Loyalty in Online Communities

William L. Hamilton[*]
Stanford University
wleif@stanford.edu

Justine Zhang[*]
Cornell University
jz727@cornell.edu

Cristian Danescu-Niculescu-Mizil
Cornell University
cristian@cs.cornell.edu

Dan Jurafsky
Stanford University
jurafsky@stanford.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

Abstract

Loyalty is an essential component of multi-community engagement. When users have the choice to engage with a variety of different communities, they often become loyal to just one, focusing on that community at the expense of others. However, it is unclear how loyalty is manifested in user behavior, or whether loyalty is encouraged by certain community characteristics.

In this paper we operationalize loyalty as a user-community relation: *users loyal* to a community consistently prefer it over all others; *loyal communities* retain their loyal users over time. By exploring this relation using a large dataset of discussion communities from Reddit, we reveal that loyalty is manifested in remarkably consistent behaviors across a wide spectrum of communities. Loyal *users* employ language that signals collective identity and engage with more esoteric, less popular content, indicating they may play a curational role in surfacing new material. Loyal *communities* have denser user-user interaction networks and lower rates of triadic closure, suggesting that community-level loyalty is associated with more cohesive interactions and less fragmentation into subgroups. We exploit these general patterns to predict future rates of loyalty. Our results show that a user's propensity to become loyal is apparent from their first interactions with a community, suggesting that some users are intrinsically loyal from the very beginning.

1 Introduction

The Internet offers a staggering variety of virtual communities for the intrepid wanderer to explore. Faced with this abundance of options, a user may have fleeting relationships with some communities, choosing to allocate only a small proportion of her time there (Tan and Lee 2015). Alternatively, a user may commit to forming a more steadfast relationship with one particular community, establishing her loyalty to that group by consistently preferring it above all others.

While there is a rich literature studying various flavors of user-to-user relationships within individual communities (Fiore, Tiernan, and Smith 2002; Arguello et al. 2006; Danescu-Niculescu-Mizil, Gamon, and Dumais 2011; Ellis

et al. 2016), less focus has been directed to understanding relationships that exist between *users* and *multiple communities*.

Loyalty is a fundamental example of such a relationship, between a user and multiple communities. In multi-community platforms like Reddit, users have no shortage of alternative communities to peruse (Tan and Lee 2015; Hessel, Tan, and Lee 2016), so understanding why a user chooses to be loyal to one community and not others is a central problem. Characterizing these loyal relationships in terms of the traits of their two components—users and communities—can offer insights into how community identity arises online (Nguyen and Rose 2011), and also guide community maintainers towards fostering better user involvement, with their most faithful users in mind.

Loyalty is fundamentally about the preferences and commitments of active, engaged users, and while there has been considerable effort spent studying how to attract and retain new users (Karnstedt et al. 2010; Dror et al. 2012) or improve user engagement (Arguello et al. 2006; Backstrom et al. 2013), there is comparatively little understanding of how already-active users choose to allot their time between communities (Tan and Lee 2015). In particular, loyalty can only exist within the context of multi-community dynamics. Unlike the well-studied issue of user churn or retention, which is usually defined in terms of a single, isolated community (Dror et al. 2012; Danescu-Niculescu-Mizil et al. 2013), understanding loyalty requires examining user preferences within a large multi-community setting.

Present work. Our aim is to provide a thorough characterization of user loyalty in the massive, multi-community platform, Reddit. To this end, we develop of measure of loyalty in terms of user preference and commitment; loyal users prefer one community above all others and sustain this preference over time. Applying this measure to Reddit, we characterize loyal *users*, as well as *communities* that foster user loyalty. The large-scale, multi-community nature of Reddit, where users can peruse thousands of possible communities, makes it an ideal setting to study loyal behavior.

