

Understanding Actor Loyalty to Event-Based Groups in Affiliation Networks*

Hossam Sharara
hossam@cs.umd.edu

April 26, 2010

Abstract

In this paper, we introduce a method for analyzing the temporal dynamics of affiliation networks. We define affiliation groups which describe temporally related subsets of actors and describe an approach for exploring changing memberships in these affiliation groups over time. To model the dynamic behavior in these networks, we consider the concept of loyalty and introduce a measure that captures an actor's loyalty to an affiliation group as the degree of 'commitment' an actor shows to the group over time. We evaluate our measure using three real world affiliation networks: a publication network, a senate bill co-sponsorship network and a dolphin network. The results show the utility of our measure for analyzing the dynamic behavior of actors and quantifying their loyalty to different time-varying affiliation groups.

1 Introduction

Across many fields, researchers are interested in understanding an individual's commitment to a group (e.g., [1]), the social structure of groups (e.g., [2]), and the changing dynamics of group structure (e.g., [3]). In marketing, researchers investigate customer behavior, comparing the purchasing behavior of different customer groups in an attempt to determine customer satisfaction and brand loyalty (e.g., [4]). In sociology, researchers investigate commitment (e.g., [1]), community cohesion (e.g., [5]) and structural embeddedness of social groups (e.g., [6]). In computer science, researchers

*This work is based on a paper submitted to the Journal of Advances in Social Networks Analysis and Mining, co-authored with Lisa Singh, Lise Getoor, and Janet Mann.

have also modeled time-varying links to improve automatic discovery of relational communities or groups (e.g., [7, 8]). While some statistical models have been developed for longitudinal analysis of social networks (see Snijders [9] for an overview), work remains to better understand the variation in actor commitment or loyalty to groups over time.

Social psychologists have investigated the role played by feelings of loyalty to groups. Druckman explains that “loyalty to a group strengthens one’s identity and sense of belonging” [10]. In this paper, we focus on an operational definition of loyalty to affiliation groups in an attempt to adequately measure this ubiquitous idea. Consistent with sociology literature [6], we believe that high loyalty may be an indicator of group cohesion.

More specifically, we investigate actor loyalty to groups in two-mode affiliation networks. A two-mode affiliation network contains two different types of nodes, one for actors and one for events. Edges between actor nodes and event nodes are used to indicate relationships between actors and events in which the actors participate [11]. Affiliation networks capture a wide variety of interesting domains, including communication data (email, cell phone calls, etc.) among people; organizational data describing peoples’ roles on teams or in companies; and epidemiological networks, describing people and the specific disease strain with which they are infected. In time-varying affiliation networks, an actor’s participation in a particular event is associated with a specific time, representing when this participation occurred. Annotating affiliation networks with temporal information allows us to capture changing actor behavior over time. In this paper, we focus on this changing behavior as it relates to groups.

Consider an author/publication network describing authors, with the publications represented as events in which the co-authors are participants. If the publications are annotated with topic areas, then we can create groups of actors who publish in the same topic area at the same time. Furthermore, we can see how loyal an author is to specific topic areas over time by examining their changing publication topics. One common scenario is that an author starts publishing in a specific area, then after some time s/he begins publishing in additional areas, and eventually s/he might end up switching areas completely. Another common scenario is that an author starts publishing in an area, and, rather than adding additional areas, remains steadfast, and continues publishing regularly in the same area over a long period of time. We introduce a measure that captures this dynamic behavior of actors in time-varying affiliation networks by introducing the concept of **affiliation group loyalty** and define an actor’s loyalty to an affiliation group as the degree of ‘commitment’ an actor shows to the group over time.

This paper extends our earlier work [12] by formalizing the concept using related theory and background from different disciplines, adding experimentation on an additional real world data set, and including extensive new experiments that compare loyalty across three real world affiliation networks, as well as to different centrality measures. The remainder of this paper is organized as follows. We provide background on loyalty and concepts related to loyalty in the next section. Section 3 presents our grouping abstraction. Section 4 defines a new measure for actor loyalty to affiliation groups and compares it to other possible measures. An analysis of actor loyalty to affiliation groups using real world affiliation networks is presented in Section 5. Section 6 compares loyalty across the three real world affiliation networks to illustrate the amount of variation in loyalty across domains. Finally, we conclude with future directions in Section 7.

2 Loyalty Background

Within literature across different disciplines, terms like *loyalty*, *commitment*, and *cohesion* have been given a number of different theoretical and operational definitions. Our goal is not to provide an exhaustive literature review on these subjects, but rather to give a context for the remainder of our discussion on loyalty in affiliation networks.

Sociologists first formalized commitment as a way to link extraneous interests with a consistent line of activity ([13]). While other definitions and theories concerning commitment exist in sociology ([1]) and social psychology ([14]), a definition proposed by ([15]) is as follows:

Commitment is a force that binds an individual to a course of action that is of relevance to a particular target.

Loyalty extends the concept of commitment. For example, [4] defines customer loyalty in terms of brand commitment (the strength of the relationship between customers and a particular brand), and gives the following multi-faceted definition:

Loyalty is a deeply held commitment to rebuy or repatronize a preferred product/service consistently in the future, thereby causing repetitive same-brand or same brand-set purchasing, despite situational influences and marketing efforts having the potential to cause switching behavior.

A well known operational definition of loyalty in business literature defines brand loyalty as the percent of purchases devoted to one's most often

purchased brand [16]. Newman et al. [17] define loyal customers as those repurchasing a brand considering only that brand, while Tellis [18] explains that loyalty has been defined as repeat purchasing frequency or relative volume of same brand purchasing. Jacoby et al. [19] state that frequent purchasing of a product is not synonymous with brand loyalty and that the notion of commitment is essential for distinguishing between brand loyalty and frequent purchasing of a product.

While business research tends to focus on the economic component dimension, based on purchasing behavior, social psychologists have investigated ways that people relate to groups. One dimension of this is the role played by feelings of loyalty to groups. Druckman [10] explains:

The feelings of attachment that comprise loyalty are not whimsical, but are generally basic to the individual's definitions of themselves. Loyalty to a group strengthens one's identity and sense of belonging.

As will be discussed later, we define a group in terms of related events. We focus on an operational dimension of loyalty to affiliation groups in an attempt to adequately measure a ubiquitous idea. We demonstrate effects of our operational definition as it compares to frequency based brand loyalty in the business literature. Consistent with sociology literature ([6]), we believe that a group containing actors with high loyalty may be an indicator of a highly cohesive group.

Because of the size and complexity of social networks, computer scientists have also begun investigating different aspects of social networks. The community detection literature uses measures of cohesion and clustering to find subsets of actors that are densely connected to each other, but less densely connected to others. The majority of research conducted on community detection focuses on static networks and constrains the problem by letting an actor belong to only a single community [20, 21, 22, 23, 24, 25].

Recently, researchers have begun to analyze the dynamics of communities over time [26, 27, 28, 29, 30, 31, 32]. Much of this research focuses on two questions: what are the communities that exist in a particular data set, and how do they change or evolve over time. In contrast, the approach that we propose is a more micro-level analysis that focuses on the dynamics of specific actors or individuals in the network. While we focus on creating groups using affiliation event attributes (as will be described in the next section), our analysis of actors can be conducted using the output from any grouping, clustering, or community detection algorithm. Once the social

groups are established, our goal is to understand the dynamics of actors and their social relationships in the context of these predefined social groups.

One approach which also proposes methods for identifying important actors in dynamic networks is the work of Habiba et al. [33]. They identify nodes in a single mode network that are likely to be good spread blockers. To accomplish this, they introduce dynamic measures for density, diameter, degree, betweenness, closeness and clustering coefficient.

