Interaction vs. Homophily in Wikipedia Administrator Selection

Allen Lavoie
Dept. of Computer Science
Rensselaer Polytechnic Institute
Troy, NY 12180
lavoia@cs.rpi.edu

***Abstract*—Less than one in 600 Wikipedia users have ever participated in administrator selection. Only a small fraction of this population participates in any given Request for Adminship (RfA). This paper investigates reasons for participation. We confirm a strong community effect in the RfA process. Further, direct interaction, not homophily, is key in determining which users participate in a given RfA. Small, interaction-selected groups of editors deciding RfAs raises questions of manipulation and fairness. More broadly, large-scale decision making, where many people have an opportunity but not an obligation to participate, can be heavily influenced by small-scale social factors. Simple mechanisms to increase participation can make these processes far more robust.**

I. Introduction

Wikipedia is a fascinating experiment in consensus based administration. Despite having millions of registered users, any user is free to participate in nearly any decision regarding content or governance of the encyclopedia. Clearly not every user does participate in every decision; nor is this a desirable situation. For many decisions, self-selecting participants seem like a logical choice: users interested in dogs should determine which content is appropriate for dog-related pages. If the Wikipedia users who are interested in dogs can reach a consensus, adding users who are primarily interested in lizards is unlikely to change anything. However, not every decision is a matter of content.

Which group of users best represent community consensus with regard to selecting administrators? Any user is free to join the consensus-building process during a Request for Adminship (RfA), and yet only a tiny fraction do. We[1] show that this small fraction is composed primarily of users who have directly interacted with the candidate. Further, this level of interaction is not explained by homophily; administrators are selected by their contacts, not by similar users who happen to have had contact with the candidate.

Homophily is a desirable property of RfA participation: who better to evaluate a candidate than someone with similar interests? Participants who have interacted with a candidate are also important, since user interaction is a significant part of an administrator's responsibilities. However, interaction currently dominates the RfA process. This is especially concerning considering that vandal fighting and participation in arbitration negatively affect the outcomes of RfAs [4]; when RfAs are decided by small groups that the candidate has interacted with, a few negative interactions may carry undue weight.

While Wikipedia is not a democracy, its consensus-based governance is not immune to apathy by potential participants. For self-governing online communities, simply allowing participation may not be enough to ensure robust decision making. Instead, even highly active communities can benefit from structured participation in governance, using deliberate sampling to offset the negative effects of social networking on the diversity of participants.

[1]We apologize for the excessive nosism in this paper.

II. Related work

There is a significant body of work on the Wikipedia promotion process [10], [5], [4], [11]. Burke et al. [4] use several macro-scale factors to predict the outcome of RfAs. They find some similarities and some significant differences between stated Wikipedia's stated promotion criteria and actual predictors of successful RfAs. Leskovec et al. [10] investigate RfAs as a sequential voting process, examining direct voter-candidate contact, herding, and several social network effects. They find a strong relationship between direct user-candidate contact and whether that user casts a positive vote. Leskovec et al. [11] investigate triadic closure on social networks, including RfA voting. Cabunducan et al. [5] use features of the social network of RfA participants to predict voting and election outcomes, finding that a user's contacts are a good predictor of RfA voting. This paper seeks to determine the relative roles of homophily and interaction in these findings. More communication is correlated with positive voting [10], which seems to indicate a strong true interaction effect. On the other hand, a user's contacts predicting their RfA voting behavior [5] might indicate that homophily also has a significant effect.

Kittur et al. [9] investigate conflict and coordination mechanisms on Wikipedia, which often involve consensus-based processes. Goldspink [7] studies social influences on Wikipedia governance, finding that social norms and rules play a surprisingly small role in content disputes. Yuan et al. [14] find that homophily played a significant role in the adoption of a recommendation system for Wikipedia editors.

Aral et al. [1] differentiate homophily and influence in product adoption, and we use a similar methodology in this paper. Several others study homophily and influence in online social networks; in product reviews [2], on financial message boards [8], on Twitter [12], and on two social networking sites [3]. This paper also examines homophily in an online social network, but with the goal of determining its effects on the robustness of consensus processes built on this social network.