We reveal behavioral traits that systematically distinguish loyal users across a diverse set of communities, and show that loyal communities share similar structural features. Together these observations provide a cross-community char-

---

[*]The two first authors contributed equally and are ordered alphabetically.
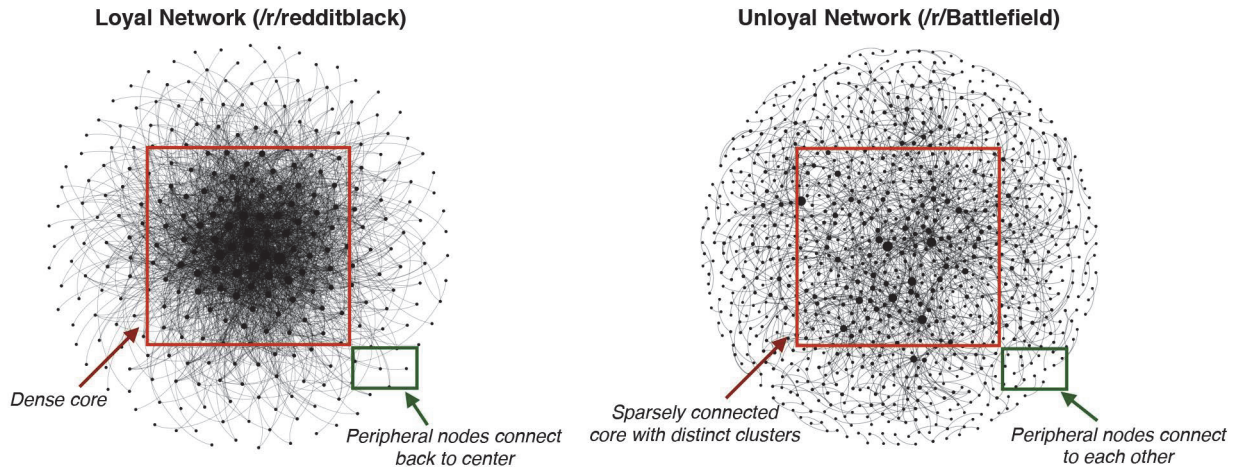
Figure 1: Example loyal and unloyal networks built from interactions in March, 2014 in two war-related gaming communities.

acterization of loyalty that lends a new perspective on user-community dynamics.

2 Operationalizing Loyalty

We define loyalty as a combination of *preference* and *commitment*. Loyal users exhibit a clear preference for one community and sustain this preference over time. This conceptualization of loyalty is closely related to traditional notions of user engagement and retention; however, a key distinguishing factor is that we focus on relative preference, rather than absolute activity levels. Our goal is to study a strong form of user-community loyalty, where users identify with, and consistently participate in, one community more than others.

The following sections describe our dataset, Reddit, and how we operationalize this definition of loyalty in the data. See the extended version of this paper for further motivation, background, and dataset details (Hamilton et al. 2017).

Dataset. We perform our analysis using a dataset of posts and comments from Reddit, a popular website where users form topical discussion-based communities (called subreddits). Our full dataset consists of all comments and posts made to Reddit in 2014: approximately $\sim 10^8$ comments made by $\sim 10^7$ users in $\sim 10^4$ communities.[1]

Loyal users. We define user loyalty on Reddit based on commenting behavior, which we view as a strong proxy for latent engagement. To focus our attention on loyalty instead of platform-level retention, all of the following definitions are restricted to users who commented at least 10 times within the relevant time period. Additionally, we only consider top-level comments that are initial responses to a post. Top-level comments demonstrate a user's evaluative choice to comment in a particular community, while lower-level comments may result from the social obligation to maintain a conversation.