The graph summarization method proposed in [34] also uses a measure similar to ours to build a classifier for predicting evolving domains. While both of our measures attempt to quantify temporal aspects of the network, there are differences between their work and ours. First, the graph summarization method is used to create an aggregation of network snapshots over time by weighting the edges according to the point in time in which they occur. In contrast, our loyalty measure is used to quantify actor’s participation pattern in different affiliation groups. Second, though the authors in [34] mention that their proposed weighting kernels are able to model both temporal recurrence and temporal locality, which represents the aspects of consistency and recency we are addressing in our loyalty measure, it is unclear how the weighting kernels used account for temporal recurrence. In contrast, the recursive formulation of our proposed loyalty measure encodes this aspect directly.

3 Modeling time-varying event-based groups

There are many ways to define a group. Groups can be formed using community detection algorithms, clustering algorithms, etc. Because we are interested in understanding groups based on affiliation networks, we describe an approach that defines groups based on a participation relationship between actors and events. Formally, an affiliation network can be represented as a bi-partite graph $\mathcal{G}(\mathcal{A}, \mathcal{E}, \mathcal{P})$ containing a set of actor nodes \mathcal{A} , a set of event nodes \mathcal{E} , and a set of participation edges \mathcal{P} that connect actors in \mathcal{A} to events in \mathcal{E} :

$$\begin{aligned}\mathcal{A} &= \{a_1, a_2, a_3, \dots, a_n\} \\ \mathcal{E} &= \{e_1, e_2, e_3, \dots, e_m\}, \text{ and} \\ \mathcal{P} &= \{(a_i, e_j) | a_i \in \mathcal{A}, e_j \in \mathcal{E}\}.\end{aligned}$$

We denote participation of actor a_i in event e_j as $p_{i,j}$. For clarity, we use a running example of an author publication network in which the actors are authors, the events are publications, and the participation relation is paper

authorship. Figure 1 shows an example network with three author nodes, $\mathcal{A} = \{a_1, a_2, a_3\}$, fifteen publication nodes, $\mathcal{E} = \{e_1, e_2, \dots, e_{15}\}$, and twenty paper authorship edges. As an example, participations involving actor a_1 are the following:

$$\mathcal{P}_{a_1} = \{p_{1,1}, p_{1,2}, p_{1,3}, p_{1,4}, p_{1,5}, p_{1,7}, p_{1,8}, p_{1,9}, p_{1,10}, p_{1,11}, p_{1,13}, p_{1,14}, p_{1,15}\}.$$

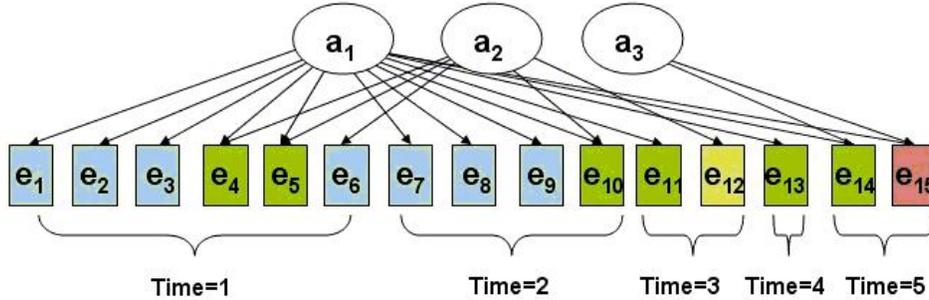


Figure 1: An affiliation network example with 3 actors, 15 events and 20 relationships across 5 time points.

Each actor node and event node can have attributes associated with them. For example, each author in Figure 1 may have a name and an age. For author a_1 , we may have the following attribute values $a_1.name = \text{'Peter Pan'}$ and $a_1.age = 50$. Each publication event may have a title attribute, e.g. $e_1.title = \text{'Static networks as non-evolving dynamic networks'}$ and a topic attribute, $e_1.topic = \text{'social networks'}$. In Figure 1, we use shading to indicate topic. Since e_1 is shaded blue, all the events shaded blue have the same value for topic, e.g. 'social networks'. For ease of exposition, we map each color to the following topics: blue - $topic_1$, green - $topic_2$, red - $topic_3$, yellow - $topic_4$.

Because our affiliation networks are temporal, a time point attribute $time$ is associated with each event e_j , and is denoted as $e_j.time$. For affiliation networks, this time is the same as the time of the participation relationship. In our example, the time attribute is the date of publication. We have labeled the time point associated with each event in Figure 1.

While the publication event serves as a grouping of a subset of actors, this event only occurs at one particular time. Because our goal is to understand the dynamics of affiliation networks over time, we are interested in analyzing actor participation in groupings of similar events across time. We propose grouping events based on values of an event attribute. In other

words, a social group is defined based on a *shared event attribute value*. The choice of a specific method for grouping actors depends on the semantics of the underlying analysis task. Using shared event attributes is particularly meaningful for affiliation networks since it incorporates the semantics of events into the data model. For other types of social networks, particularly uni-mode networks, it is reasonable to use other methods for defining social groups.

Each event feature or attribute F has an associated domain $Domain = \{g_1, g_2, \dots, g_p\}$, where p is the number of distinct values of F . We denote a particular value g_l of an event e_j for event attribute F as $e_j.F = g_l$. Based on this, we define an affiliation group to be a subset of actors having the same group value g_l at time t for an event e_j : $G(g_l, t) = \{a_i | a_i \in \mathcal{A}, (a_i, e_j) \in \mathcal{P}, \text{ where } e_j.F = g_l \text{ and } e_j.time = t\}$. In our example, suppose our grouping attribute is *topic*. Referring back to Figure 1, $G(topic_1, 1) = \{a_1, a_2\}$ is the set of actors in topic group $topic_1$ at time 1.

We pause to mention a few advantages of our grouping formulation. First, actors can belong to multiple affiliation groups at a particular time. In other words, membership in different groups can be *overlapping*. In our example, author a_1 participates in five events at time 1. Also, actors are not required to be part of an event (or group) at every time t . This is also illustrated in our example. Author a_1 participates in an event at every time step. Authors a_2 and a_3 do not. In our experience, these assumptions better capture the dynamics of real world affiliation networks.

4 Loyalty of Individuals to Affiliation Groups

In order to better understand the loyalty of an actor to groups based on event affiliation, we need to quantify the participation of an actor in different groups over time. Based on our example in Figure 1, Figure 2 shows actor a_1 's membership in topic groups, $topic_1$, $topic_2$, and $topic_3$ across five time steps. The rectangles represent different topic groups and an edge from the author to a topic means that an author has published on the linked topic. The count on the edge represents the number of publications an actor a_i has published on this topic during a particular time period. For example, the network snapshot of the first time period shows author a_1 having three publications with $topic_1$ and two publications with $topic_2$. As time continues, author a_1 stops publishing on $topic_1$, continues publishing on $topic_2$ at each time step, and begins publishing on $topic_3$ in the last time step. Intuitively, if we consider the loyalty of the author at time step 5, we would like to see

a higher loyalty score for $topic_2$ since the author has published in this topic since time step 1. At time step 2, a topic shift occurs from $topic_1$ to $topic_2$. Our goal is to create a measure that is sensitive to both continual group membership and changing group membership over time.

We begin by considering two simple loyalty measures: frequent participation and recent participation, illustrating how well they can capture the nuanced nature of loyalty in dynamic networks.

