The Dynamics of Actor Loyalty to Groups in Affiliation Networks

Hossam Sharara
Computer Science Dept.
University of Maryland
College Park, MD
hossam@cs.umd.edu

Lisa Singh
Computer Science Dept.
Georgetown University
Washington, DC
singh@cs.georgetown.edu

Lise Getoor
Computer Science Dept.
University of Maryland
College Park, MD
getoor@cs.umd.edu

Janet Mann
Biology Dept.
Georgetown University
Washington, DC
mannj@georgetown.edu

***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 two real world affiliation networks: a senate bill co-sponsorship network and a dolphin network. The results show how the behavior of actors in different affiliation groups change dynamically over time, reinforcing the utility of our measure for understanding the loyalty of actors to time-varying affiliation groups.**

I. 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 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].

We will 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 will 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 and commitment to groups over time.

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 capture 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.

II. MODELING TIME-VARYING EVENT-BASED GROUPS

An affiliation network $\mathcal{G}(\mathcal{A}, \mathcal{E}, \mathcal{P})$ contains 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 $\mathcal{E}$:

$$\mathcal{A} = \{a_1, a_2, a_3, \ldots, a_n\}$$

978-0-7695-3689-7 2009

DOI 10.1109/ASONAM.2009.27

$$\mathcal{E} = \{e_1, e_2, e_3, \ldots, e_m\}, \text{and}$$

$$\mathcal{P} = \{(a_i, e_j) | a_i \in \mathcal{A}, e_j \in \mathcal{E} \forall (a_i, e_j)\}.$$

We denote participation of actor a_i in event e_j as $p_{i,j}$. For clarity, we will 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, $\mathcal{E} = \{e_1, e_2, \ldots e_{15}\}$, and twenty paper authorship edges. As an example, those involving actor a_1 are the following: $\mathcal{P}_{a_1} = \{p_{1,1}, p_{1,2}, p_{1,3}, p_{1,4}, p_{1,5}, p_{1,7}, p_{1,8}, p_{1,9}, p_{1,10}, p_{1,11}, p_{1,13}, p_{1,14}, p_{1,15}\}$.

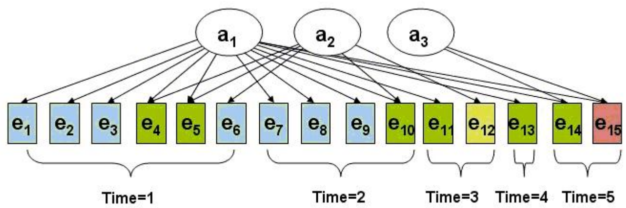

Figure 1. An affiliation network example with 3 actors, 15 events and 20 relationships across 5 time points.

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 $a_1.name$ = 'Peter Pan' and $a_1.age$ = 50. Each publication event may have a title attribute, e.g. $e_1.title$ = 'Static networks as non-evolving dynamic networks' and a topic attribute, $e_1.topic$ = '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 will 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 each event serves as a grouping of a subset of actors, it 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.

Each event feature or attribute F has an associated domain $Domain = \{g_1, g_2, \ldots, g_p\}$, 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: $G(g_l, t) = \{a_i | a_i \in \mathcal{A}, (a_i, e_j) \in \mathcal{P},$ where $e_j.F = g_l$ 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.

III. 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 represents different topic groups and an edge from the author to a topic represents that an author has published on this topic. The count on the edge represents the number of publications with this topic. For example, the network snapshot of the first time period shows the author a_1 having three publications with $topic_1$ and two publications with $topic_2$. As time continues, author a_1 stops publishing in $topic_1$, continues publishing in $topic_2$ at each time step, and begins publishing in $topic_3$ 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.

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.

Focusing on the temporal aspect of the data, a *recency-based loyalty* measure considers an actor loyal if she 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. Then 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:

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

Using our example at time point 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.

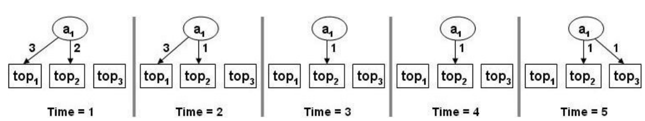

Figure 2. Single actor dynamic affiliation example

From this simple example, we see that a temporal measure of affiliation group loyalty should incorporate group frequency, consistency, and recency. In order to capture all of these, we extend the previous loyalty measures. 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 at time t on his first participation in it is 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 $Loy(a_i, g, t_{prev})$ which takes into consideration that latest recorded loyalty for an actor in a specific group and is used to favor recent participations in a 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 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 1 for T_{tot} means that the recent participation component of the measure is dominant.

For consistency with the group membership notation, 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:

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

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

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

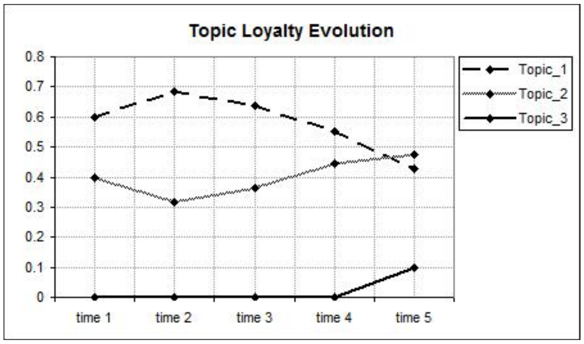

Figure 3. The evolution of loyalty over time for our simple example.