We say that a user *X prefers* a community *A* in month *t* if at least 50% of the comments that *X* submits across Reddit in *t* are to *A*. *X* is then *loyal* to *A* at *t* if *X* prefers *A* at both *t* and *t* + 1 (i.e., exhibits *commitment*). We use monthly time windows, following common practice in studies of user engagement and churn (e.g., (Oentaryo et al. 2012; Danescu-Niculescu-Mizil et al. 2013)). To provide a reference point for the behavior of loyal users, we contrast loyal users with *vagrant* users who fleetingly interact with a community before wandering off. We define a *vagrant* of *A* as a user who comments between 1 and 3 times in *A* at *t*, and, while still active on Reddit at time *t* + 1, does not contribute to *A* in *t* + 1.

In order to have enough statistical power for within-community analyses, we restrict our user-level studies to communities with at least 25 loyal and 25 vagrant users per month, resulting in 242 communities.

Loyal communities. As a natural extension of our user-level definition, we say a *community* is loyal if it tends to retain a high proportion of loyal users month after month. We focus on a community's success in retaining loyal users, rather than counts of loyal users at particular points in time, in order to minimize confounds due to differences in community sizes and activity levels.

We compute a *loyalty-rate* for each community *A* as the expected proportion of users who prefer[2] *A* at *t* and sustain this preference at *t* + 1. Users who prefer *A* at *t* and then leave Reddit altogether are ignored, since we seek to model inter-community loyalty and not platform-level churn.

We analyze all communities with at least 25 loyal users in one month, resulting in 1440 communities (with a median

---

[1]https://archive.org/details/2015_reddit_comments_corpus.

[2]To achieve sufficient statistical power, in community-level experiments we only require that users comment in *A* more than any other community that month, without the 50% threshold used in the user-level experiments.

loyalty rate of 60.7%). For the purpose of this study, we denote *loyal communities* as the top-25% of this distribution, and *non-loyal* communities as the bottom-25%.

3 Community-level Loyalty

We begin by analyzing the communities that tend to foster high rates of loyalty. We find that loyal communities exhibit consistent structural features in their user-user interaction networks. The results presented in this section summarize a subset of the findings in the extended version of this paper (Hamilton et al. 2017), which further characterizes the structural characteristics of loyal communities and how these properties relate to a community's topical focus.

For each community, we construct monthly interaction networks where users are connected if they comment in the same linear comment chain within three comments of each other (i.e., separated by at most two comments).[3]

Loyal communities tend to be smaller. Comparing total numbers of commenters, loyal communities are significantly smaller than non-loyal communities ($p < 10^{-5}$, U Test); the median loyal community is 39% smaller than the median non-loyal one. For example, many communities that are highly successful at retaining loyal users are small fan-fiction or role-playing communities, such as /r/HarryPotterRP or /r/randomsuperpowers (a community where users construct individualized superhero identities).

Loyal communities have denser interaction networks. The user-user interaction networks of loyal communities have significantly higher edge density, even after controlling for activity levels, meaning that the average user in a loyal community interacts with a greater number of other users. To test differences in density while controlling for activity, we compare matched pairs of communities that are closest in their activity levels (i.e., $\log\left(\frac{\text{\# comments}}{\text{\# users}}\right)$). After discarding pairs that differ by more than one-tenth of a standard deviation, we find a significant disparity in edge density ($p < 10^{-4}$, Wilcoxon Test).

Loyal communities are less assortative. The interaction networks of loyal communities also have significantly less activity assortativity ($p < 10^{-5}$, U test). In communities that foster loyalty, highly active users tend to engage with others who have a wide variety of activity levels, while in non-loyal communities users tend to comment near other users of similar activity levels.

Loyal communities are less clustered. When compared to a random null baseline (to control for density effects), loyal networks are significantly less clustered ($p < 10^{-5}$, U test). Random null-model networks were generated by randomly rewiring edges within a community's interaction network while maintaining node degrees, with the number of rewiring iterations set at $10^4\times$ the edge count in the empirical network. For each community, we compute statistics by taking the relative difference of the median monthly empirical statistic compared to the median monthly null statistic.