Loyalty based on frequent participation, which we refer to as *frequency-based loyalty* considers an actor loyal if s/he appears in a group frequently. Let $n(a_i, g_l)$ represent the number of participations of actor a_i in group g_l and $n(a_i, *)$ represent the number of participations of actor a_i in all groups. Then the frequency-based loyalty of actor a_i is defined as the number of participations in a particular group g_l divided by the number of participations across all groups:

$$Loy_{FP}(a_i, g_l) = \frac{n(a_i, g_l)}{n(a_i, *)}$$

Using our example, author a_1 publishes in $topic_1$ six times, $topic_2$ six times, and $topic_3$ one time. Therefore, $Loy_{FP}(a_1, topic_1) = Loy_{FP}(a_1, topic_2) = 6/13$ and $Loy_{FP}(a_1, topic_3) = 1/13$. $topic_1$ and $topic_2$ are considered equally important even though the author has not published in $topic_1$ since time step 2. Thus, considering frequency alone ignores the temporal component of the group affiliation and results in assigning higher loyalty values to groups that the actor was once active in, but may not be active in any longer. Frequency-based loyalty can be viewed as a static measure of commitment.

Focusing on the temporal aspect of the data, a *recency-based loyalty* measure considers an actor loyal if s/he has participated recently in a specific group. Let $n(a_i, g_l, t)$ represent the number of participations of actor a_i in group g_l at time step t . The recency-based loyalty of actor a_i is defined as the number of participations in a particular group g_l at the last time step t_f divided by the number of participations across all groups at time t_f :

$$Loy_{RP}(a_i, g_l) = \frac{n(a_i, g_l, t_f)}{n(a_i, *, t_f)}$$

In our example, when $t_f = 5$, $Loy_{RP}(a_1, topic_2) = Loy_{RP}(a_1, topic_3) = 1/2$ and $Loy_{RP}(a_1, topic_1) = 0$. Author a_1 is equally loyal to $topic_2$ and $topic_3$ even though $topic_3$ only appears in the current time step. If we consider the last two time steps (using a recent window as opposed to a recent time point), then a_1 is most loyal to $topic_2$. While this is accurate, the strong early participation of actor a_1 to $topic_1$ is not captured at all

since $Loy_{RP}(a_1, topic_1) = 0$. Using recent participation leads to assigning an actor high loyalty values for groups that the actor participates in during current time steps, but it disregards earlier participation.

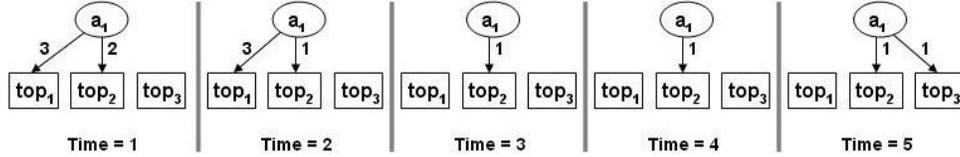


Figure 2: Single actor dynamic affiliation example

From this simple example, we see that a temporal measure of affiliation group loyalty should incorporate participation frequency for giving higher preference to actors with a large number of participations in the affiliation group, consistency for putting more bias toward actors with regular group participations across time over those with more sparse, isolated participation, and recency for favoring actors with current participations. In order to capture all of these aspects, we incorporate frequency and recency based loyalty into a more comprehensive measure of loyalty.

Let T_{tot} represent the total number of time points the dynamic affiliation is defined over. The loyalty of an actor to a group that s/he has not participated in yet is equal to zero. In order to keep track of consistent participation over time, we need to keep track of the actor's loyalty in the time step that precedes the current one. Thus, we define t_{prev} as the previous time point (relative to the current time point t) that actor a_i participated in group g_l . Let $n(a_i, g_l, \Delta t)$ be the number of participations of actor a_i in group g_l from the starting time point t_0 until the current time point t , and let $n(a_i, *, \Delta t)$ be the number of participations of actor a_i to all groups from t_0 until time t . We define the loyalty of an actor a_i to a group g_l on his first participation in it as

$$Loy(a_i, g_l, t_0) = \frac{n(a_i, g_l, t_0)}{n(a_i, *, t_0)}$$

where $t = t_0$ and the loyalty on any consecutive participation is given by

$$Loy(a_i, g_l, t) = \frac{n(a_i, g_l, \Delta t)}{n(a_i, *, \Delta t)} \times Loy(a_i, g_l, t_{prev})^\alpha \frac{t - t_{prev}}{T_{tot}}$$

where α represent a smoothing parameter that will be described shortly.

Examining the different components of the loyalty measure, we see that the first term, $\frac{n(a_i, g_l, \Delta t)}{n(a_i, *, \Delta t)}$, accounts for the frequency of participation of an actor into a specific group. The second term includes the component $Loy(a_i, g_l, t_{prev})$ which takes into consideration that latest recorded loyalty for an actor in a specific group, g_l , and is used to favor recent participation in that group. Finally, to favor continuous actor participation, the second term includes an exponent term for the recent loyalty. This decreases the effect that the loyalty in the previous time step has on the calculated loyalty in the current time step based on how long in the past this previous participation occurred. The more recent and continuous the participation, the larger the effect of this component on the overall loyalty of the actor to the group.

The smoothing parameter α is introduced to control the overall effect of time. The value of α can be varied from 0 to T_{tot} . A value of 0 means $Loy = Loy_{FP}$, focusing on the frequent participation component of the measure. A value of T_{tot} means that the recent participation component of the measure is dominant. For exploratory analysis, setting the value of α to 1 represent a good initial point to start off, where the loyalty accounts for both the frequency and the recency factors.

For consistency with the group membership notation, where the actor membership values in various groups sum up to 1, we normalize the values of loyalty of a specific actor to various groups that s/he participated in over the considered time period. As a result, the final loyalty value of actor a_i to group g_l at the final point in time t_f can be defined as follows

$$Loyalty(a_i, g_l, t_f) = \frac{Loy(a_i, g_l, t_f)}{\sum_j Loy(a_i, g_j, t_f)}$$

where the summation parameter j ranges over all the groups that actor a_i participated in during the entire time period.

Returning to our earlier example, we see that our proposed measure results in the desired effect. Setting the value of ($\alpha = 1$), the results for actor a_1 loyalty to different topics are as follows:

$$Loyalty(a_1, topic_1, t_5) = 0.429$$

$$Loyalty(a_1, topic_2, t_5) = 0.474$$

$$Loyalty(a_1, topic_3, t_5) = 0.097$$

The evolution of the author's loyalty for each topic at each time step with $\alpha = 1$ is illustrated in Figure 3. $topic_1$ begins with the highest loyalty

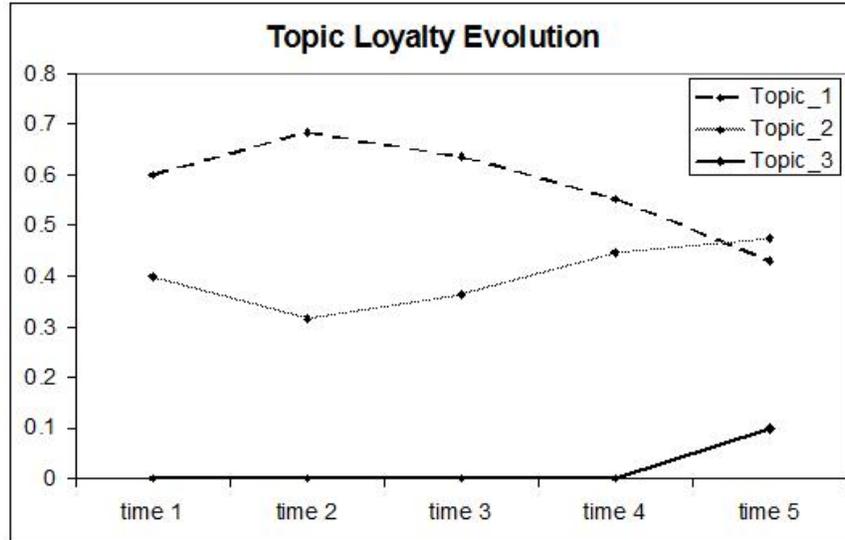


Figure 3: The evolution of loyalty over time for our simple example

at time 1. Its loyalty increases at time 2 and then begins to decline. After time 4, author a_1 's loyalty to topic $topic_2$ overtakes that of $topic_1$ because of the effect of recency.