III. Data set

We examine a set of 178,000 votes by 8,800 users cast in 3,700 RfAs between January 2004 and February 2011. Although we refer to these as votes, they are more accurately described as individual contributions to consensus building, self-described as either positive, negative, or neutral. In practice, the descriptions are tallied and used to decide the outcome of RfAs; in effect, they are tallied as votes.

This RfA data set is accompanied by meta-data about each user's contributions to Wikipedia. Specifically, we use the date of a user's edits and the pages they've edited to determine interactions between users. Further, we make use of page meta-data to determine pages similarities: this includes category, link, and user data.

IV. Evidence for a community effect

Editors participate more frequently in RfAs that their contacts participate in [5]. To what extent does voting occur within well-defined communities?

A. Interaction based clustering

To investigate this question, we construct a clustering of RfA participants based on their interactions with other participants. We define an interaction as two users editing the same page within a six hour period. This is not a perfect indicator of interactions. In the case of talk page edits, the period between edits is irrelevant; an interaction has occurred regardless. For more heavily edited pages, it is quite possible that two users edit the same page within 6 hours and yet do not directly interact. Nevertheless, 6 hours is a good compromise between stringent and weak indicators of interaction.

Define a graph with each node representing an RfA participant (a voter or candidate). Let each edge weight be the number of interactions between any two participants. We use a spectral embedding based on the unnormalized Laplacian[13], followed by k-means to construct clusters. This clustering is in no way informed by who voted for whom, and is only related to the RfA process because we have limited the nodes to RfA participants. This clustering may find communities of RfA participants who communicate regularly; we would like to determine what effect these communities have on RfA participation.

B. Voting within and between clusters

When analyzing the clustering of RfA participants with respect to voting patterns, we are interested in finding surprisingly high rates of intra-cluster voting. However, we can easily maximize intra-cluster voting by including the entire graph in a single cluster. Therefore, we compare intra-cluster voting in our clusterings to a random graph model with the same size clusters.

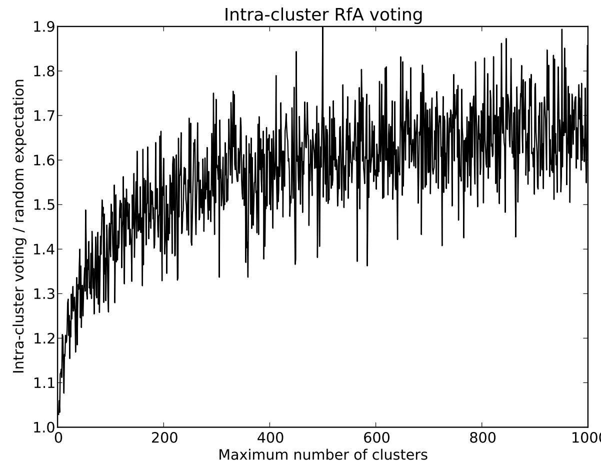

Fig. 1. Intra-cluster voting divided by the random expectation for various numbers of clusters, where clustering is based on user interactions. Much more voting occurs inside clusters than we would expect from random clusters of the same size.

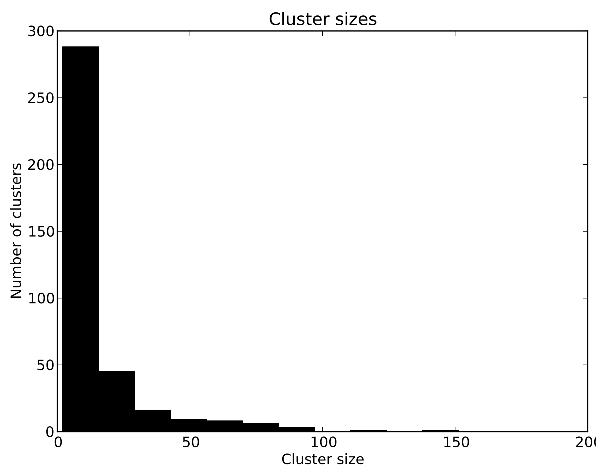

Fig. 2. Cluster size distribution for a clustering of RfA participants with 378 non-empty clusters.

