the other members.

At the same time Backstrom et al. pointed out that cogent arguments [15, 9] also support the opposite finding; this theory based on weak ties suggested that there is an informational advantage to having loosely connected members. This provides an individual multiple "independent" perspectives; he/she could join based on any one of the ways.

Empirical evidence based on the Live Journal [3] dataset used by Backstrom et al. made them conclude that trust advantage had a stronger effect than informational advantage. Kairam et al. [18] shed further light on the group evolution and growth process. They too touched upon this problem; empirically they came to the same conclusion i.e. probability of joining increases as the density of linkage increases among the individual's friends in the community. They also tried to solve the paradoxical finding of why highly clustered groups tend to have lower growth rates overall. Their findings suggested that some groups grow by appealing to common interests and identities (non-diffusion growth) while other groups grow by virtue of its extra-group connections (diffusion growth). Furthermore they conclude that if a group relies on diffusion growth its scope for growth is limited to the number of ties its members have to non-members. Thus such groups will eventually suffer from lack of new members. Thus, even though high clustering in a group will lead to increased membership it will also lead to diminishing returns (with respect to growth) down the road. In their findings (based on the Ning [5] dataset), they are able to show that groups that grow to small sizes are those that rely on diffusion growth whereas groups that grow to large sizes are those that rely more on non-diffusion growth.

We try to validate the theories and findings put forward in [6, 18] using our datasets as follows,

- WoW Dataset: We compute the correlation between the features "average clustering coefficient in guild" and "number of new members". The "average clustering coefficient in guild" allows us to quantify the density of linkage amongst a guild's members. The correlation coefficient is -0.0584 i.e. weakly negatively correlated which tends to suggest support for informational advantage. This is an interesting finding which hasn't been observed in previous studies. We also compute the correlation between "average clustering coefficient in guild" and "percentage change in number of members over the previous snapshot". This correlation comes in at -0.00959 which indicates that density of linkage does not play any role in determining guild growth.
- DBLP Dataset: Again we compute the correlation between the features "average clustering coefficient in guild" and "number of newly active members". The value of correlation is 0.2530 which shows support for trust advantage over informational advantage. The correlation between "average clustering coefficient in guild" and "percentage change in membership score over the previous snapshot" turns out to be 0.0061 indicating again that density of linkage does not play any role in determining guild growth.

Our findings indicate as far as WoW data is concerned, individuals join guilds due to common interests and identities; thus guilds in WoW are characterized by non-diffusion growth. On the other hand in the DBLP data most of the growth can be characterized as diffusion based growth. These findings are also due to the nature of the social networks. WoW is a multiplayer game where individuals work towards an objective of being successful at playing the game. Gamers are likely to join guilds based on common objectives rather than based on trust factors. On the other hand DBLP data is a co-authorship network where edges indicate collaborations at a particular conference or in a given journal. Thus in this case links amongst an individual's friends indicates stronger endorsement for that group from your peers.

6. CONCLUSION

Our analysis has shown that it is possible to predict group stability with high accuracy using a range of features that describes the group composition, activities within the group & structural aspects of a group. We have experimented with two large social networking datasets and have been able to achieve similar accuracy levels on both datasets. We have also defined an efficient measure of gauging group membership in scenarios where a person is likely a member of several groups. Our analysis can easily be extended to other online social networks and is also scalable to large networks. The study also shows that it is important to choose features from multiple perspectives, in fact combining diverse features is essential to predictor performance.

Acknowledgements

This research is funded in part by DARPA/ADAMS program under contract W911NF-11-C-0216. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the government funding agencies. We would also like to acknowledge contributions from our colleagues: Nicholas Ducheneaut, Nick Yee, and Oliver Brdiczka (PARC) for providing WoW game census data and insights about game dynamics, and Hossam Sharara (University of Maryland, College Park) for help on processing WoW data and social network analysis.

7. REFERENCES

[1] The dblp computer science bibliography. `http://dblp.uni-trier.de`.

[2] Facebook. `https://www.facebook.com`.

[3] Live journal. `http://www.livejournal.com`.

[4] Microsoft academic search. `http://academic.research.microsoft.com`.

[5] Ning. `http://www.ning.com`.

[6] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '06, pages 44–54, New York, NY, USA, 2006. ACM.

[7] L. Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

[8] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[9] R. Burt. *Structural holes: The social structure of competition*. Harvard University Press, 1995.

[10] D. Centola and M. Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734, Nov. 2007.

[11] J. Coleman. Social capital in the creation of human capital. *American journal of sociology*, pages 95–120, 1988.

[12] J. Coleman. *Foundations of social theory*. Belknap Press, 1994.

[13] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore. The life and death of online gaming communities: a look at guilds in world of warcraft. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 839–848, New York, NY, USA, 2007. ACM.

[14] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.

[15] M. Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.

[16] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[17] J. Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United states of America*, 102(46):16569, 2005.

[18] S. Kairam, D. Wang, and J. Leskovec. The life and death of online groups: Predicting group growth and longevity. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 673–682. ACM, 2012.

[19] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, IMC '07, pages 29–42, New York, NY, USA, 2007. ACM.

[20] A. Patil, J. Liu, B. Price, H. Sharara, and O. Brdiczka. Modeling destructive group dynamics in on-line gaming communities. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, ICWSM '12, Dublin, Ireland, June 2012.

[21] M. Porter, J. Onnela, and P. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.

[22] J. R. Quinlan. *C4. 5: programs for machine learning*, volume 1. Morgan kaufmann, 1993.

[23] U. Sharan and J. Neville. Temporal-relational classifiers for prediction in evolving domains. In *ICDM'08. Eighth IEEE International Conference on Data Mining, 2008*, pages 540–549. IEEE, 2008.

[24] J. Ugander, L. Backstrom, C. Marlow, and J. Kleinberg. Structural diversity in social contagion. *Proc. National Academy of Sciences*, 109(16):5962–5966, April 2012.

[25] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.

[26] D. Watts and S. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393(6684):440–442, 1998.

[27] E. Zheleva, H. Sharara, and L. Getoor. Co-evolution of social and affiliation networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, pages 1007–1016, New York, NY, USA, 2009. ACM.