To further illustrate the effect of the smoothing factor, Figure 4 shows the different values for the loyalty of the author to all the topics at the final time step by varying the value of α . When $\alpha = 0$, the loyalty values are the same as if we consider only (normalized) frequency, Loy_{FP} . As the value of alpha increases, we can see the effect of recency starting to dominate the frequency. At the maximum value of ($\alpha = 5$), we see that the highest loyalty score is for $topic_3$ (which corresponds to the most recent group).

The interpretation here is that as the value of α increases, the measure favors new topics occurring at the last time point. Therefore, $topic_3$, which has just occurred at the last time point, dominates all other topics because its loyalty is not decreased by previous occurrences. The loyalty values of $topic_1$ and $topic_2$ are overwhelmed by the large exponential factor resulting from the large α .

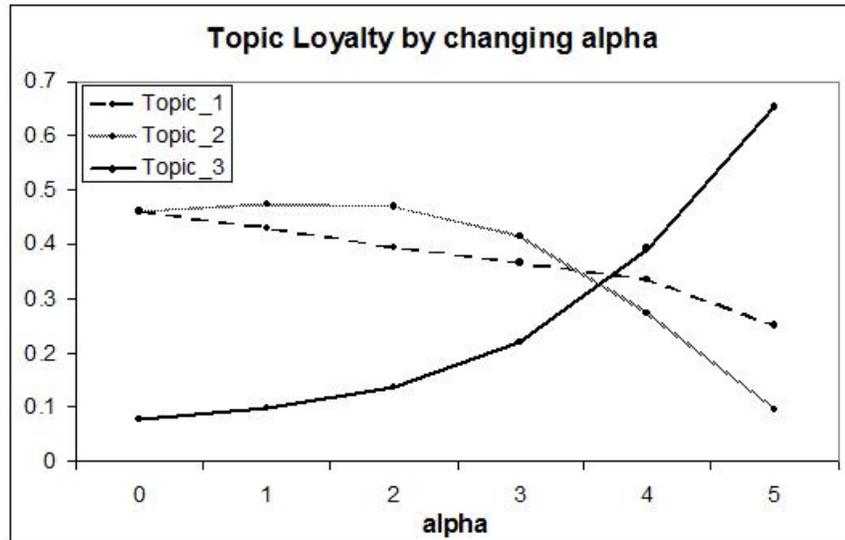


Figure 4: The effect of the smoothing factor in calculating group loyalty

5 Loyalty Analysis on Individual Data Sets

We analyze our proposed loyalty measure on three data sets - a scientific publication network, a senate bill sponsorship network and a dolphin social network. In order to consider frequency, consistency, and recency, we set $\alpha = 1$.

5.1 Scientific Publication Network

The scientific publication network is based on publications in the ACM Computer-Human Interaction (ACMCHI) conference from 1982 until 2004. Similar to our running example, this data set describes an author/publication affiliation network. The data set was extracted from the ACM Digital Library and contains 4,073 publications and 6,358 authors. There are 12,727 participation relationships (edges) between authors and publications. In this data set, we filtered 5230 authors having only one publication over the entire period of time since no 'dynamic' group loyalty exists for these actors. Also, by removing them, we avoid biasing the average loyalty statistics calculated for the data set. The remaining 1,128 authors had 4,688 relationships with publication events.

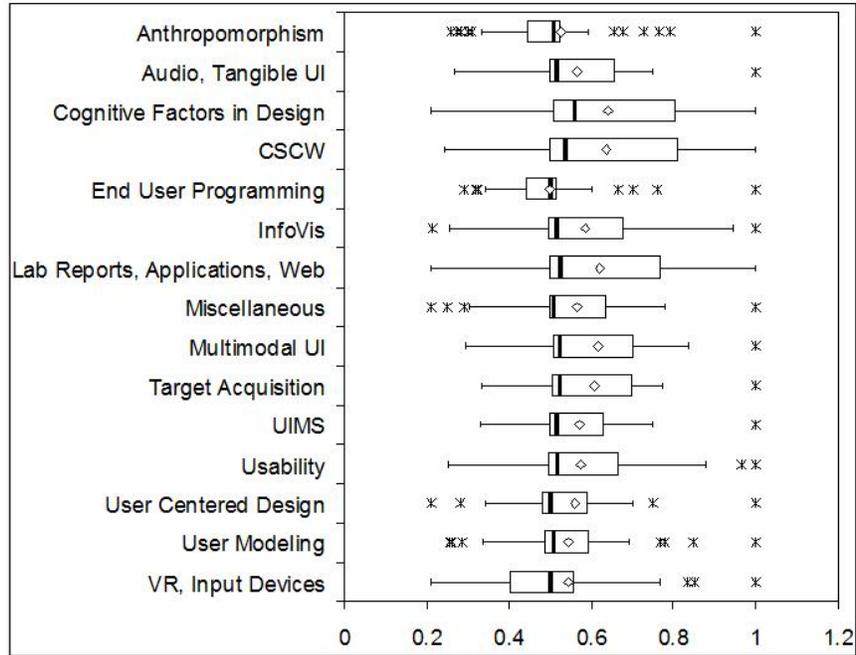


Figure 5: The average topic loyalty for the scientific publication network

There are a number of features that the publication events can be grouped on; for this analysis, like the simple running example, publications were grouped by their topic. There are 15 different topic values, and the loyalty of authors to different topics was measured. The results of applying the proposed loyalty measure on the ACMCHI data set are shown in Figure 5. The topic loyalty of authors range from 0.2 to 1, while the average topic loyalty ranges from 0.5 to 0.65 for all 15 topics. While there are a number of interesting observations to be made, we highlight two of them. First, the average topic loyalty is fairly uniform across the topics. This is an indication of the continued importance of these topics at the ACMCHI conference. Second, the average loyalty of authors to topic groups is very high across all the topics. This is an indication that, in general, authors in this data set consistently published in a particular research area as opposed to oscillating among multiple areas.

To better understand the distribution of author loyalty as it relates to an author’s employer type, Figure 6 shows the average loyalty of authors categorized by employer type (corporate institutions, universities, research

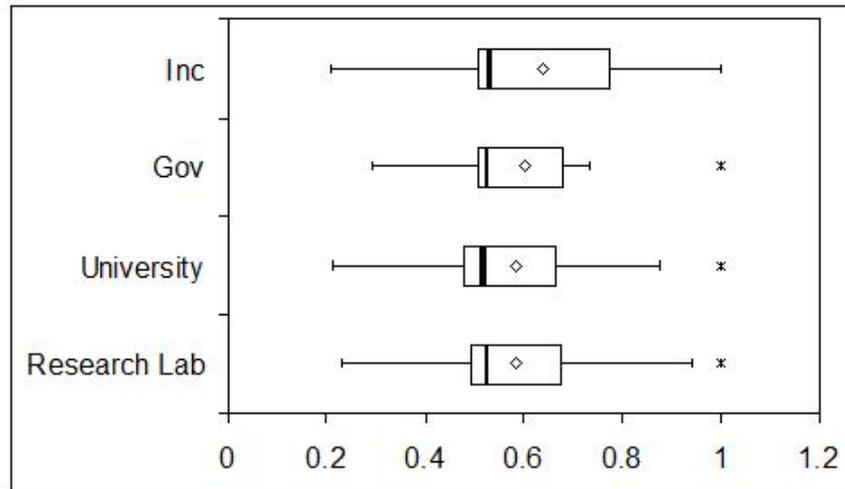


Figure 6: The average topic loyalty grouped by institution type for the scientific publication network.

laboratories, and government). One interesting result is that authors from corporate institutions, i.e. Inc., have a statistically significant higher average loyalty to their topic areas than the authors from academic institutions (like universities and research laboratories). One possible explanation for this is that authors from corporate institutions are more likely to publish in an area that will serve the interest of their company, while authors from academia have more flexibility to explore new research areas.