For each cluster, let p_i be the fraction of RfA participants in that cluster. If we assign users to these clusters randomly, the probability of any two participants being in the same cluster is $p_{same} = \sum_i p_i^2$. Distributing votes randomly, this is the probability that any given vote will fall within a cluster (both the voter and the candidate being in the same cluster). Let the total number of votes cast be v. Then the expected number of intra-cluster votes under this model is vp_{same}.

Figure 1 shows the number of observed intra-cluster votes divided by the random clustering expectation for different cluster sizes. Note that the x-axis is the maximum number of clusters requested from the k-means algorithm, and not the number of non-empty clusters produced. This graph clearly shows a community effect in RfA voting; clustering based on interactions yields 60-70% more intra-cluster votes than we would expect from a similar random graph model.

It is useful to examine one clustering more closely. A clustering with a maximum of 700 clusters produced 378 non-empty clusters, with an average cluster size of 17 users. Cluster sizes roughly follow a power law distribution; see Figure 2. Not shown on this plot is one exceptionally large cluster of 1200 users. This large cluster could represent a “core” set of RfA participants who interact with each other frequently, many administrators themselves; 50% of the users in the 1200 user cluster had successful RfAs, while only 30% of users from other clusters where administrators.

V. Interaction vs. homophily

Given that there is a strong community effect in RfA participation, we would like to determine causes. Wikipedia has some explicit community structure surrounding different topic areas; WikiProjects are composed of editors interested in one well-defined topic area, for example. Given this structure, homophily is a reasonable explanation for the observed community effect. Given an RfA candidate focused on military history, it would be surprising if other users interested in military history did *not* comment on the candidate’s contributions. In this sense, homophily is a rather innocuous explanation for the observed community effect; it is nearly unavoidable.

Another explanation for the community effect is an explicit social network, formed through direct interactions between candidates and potential voters. Of course, interaction is closely intertwined with homophily; users both interested in the same topic are far more likely to edit the same pages. Pure interaction not due to homophily is a more concerning explanation. Under this explanation, not only is Wikipedia administration selected by very small groups of editors, but these editors are the candidate’s own contacts.

A. Measuring homophily

To address the question of homophily vs. pure interaction, we first need a measure of similarity between users. A user’s public edit history provides a rich source for this measure, comprising all of a user’s interests and activities on Wikipedia. We propose average pairwise similarity (APS) as a measure of user similarity, useful in testing for the effects of homophily. Let S_1 and S_2 be the set of pages edited by user 1 and user 2 respectively. Define $e_i(s)$ as the fraction of user i’s edits on page s. w_{st} is a measure of the similarity between page s and page t. Then APS for two users i and j is defined as follows:

$$\mathrm{APS}(i,j) = \frac{\sum_{s\in S_i}\sum_{t\in S_j} w_{st}e_i(s)e_i(t)}{\sum_{s\in S_i}\sum_{t\in S_j} e_i(s)e_i(t)} \qquad (1)$$

APS computes an overall topical similarity, not an edit set similarity. Being a weighted average, if $w_{st} \in [0,1]$ then so is APS. Note that $\mathrm{APS}(i,i)$ will not be 1 in general. To make this property hold, we could compute a score for pages $S_i \cap S_j$ and $S_i \cup S_j - S_i \cap S_j$ separately, computing a pairwise score only among the latter. However, $|S_i \cap S_j|/|S_i \cup S_j|$ is quite small even for similar users, as any given user contributes to a very small fraction of Wikipedia pages. While topical self-dissimilarity is counter-intuitive, $\mathrm{APS}(i,i)$ will only be low if pages in S_i are dissimilar, meaning that i’s edits are not topically coherent. An intuitive measure of similarity between two topically incoherent users is somewhat paradoxical, and is tangential to this investigation besides.

Voter		Candidate	
36	Limnodynastes dumerilii	127	Frog
33	List of Anuran families	118	Cane toad
22	Green and Golden Bell Frog	47	List of Anuran families
19	Whistling Tree Frog	24	Litoria
16	Revealed Frog	22	Central Bearded Dragon
16	Smooth Toadlet	15	Australian Green Tree Frog
15	Red-crowned Toadlet	14	Animal
14	Leaf Green Tree Frog	14	Vocal sac
13	Blue Mountains Tree Frog	13	Fauna of Australia
13	Cane toad	13	List of Lacertilia families

TABLE I
TOP ARTICLES EDITED BY THE MOST SIMILAR (VOTER, CANDIDATE) PAIR.