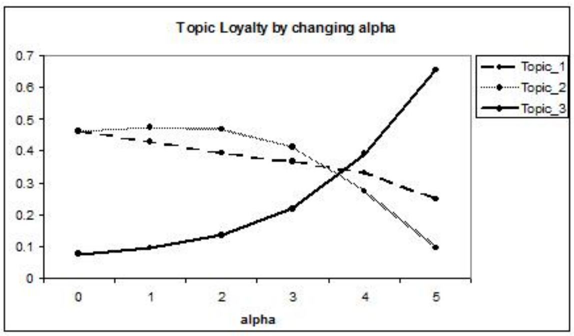

Figure 4. The effect of the smoothing factor in calculating group loyalty.

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 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). If we let α approach infinity the loyalty of $topic_3$ will approach 1 and the loyalty of the other two topics will approach zero.

IV. Loyalty Analysis on Individual Data Sets

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

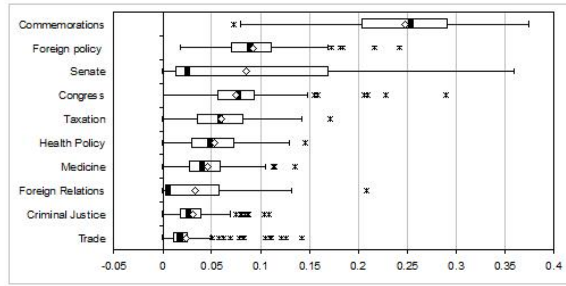

Figure 5. Average topic loyalty across all topics in the senator bill sponsorship network.

A. Senate Bill Sponsorship Network

The senate bill sponsorship network is based on data collected about United States senators and the bills they sponsor ([12]). 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 Commemorations, Senate, and Congress. This average loyalty ranges from 0.22 to 0.27. By investigating the dataset, 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 5 shows the 10 bill topics with the highest average loyalty across all the topics 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 also 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 means 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 6, we calculated his group loyalty at 5 different time points. Al-

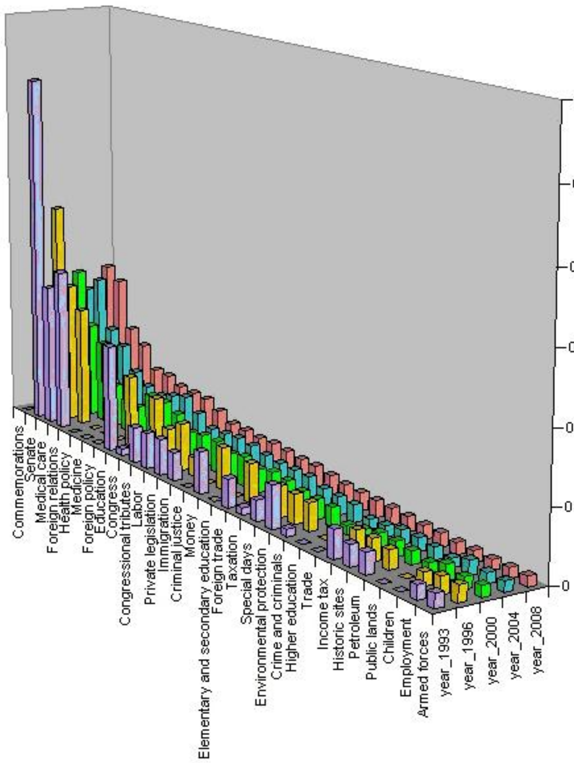

Figure 6. Changing loyalty over time for Edward Kennedy in the senate bill sponsorship network.

though he sponsors bills across 130 topics, our graph shows the thirty topics with the highest 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 by the changing sizes of the bars.

B. Dolphin Social Network

We also consider an affiliation network based on a dataset describing a long-term study of a wild bottlenose dolphin (Tursiops sp.) population in Shark Bay Australia ([13]). 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. For this analysis, we focus on observational surveys, collected by researchers on the Shark Bay Dolphin Research Project (SBDRP). 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 7 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*. This is likely due to 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 [14]. *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.

Previous work by project biologists indicates calves are most tied to the locations of their mothers and maternal foraging type [15]. After weaning, juveniles might range further and develop bonds with others separate from the mother. Figure 8 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 the 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.

V. 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

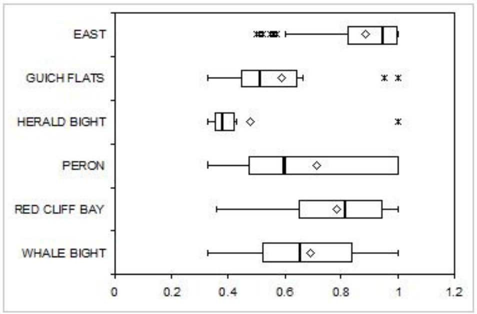

Figure 7. Average location loyalty for dolphins

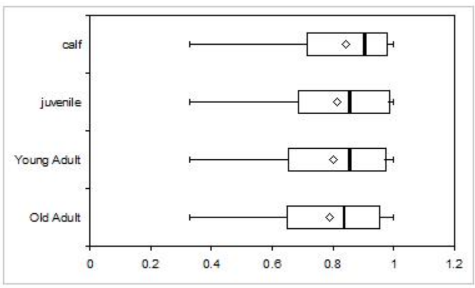

Figure 8. Average location loyalty grouped by age groups for dolphins

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 then demonstrate its utility on two real world affiliation networks. In general, the average loyalty of senators to groups based on topics of bills they sponsor was less than the average loyalty of Shark Bay dolphins to groups based on locations they are observed in. The varying characteristics across data sets reinforces the utility of a measure that captures changing loyalty of actors to affiliation 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] Govtrack, "Senate bill sponsorship data," website: http://www.govtrack.us, 2008.

[13] 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.

[14] 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.

[15] 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.