5.2 Senate Bill Sponsorship Network

The senate bill sponsorship network is based on data collected about United States senators and the bills they sponsor ([35]). The data contains senators' demographic information and the bills each senator sponsored or co-sponsored from 1993 through February 2008. Each bill has a date and topics associated with it. We group the bills using their high-level topic, and then measure the loyalty of senators to different topics. After removing the senators that do not sponsor a bill or sponsor only a single bill and removing bills that do not have a topic, our analysis uses 181 senators, 28,372 bills, and 188,040 participation relationships spanning 100 high level topics.

When considering only the topics that each senator is most loyal to, the three bill topics that have the highest average loyalty values are Commem-

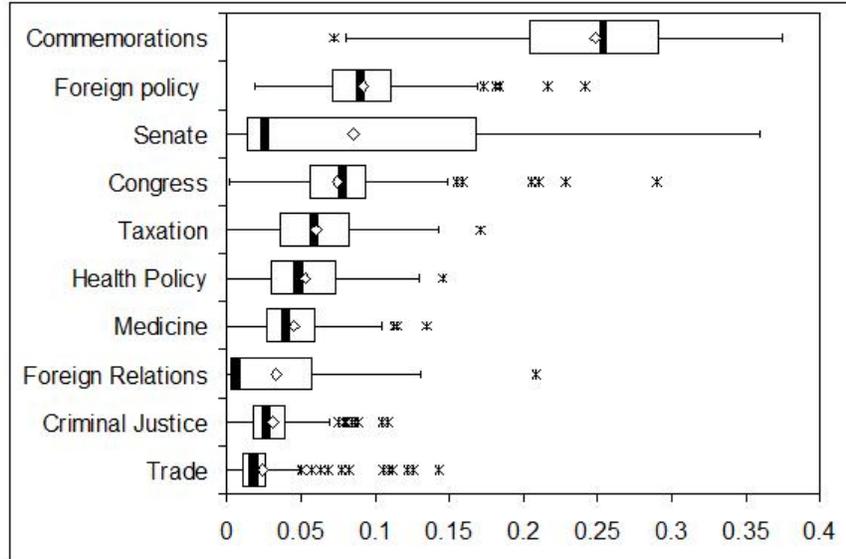


Figure 7: Average topic loyalty across all topics in the senator bill sponsorship network

orations, Senate, and Congress. This average loyalty ranges from 0.22 to 0.27. By investigating the data set, we found that these three topics constitutes 56,035 (approximately 30%) of the total number of sponsorship/co-sponsorship relationships. This finding seems consistent since bills with these topics occur frequently, regularly, and have a large number of senators sponsoring them. Figure 7 shows the 10 bill topics with the highest average loyalty across all the topic groups each senator sponsors a bill in. When looking across all topics for each senator, foreign policy has the second highest average loyalty value. This seems reasonable since the United States has been at war in recent years. In this category, Senator Joe Biden has the highest senator loyalty. Still, the average loyalty of senators to bill topics is generally low. This results because of the large number of bills sponsored by senators across a large number of topics. Many may find comfort in this result since senators supporting bills across topics can be interpreted to mean that they are servicing a wider constituency.

To better understand the changes in loyalty over time, we investigate the changing dynamics of a particular senator’s loyalty over time. We selected the senator that sponsored the largest number of bills - Senator Edward Kennedy, a democrat from Massachusetts. As illustrated in Figure 8, we

calculated his group loyalty at 5 different time points. Although he sponsors bills across 130 topics, our graph shows nine topics with the highest means and standard deviations for loyalty values across the entire time period. During each time period, he consistently sponsors or co-sponsors roughly 10% of the Senate bills. The figure illustrates that Senator Kennedy starts out with a distribution of loyalty that favors a small number of bill topics. He does not sponsor bills across all the topics listed. Over time his loyalty to some of the topics decreases and increases to others as highlighted in the figure. It is also interesting to note that the variance of his loyalty across the topics decreases over time.

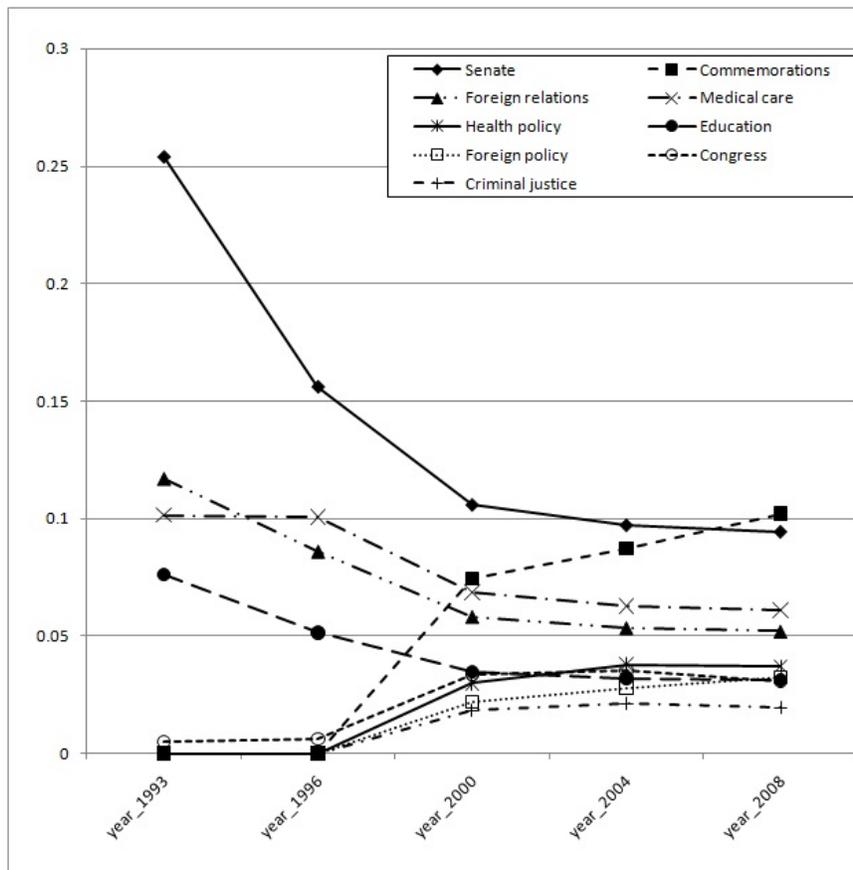


Figure 8: Changing loyalty over time for Edward Kennedy in the senate bill sponsorship network

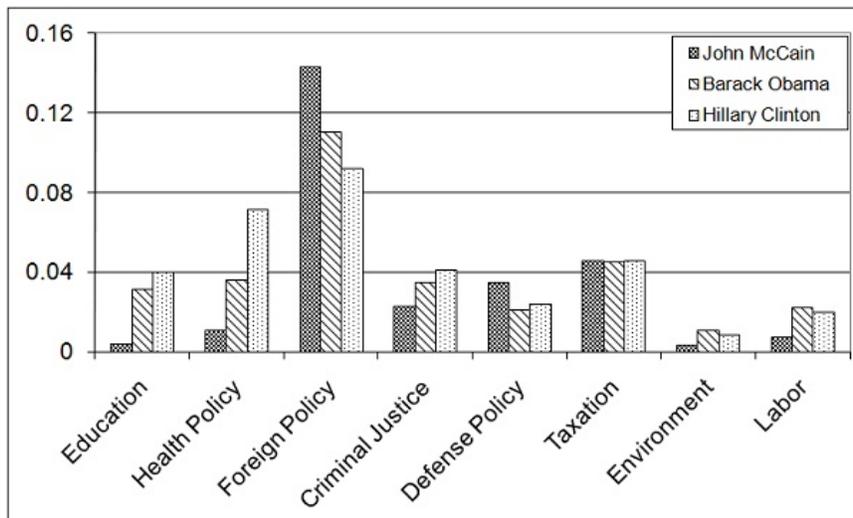


Figure 9: Average topic loyalty of 2008 presidential candidates in the senate bill sponsorship network