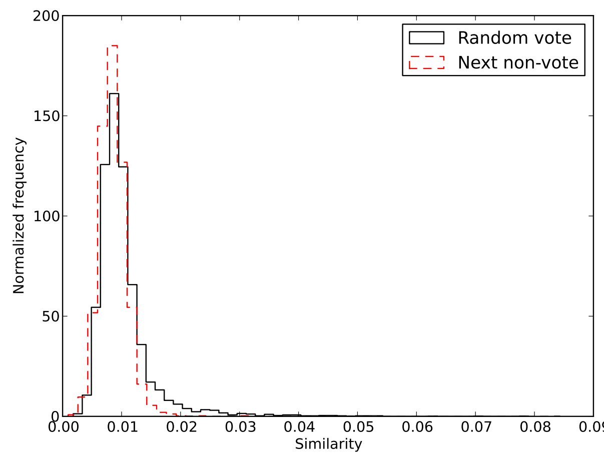

Fig. 3. The frequency of various levels of similarity between voters and RfA candidates in a sample of random votes, and the next RfA following that vote in which the same user did not participate. Users are more similar to candidates whose RfAs they voted in.

We have defined APS in terms of page similarity w_{st}. In this paper, we use the similarity measure defined in [6]: an average of the Jaccard coefficients for the sets of users who have edited either page, incoming links, outgoing links, and categories. Thus, $w_{st} \in [0,1]$ and so is APS.

Table I shows one example of users with very high APS. Given their extremely similar and focused topic areas, it is not surprising that one would participate in the other’s RfA.

Figure 3 shows the distribution of APS scores for (user, candidate) pairs where the user participated in the candidate’s RfA, and for those who did not. The AUC of APS when predicting which pairs had RfA participation is 0.6; APS is at least somewhat predictive, but could simply be predicting interaction.

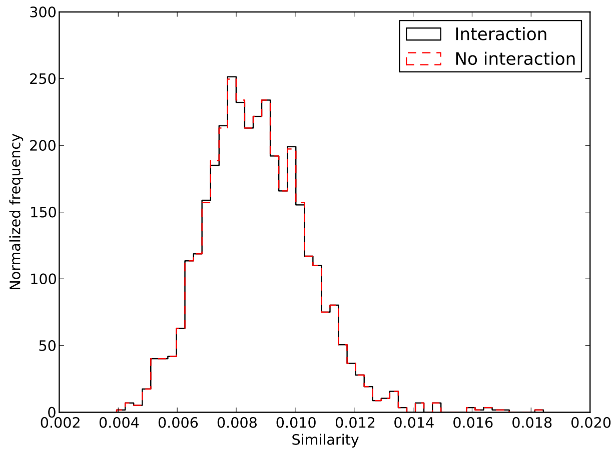

Fig. 4. The frequency of different levels of similarity between users and RfA candidates in the matched sample; the distributions are nearly identical by construction. A (user, candidate) pair who had interacted was matched with another (user, candidate) pair with a similar APS who had not interacted.

B. Matched sample

In order to determine the relative influences of homophily and interaction, we use a matched sample with a similar goal to the matched sample of Aral et al. [1]; here we use a less complex single-dimensional matching based on APS. We construct a matched sample of (user, candidate) pairs with the same levels of similarity, differing only in that one pair has interacted while the other has not.

In the matched sample, homophily-induced RfA participation predicts that the fraction of (user, candidate) pairs where the user participated in the candidate's RfA is the same for pairs who have interacted and those who have not. Interaction-induced RfA participation, on the other hand, predicts that the proportions will be approximately the same as in the un-matched sample.

Without matching, 4029 (user, candidate) pairs interacted and 4897 did not interact. In 76% (95% confidence interval $[74.9, 77.4]$) of the pairs who interacted, the user participated in the candidate's RfA after the interaction. Among the pairs who did not interact, only 28% ($[27.2, 29.7]$) of the users had participated in the candidate's RfA.