Figure 1 shows two example networks highlighting how loyalty is reflected in user-user interactions. See the ex-

---

[3]Networks and details available at: http://snap.stanford.edu/data/web-RedditNetworks.html.
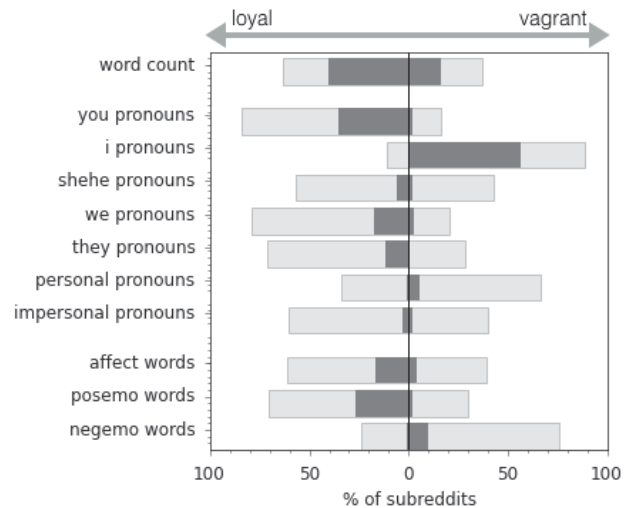

Figure 2: Linguistic features of comments written by loyal and vagrant users. Light bars indicate percentages of subreddits where a feature is exhibited more by loyal vs vagrant commenters; dark bars indicate the proportion of subreddits in which this effect is significant (Wilcoxon signed-rank test at the $p < 0.01$ level; see extended version for details on how we sampled the comments (Hamilton et al. 2017)).

tended version for more detailed statistics and further analyses on how these network properties are predictive of loyalty, independent of a community's topic (Hamilton et al. 2017).

4 User-level Loyalty

In this section, we analyze loyalty at the level of individual users. We show that markers of loyalty are present in users' very first contributions to a community. The extended version of this work contains more detailed motivation for the features we explore and additional results that show how these features are consistently predictive of loyalty across a diverse set of communities (Hamilton et al. 2017).

We consider the prediction task of determining whether a user will become loyal to a community based on their first comments to that community. We balance between a positive class of users who become loyal within 2 months of arrival, and users who never become loyal beyond their initial activity. To focus on evaluating the predictive power of these small snapshots, we train one classifier *per community* on the initial $k = 3$ comments of users who make their first contribution in January to June of 2014, and predict on users arriving to that community in July to October. We average feature values over the comments, and enforce that users in both classes must have at least k comments. We use random forest classifiers from the scikit-learn (Pedregosa and others 2011) package with ensembles of size 100, setting the minimum number of samples required to split a node to 10.

As features we use the score of the posts that the user comments on as well as the following linguistic features, which we apply to both the text of the user's comments and the posts they reply to: verbosity (i.e., word count), dis-

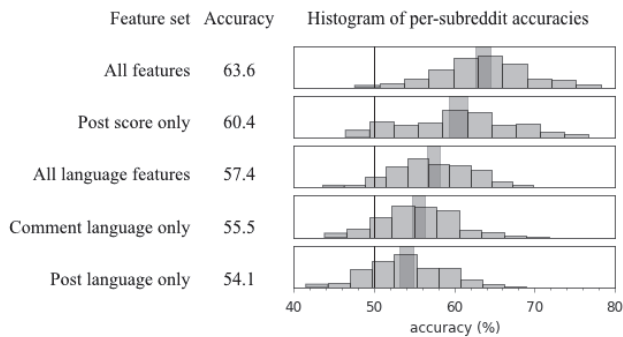| Feature set	Accuracy	Histogram of per-subreddit accuracies
All features	63.6	
Post score only	60.4	
All language features	57.4	
Comment language only	55.5	
Post language only	54.1	

Figure 3: Predicting loyalty from first contributions. Histograms depict test accuracies per subreddit, and average accuracy, for predicting future loyalty given a small window of 3 comments at the start of a user's activity in a community. Shaded regions show bootstrapped 95% confidence intervals.