Finally, we briefly consider the 2008 presidential election. Examining results in the spring of 2008, in the time period preceding the 2008 fall presidential election, we compared the loyalty of the presidential candidates, John McCain, Barack Obama, and Hillary Clinton across a subset of bill sponsorship topics. The results are shown in Figure 9. These bill sponsorship loyalty values are consistent with priorities emphasized on the campaign trail. All the candidates have strong positions on foreign policy. Senator McCain made it a centerpiece of his campaign. Senator Clinton had highlighted her commitment to health care. Both Senators Obama and Clinton also spent a lot of time discussing education. Interestingly, Senator McCain’s loyalty to sponsoring education bills is very low.

5.3 Dolphin Social Network

We also consider an affiliation network based on a data set describing a long-term study of a wild bottlenose dolphin (*Tursiops* sp.) population in Shark Bay Australia ([36]). It is the most comprehensive dolphin data set in research today with over 20 years of behavioral, reproductive, demographic and ecological data on wild bottlenose dolphins.

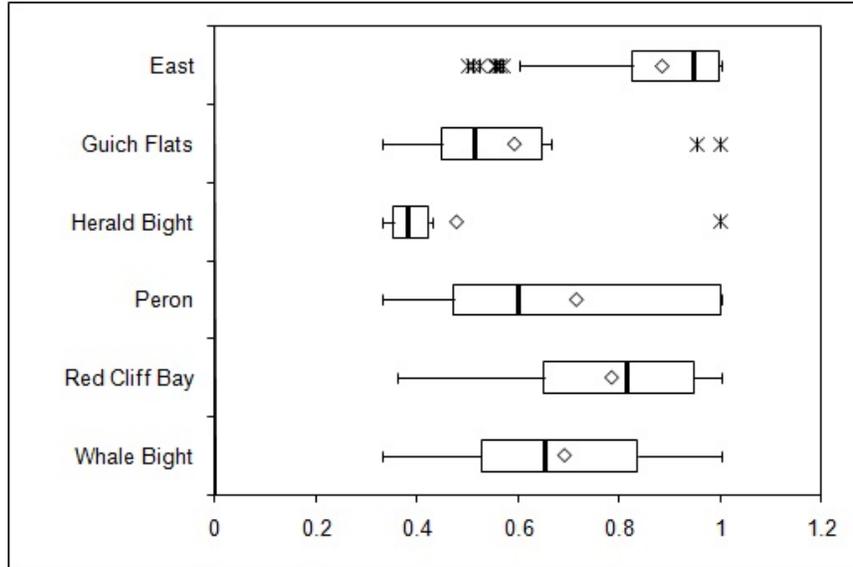


Figure 10: Average location loyalty for dolphins

For this analysis, we focus on observational surveys, collected by researchers on the Shark Bay Dolphin Research Project (SBD RP). Data gathered includes location, animal behaviors, associates, habitat, photographic information, and physical data (e.g., scars, condition, speckles). These surveys are brief, typically lasting 5 to 10 minutes. They are used to present a “snapshot” of associations and behaviors among dolphins.

The affiliation network is defined by using dolphins as actors and surveys as events. Dolphins observed in a survey constitutes the participation relationship. We group survey observations together by the location the observation takes place. There are six different general regions in this data set. Similar to the other analysis, we remove dolphins with few sightings (less than 5) and we remove surveys with no location. After doing this, our analysis includes 560 dolphins, 10,731 surveys, and 36,404 relationships between dolphins and surveys for the loyalty analysis.

Figure 10 show the average loyalty of dolphins to different locations based on the observational surveys. First, the average loyalties of dolphins across all locations ranges from 0.45 to 0.9. Some locations appeared to invite higher loyalty than others, e.g. *East* and *Red Cliff Bay*. One explanation for this is the varying habitat structure. For example, *East*, which has the highest loyalty, is mostly deep channels bisected by shallow sea grass banks.

Many dolphins spend a large amount of time foraging. The extensive habitat heterogeneity might limit the region to dolphins with certain foraging specialization (channel foragers or sea grass bed foragers). For example, a subset of the dolphins in this population use sponges as foraging tools, and will forage almost exclusively in the *East* channels [37]. *Peron* is at the tip of the peninsula and is a very open area where the western and eastern gulf meet. This open habitat (to the Indian Ocean) may allow for great mobility and less loyalty when compared to other areas.

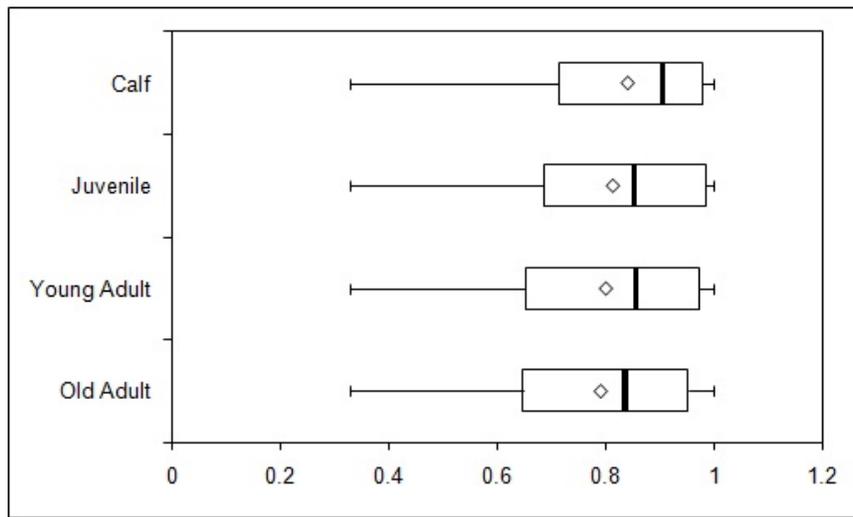


Figure 11: Average location loyalty grouped by age groups for dolphins

Previous work by project biologists indicates calves are most tied to the locations of their mothers and maternal foraging type [38]. After weaning, juveniles might range further and develop bonds with others separate from the mother. Figure 11 looks at the distribution of location loyalty among different age groups: calves (0-4 years), juveniles (5-11 years), young adults (12-24 years), and old adults (25+ years). The results indicate that loyalty decreases with age, but still remains very high. This may occur because older dolphins travel more during the course of their life and they explore more places, while calves tend to have higher loyalty to a small number of locations (which happen to be the ones their mothers are also in). Location loyalty is a nice indication of long-term residency in the population and allows researchers to track individuals over long periods of time.