The matched sample consisted of 1984 (user, candidate) pairs. Of those who interacted, 73% ($[70.8, 74.7]$) also had RfA participation. Of the non-interacting pairs, only 27% ($[25.1, 29.0]$) had RfA participation. Figure 4 shows the distribution of APS scores for the matched sample.

The matched sample shows that the predictive power of APS in determining which (user, candidate) pairs have RfA participation is almost entirely due to the increased chance that a pair of users with high APS will have interacted at least once. This rules out homophily as a major cause of RfA participation. Although the decrease in the probability of RfA participation given interaction going from the un-matched pairs to the matched sample is statistically significant, the magnitude of this effect is insignificant when compared to interaction effects.

Overall, 69% ($[67.6, 69.9]$) of RfA participants had at least one prior interaction with the candidate. Since only a small fraction of this participation can be explained by homophily, interaction effectively dominates RfA participation.

VI. Discussion

A. Increasing participation

The problems of sample-size and selection are not fundamental to the RfA process. Consider a system like Slashdot's[2] meta-moderation: active users are reminded periodically that they can review other users' ratings, and then given a sample of moderations to rate. Such a prompt for a small number of active but otherwise disinterested Wikipedia users would go a long way towards ensuring the RfA process, and related governance issues, properly reflect community consensus. Automated changes are not necessary; even an informal process for selecting a small group of neutral reviewers would be sufficient.

B. Conclusions

We have differentiated the effects of homophily and interaction in determining RfA participation on Wikipedia. While homophily does play a very small role, its effects are insignificant when compared to interaction. This raises questions about the robustness of Wikipedia's administrator selection process, which is then comprised of a very small interaction-selected group of editors.

This issue is certainly not unique to Wikipedia; low levels of participation are a significant issue in any form of self-governance. These findings indicate that even when huge numbers of users *can* participate in decision making, very few do without a social motive. Luckily, social motives are cheap: just ask!

Acknowledgment

Many thanks to Malik Magdon-Ismail for his helpful comments.

References

[1] S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009.

[2] C.-m. Au Yeung and T. Iwata. Strength of social influence in trust networks in product review sites. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 495–504, New York, NY, USA, 2011. ACM.

[3] H. Bisgin, N. Agarwal, and X. Xu. Investigating homophily in online social networks. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, volume 1, pages 533 –536, 31 2010-sept. 3 2010.

[4] M. Burke and R. Kraut. Mopping up: modeling wikipedia promotion decisions. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW '08, pages 27–36, New York, NY, USA, 2008. ACM.

[2]http://slashdot.org

[5] G. Cabunducan, R. Castillo, and J. Lee. Voting behavior analysis in the election of wikipedia admins. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, pages 545 –547, july 2011.

[6] S. Das, A. Lavoie, and M. Magdon-Ismail. Pushing Your Point of View: Behavioral Measures of Manipulation in Wikipedia. *ArXiv e-prints*, Nov. 2011.

[7] C. Goldspink. Normative behaviour in wikipedia. *Information, Communication & Society*, 13(5):652–673, 2010.

[8] B. Gu, P. Konana, and H.-W. M. Chen. Melting-Pot or Homophily? - An Empirical Investigation of User Interactions in Virtual Investment-Related Communities. *Social Science Research Network Working Paper Series*, Aug. 2008.

[9] A. Kittur, B. Suh, B. A. Pendleton, and E. H. Chi. He says, she says: Conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 453–462, New York, NY, USA, 2007. ACM.

[10] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Governance in Social Media: A case study of the Wikipedia promotion process. Apr. 2010.

[11] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Signed networks in social media. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 1361–1370, New York, NY, USA, 2010. ACM.

[12] S. A. Macskassy and M. Michelson. Why do People Retweet? Anti-Homophily Wins the Day! In *Proceedings of the Fifth International Conference on Weblogs and Social Media*, Menlo Park, CA, USA, July 2011. AAAI.

[13] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17:395–416, 2007.

[14] Y. C. Yuan, D. Cosley, H. T. Welser, L. Xia, and G. Gay. The diffusion of a task recommendation system to facilitate contributions to an online community. *Journal of Computer-Mediated Communication*, 15(1):32–59, 2009.