tribution of personal pronouns, and the quantity of words pertaining to affect (using pronoun and affect word counts from the standard LIWC lexicon (Tausczik and Pennebaker 2010)). The choice of features was motivated by the observation that loyal users tend to comment on low scoring posts and use language that signals collective identity. Figure 2 shows how the linguistic features differ between comments written by loyal and vagrant users. The extended version of this paper contains full details on how we compute these features, how we used matched sampling to control for topical variation, and further analyses on how these features are differentially used by vagrant vs. loyal users (Hamilton et al. 2017).

We find that several groups of these features are significantly predictive of future loyalty in many subreddits (Figure 3). Using all features, 86% of subreddits have accuracies significantly above the random baseline (averaging 63.6% accuracy). Importantly, all the above markers of loyalty are present in users' first few contributions, meaning that these markers are not simply explained by differences in activity levels.

5   Conclusion

In this work we revealed how user-community loyalty manifests in the structural properties of user-user interactions, and we showed that future user loyalty can be predicted from users' first contributions. Our results suggest that some users are intrinsically loyal from the very beginning and that community maintainers could use this information to identify and understand the types of users who are most likely to become faithful contributors.

Acknowledgements

References

Arguello, J.; Butler, B. S.; Joyce, E.; Kraut, R.; Ling, K. S.; Rosé, C.; and Wang, X. 2006. Talk to me: Foundations for successful individual-group interactions in online communities. In *CHI*.

Backstrom, L.; Kleinberg, J.; Lee, L.; and Danescu-Niculescu-Mizil, C. 2013. Characterizing and curating conversation threads: Expansion, focus, volume, re-entry. In *WSDM*.

Danescu-Niculescu-Mizil, C.; West, R.; Jurafsky, D.; Leskovec, J.; and Potts, C. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *WWW*.

Danescu-Niculescu-Mizil, C.; Gamon, M.; and Dumais, S. 2011. Mark my words!: Linguistic style accommodation in social media. In *WWW*.

Dror, G.; Pelleg, D.; Rokhlenko, O.; and Szpektor, I. 2012. Churn prediction in new users of Yahoo! answers. In *WWW*.

Ellis, K.; Goldszmidt, M.; Lanckriet, G.; Mishra, N.; and Reingold, O. 2016. Equality and social mobility in Twitter discussion groups. In *WSDM*.

Fiore, A. T.; Tiernan, S. L.; and Smith, M. A. 2002. Observed behavior and perceived value of authors in usenet newsgroups: bridging the gap. In *CHI*.

Hamilton, W. L.; Zhang, J.; Danescu-Niculescu-Mizil, C.; Jurafsky, D.; and Leskovec, J. 2017. Loyalty in online communities (extended version). *arXiv:1703.03386 [cs.SI]*.

Hessel, J.; Tan, C.; and Lee, L. 2016. Science, AskScience, and BadScience: On the coexistence of highly related communities. In *ICWSM*.

Karnstedt, M.; Hennessy, T.; Chan, J.; and Hayes, C. 2010. Churn in social networks: A discussion boards case study. In *SocialCom*.

Nguyen, D., and Rose, C. 2011. Language use as a reflection of socialization in online communities. In *ACL WKSP on Languages in Soc. Media*.

Oentaryo, R. J.; Ee-Peng, L.; Lo, D.; Zhu, F.; and Prasetyo, P. K. 2012. Collective churn prediction in social network. In *ASONAM*.

Pedregosa, F., et al. 2011. scikit-learn: Machine learning in python. *JMLR*.

Tan, C., and Lee, L. 2015. All Who Wander: On the prevalence and characteristics of multi-community engagement. In *WWW*.

Tausczik, Y. R., and Pennebaker, J. W. 2010. The Psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. and Soc. Psychol.*