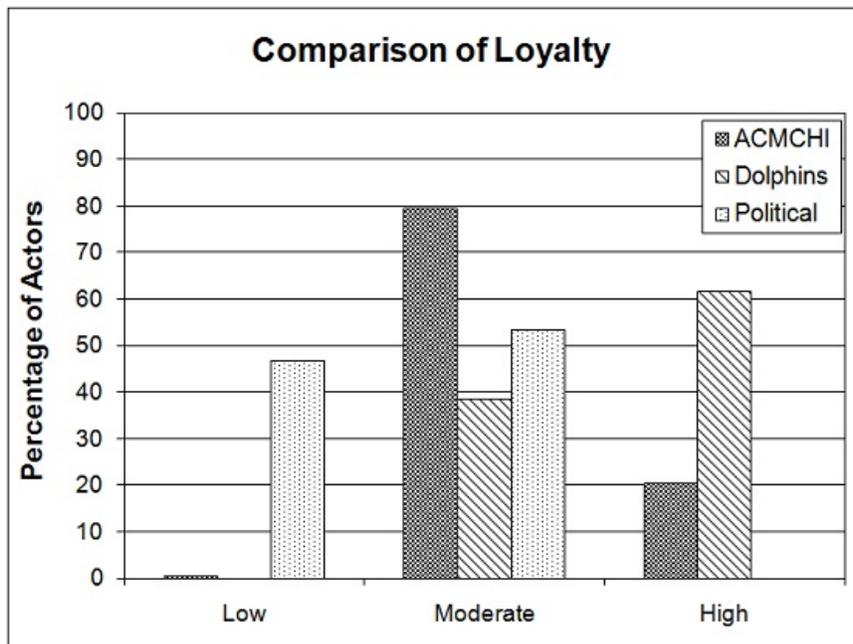


Figure 12: Loyalty Comparison Across Data Sets

6 Comparative Loyalty Analysis

We now compare the average actor loyalty across these different communities. As the loyalty metric values can vary from zero to one, we divided the range of loyalty into three bins; low loyalty (scores from 0 to 0.25), moderate loyalty (scores from 0.25 to 0.75), and high loyalty (scores from 0.75 to 1).

The results in Figure 12 shows the percentage of actors with loyalty scores falling in each of the three bins for the scientific publication network, the dolphin survey network, and the political bill sponsorship network, respectively. This figure highlights the different distribution of actor loyalty in the different data sets.

For the ACMCHI publication network, we can notice that 79.2% of the authors have moderate loyalty and most of the rest (20.4%) have high loyalty to the topic of their publications. In the political data set, we find that 53.3% of the senators have moderate loyalty to the topic of the bill they sponsor, and the rest fall in the low loyalty category. For the dolphin affiliation network, we can observe that most of the dolphins (61.5%) have high loyalty

to their locations. This large variation in the distribution of actor loyalty across data sets reinforces the utility of a measure that captures changing loyalty of actors to affiliation groups.

These classification results are consistent with the interpretations of community loyalty presented in the previous section. The figure highlights the varying distribution of actor loyalty to groups in each affiliation network. As a final analysis, for each affiliation network we consider the average number of events each actor participates in. This allows us to compare the loyalty of these affiliation networks to the density of the connections in the network. The averages are as follows:

1. Average number of Publications per Author = 3.61
2. Average number of Bills per Senator = 159.94
3. Average number of Observations per dolphin = 19.16

The most dense network is the senate bill sponsorship network, followed by the dolphin social network. The author publication network is much more sparse than the other two networks. Interestingly enough, the loyalty categories are not completely consistent with these frequency averages, thereby affirming that frequency alone may not be sufficient to capture loyalty.

7 Comparison with centrality measures

It is natural to want to understand how loyalty compares to existing centrality measures. Does it capture the same information, or does it provide additional insight? We begin by comparing actor loyalty to the most common centrality measures. The first centrality measure used is betweenness centrality, calculated by computing all pairs shortest paths in the network and computing the number of shortest paths that the target node occurs on. The second centrality measure used is the closeness centrality, defined as the average of shortest paths from the target nodes to all other nodes reachable from it. Lastly, eigenvector centrality measures importance of a node based on the importance of neighboring nodes. For more detail about these measures, we refer you to [11].

In order to perform the comparison, we generated the underlying single-mode, co-membership network for actors participating in each affiliation group, and computed various centrality measures on the generated networks. As can be noted in Figure 13, the scatter plot between loyalty and various centrality measures on the publication data set shows authors having all

combination of values for both measures, with no visible trend in the results. The same results holds for the other two data sets.

To further investigate loyalty to a particular topic, we take a more detailed look at the ‘Information Visualization’ topic as a sample affiliation group. For this group, Benjamin Bederson is ranked as the author with highest betweenness and eigenvector centrality. However, by examining figure 14, we notice that his publication pattern is neither consistent across time nor numerous. This is also true for Robert Spence who was ranked first according to the closeness centrality. On the other hand, the time-consistent, recent and numerous publications of the most loyal author, namely Stuart Card, as shown in the same figure, illustrates exactly what our proposed loyalty measure captures that the other centrality measures missed.

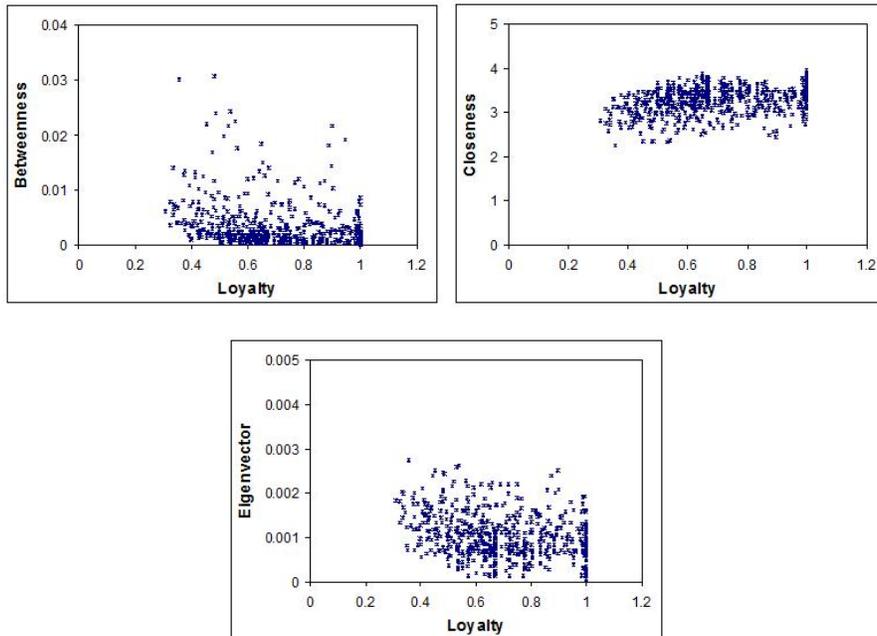


Figure 13: Loyalty vs. Centrality for Scientific Publication Network

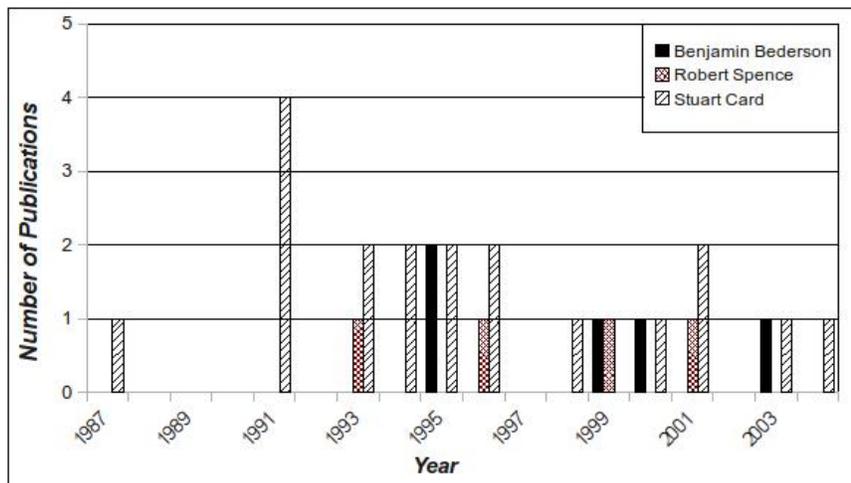


Figure 14: Author publications in "Information Visualization" topic

8 Conclusion and Future Directions

In this paper, we proposed a new measure for capturing loyalty in time-varying affiliation networks. We begin by defining affiliation groups which describe temporally related subsets of actors. This is accomplished by grouping events over time based on attribute values. To model the dynamic behavior of affiliations to groups, we consider the concept of loyalty and introduce a measure that captures an actor's loyalty to an affiliation group as the degree of 'commitment' an actor shows to the group over time. We compare our measure to both frequency-based loyalty and recency-based loyalty and find our measure to be more flexible since it incorporates components for frequency, consistency, and recency.

We demonstrate its utility on three real world affiliation networks: a publication network, a senate bill co-sponsorship network, and a dolphin network. It is interesting to note that the distribution of actor loyalty varies substantially across data sets, thereby reinforcing the utility of a measure that captures changing loyalty of actors to affiliation-based groups.

One interesting direction of future work involves studying the changing group composition over time. Do larger groups contain a higher percentage of loyal actors or do smaller groups exhibit this behavior? How cohesive are loyal group members? Can we predict group loyalty based on changes to an actor's affiliations over time or based on member actor loyalty distributions?

How do these dynamics change as the size and density of the network increases? There are still a large number of outstanding questions related to the dynamics of actors and groups in affiliation networks that are challenges for researchers across disciplines.

References

- [1] R. M. Kanter, “Commitment and social organization: A study of commitment mechanisms in utopian communities,” *American Sociological Review*, vol. 33, pp. 499–517, 1968.
- [2] D. C. Dunphy, “The social structure of urban adolescent peer groups,” *Sociometry*, vol. 26, pp. 230–246, 1963.
- [3] M. E. Shaw, “Group dynamics,” *Annual Review of Psychology*, vol. 12, pp. 129–156, 1961.
- [4] R. L. Oliver, “Loyalty: Whence consumer loyalty?” *Journal of Marketing*, 1999.
- [5] R. J. Sampson and W. B. Grove, “Community structure and crime: testing social disorganization theory,” *American Journal of Sociology*, vol. 94, pp. 774–802., 1989.
- [6] J. Moody and D. R. White, “Structural cohesion and embeddedness: A hierarchical concept of social groups,” *American Sociological Review*, vol. 68, no. 1, pp. 103–127, 2003.
- [7] D. Cai, Z. Shao, X. He, X. Yan, and J. Han, “Community mining from multi-relational networks,” *In Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, pp. 445–452, 2005.
- [8] S. Hill, D. Agarwal, R. Bell, and C. Volinsky, “Building an effective representation of dynamic networks,” *Journal of Computational and Graphical Statistics*, 2006.
- [9] T. A. Snijders, “Models for longitudinal network data,” *In P. Carrington, J. Scott, and S. Wasserman (Eds.), Models and methods in social network analysis. New York: Cambridge University Press*, p. Chapter 11, 2005.

- [10] D. Druckman, “Nationalism, patriotism, and group loyalty: A social psychological perspective,” *Mershon International Studies Review*, vol. 38, no. 1, pp. 43–68, 1994.
- [11] S. Wasserman and K. Faust, *Social network analysis: methods and applications*. Cambridge: Cambridge University Press, 1994.
- [12] H. Sharara, L. Singh, L. Getoor, and J. Mann, “The dynamics of actor loyalty to groups in affiliation networks,” *In Proceedings of the 2009 International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2009.
- [13] H. S. Becker, “Notes on the concept of commitment,” *Journal of Sociology*, vol. 66, pp. 32–42, 1960.
- [14] C. A. Kiesler, *The psychology of commitment: Experiments linking behavior to belief*. Academic Press, 1971.
- [15] J. Meyer, T. Becker, and C. Vandenberghe, “Employee commitment and motivation: A conceptual analysis and integrative model,” *Journal of Applied Psychology*, vol. 89, no. 6, pp. 991–1007, 2004.
- [16] R. M. Cunningham, “Brand loyalty-what, where, how much?” *Harvard Business Review*, vol. 34, pp. 116–128, 1956.
- [17] J. W. Newman and R. A. Werbel, “Multivariate analysis of brand loyalty for major household appliances,” *Journal of Marketing Research*, vol. 10, pp. 404–409, 1973.
- [18] G. Tellis, “Advertising exposure, loyalty, and brand purchase: A two-stage model of choice,” *Journal of Marketing Research*, vol. 25, pp. 134–144, 1988.
- [19] J. Jacoby and D. B. Kyner, “Brand loyalty vs. repeat purchasing behavior,” *Journal of Marketing Research*, vol. 10, pp. 1–9, 1973.
- [20] D. Cai, Z. Shao, X. He, X. Yan, and J. Han, “Community mining from multi-relational networks,” *In Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, 2005.
- [21] G. Flake, S. Lawrence, C. Giles, and F. Coetzee, “Self-organization and identification of web communities,” *Computer*, vol. 35, no. 3, pp. 66–71, 2002.

- [22] M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” *PNAS*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [23] J. Hopcroft, O. Khan, B. Kulis, and B. Selman, “Natural communities in large linked networks,” *In Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.
- [24] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney, “Statistical properties of community structure in large social and information networks,” in *Proceeding of the 17th international conference on World Wide Web*, 2008, pp. 695–704.
- [25] M. E. J. Newman, “Detecting community structure in networks,” *Eur. Phys. J. B*, vol. 38, pp. 321–330, 2004.
- [26] S. Asur, S. Parthasarathy, and D. Ucar, “An event-based framework for characterizing the evolutionary behavior of interaction graphs,” *In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 913 – 921, 2007.
- [27] 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*, 2006.
- [28] T. Berger-Wolf and J. Saia, “A framework for analysis of dynamic social networks,” *In Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006.
- [29] L. Friedland and D. Jensen, “Finding tribes: identifying close-knit individuals from employment patterns,” *In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2007.
- [30] T. A. Snijders, “Models for longitudinal network data,” *In P. Carrington, J. Scott, and S. Wasserman (Eds.), Models and methods in social network analysis*. New York: Cambridge University Press, p. Chapter 11, 2005.
- [31] J. Sun, C. Faloutsos, S. Papadimitriou, and P. Yu, “Graphscope: parameter-free mining of large time-evolving graphs,” *In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 687 – 696, 2007.

- [32] C. Tantipathananandh, T. Berger-Wolf, and D. Kempe, “A framework for community identification in dynamic social networks,” *In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2007.
- [33] Habiba, T. Y. Berger-Wolf, Y. Yu, and J. Saia, “Finding spread blockers in dynamic networks,” *In Proceedings of the 2nd SNA-KDD workshop on Web mining and social network analysis*, 2008.
- [34] U. Sharan and J. Neville, “Temporal-relational classifiers for prediction in evolving domains,” *In Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 2008.
- [35] Govtrack, “Senate bill sponsorship data,” website: <http://www.govtrack.us>, 2008.
- [36] J. Mann, R. C. Connor, L. M. Barre, and M. R. Heithaus, “Female reproductive success in bottlenose dolphins (*Tursiops* sp.): Life history, habitat, provisioning, and group size effects,” *Behavioral Ecology*, vol. 11, pp. 210–219, 2000.
- [37] B. L. Sargeant, A. J. Wirsing, M. R. Heithaus, and J. Mann, “Can environmental heterogeneity explain individual foraging variation in wild bottlenose dolphins?” *Behavioral Ecology and Sociobiology*, vol. 61, pp. 679–688, 2007.
- [38] J. Mann and B. Sargeant, “Like mother, like calf: The ontogeny of foraging traditions in wild indian ocean bottlenose dolphins,” *In D. Fragaszy and S. Perry (Eds.), The Biology of Traditions: Models and Evidence. Cambridge University Press*, pp. 236–266, 2003.