Social Motivation and Point of View

(Doctoral Consortium)

Allen Lavoie
Washington University in St. Louis
allenlavoie@wustl.edu

ABSTRACT

Social media facilitates interaction and information dissemination among an unprecedented number of participants. Why do users contribute, and why do they contribute to a specific venue? Does the information they receive cover all relevant points of view, or is it biased? The substantial and increasing popularity of social media makes these questions more pressing, but also puts answers within reach of automated methods. I investigate scalable algorithms for finding user behavior changes, predicting the effect of feedback on where users will make contributions, and evaluating the topics and points of view presented in peer-produced content. Users tend to take actions which in the past have led to social interaction, creating herding effects when large groups exchange feedback. In peer production, positive and negative interactions between users can reveal topical disputes, enabling inferences about points of view. Such learning from large-scale social interactions allows us to monitor the quality of information and the health of venues, but also provides fresh insights into human behavior.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]

Keywords

Social Media; Computational Social Science

1. IDENTIFYING BEHAVIOR CHANGES

Viewing a user's contributions to a collective intelligence process as a graph with "knowledge artifacts" as nodes and the similarities between these artifacts as edges, what can metrics such as the clustering coefficient tell us about a user's behavior? Even without an explicit model, this view allows us to quantify a user's concentration on controversial topics[3]. Using readily computed scores, we can identify users who are later blocked from Wikipedia for manipulative behavior, validating their use as indicators of manipulation.

On Wikipedia, administrators can exert significant influence over the encyclopedia through rule enforcement, interpretation of consensus, and social factors. Using the aforementioned measures of manipulative behavior, we can look for suspicious behavior changes after an administrator is elected, which could indicate a user or group misrepresenting themselves in order to gain influence. Some Wikipedia administrators do show suspicious behavior changes.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

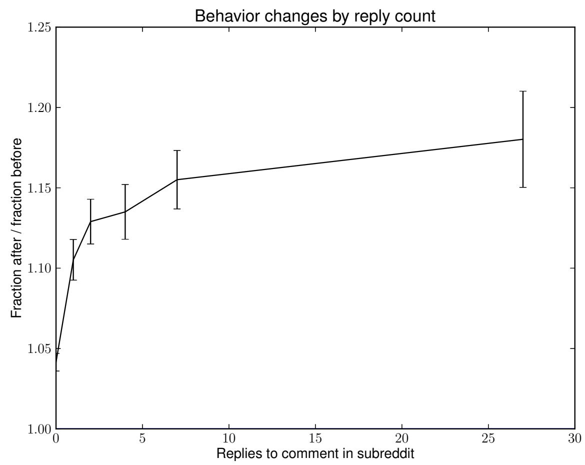

Figure 1: Splitting a reddit user's contributions at a specific comment, the ratio of relative time spent in that comment's community after the comment to the relative time spent in the same community before the comment, as a function of the number of replies received. Users who receive more replies in a community spend more of their time in that community in the future, indicative of learning in response to social feedback.

Administrators are elected via an open voting process. Can we identify administrators who go on to change their behavior, potentially misrepresenting themselves? The popular vote is not helpful: Some who go on to change their behavior significantly receive near unanimous support. However, more sophisticated methods, taking the voting history of participants into account, show that information about who will change behavior is revealed in the voting process. Alternative election processes might harness this collective intelligence more effectively, although more research is needed to find a mechanism that is both politically desirable and effective.

2. FEEDBACK INFLUENCES BEHAVIOR

We can detect extrinsically motivated behavior changes, but what about motivations which are intrinsic to an online social process? For example, we are social creatures: Do social interactions change our behavior? Figure 1 implies that they do. Users who receive more social feedback in a community are more likely to participate in that community in the future.

How exactly do these interactions affect our behavior? I make an analogy to games[2]. We can model decision making in social me-

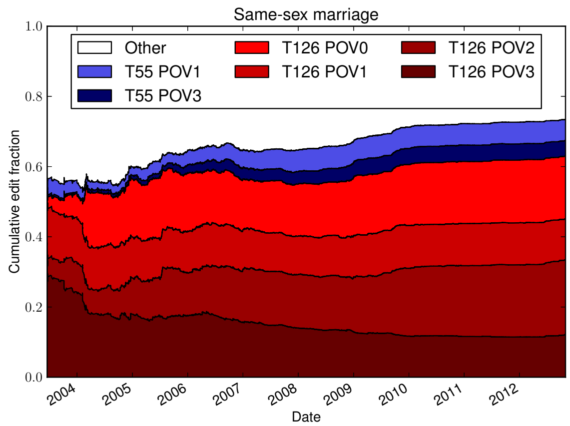

Figure 2: The most popular topics and points of view on the Wikipedia page about same-sex marriage. Topic 126 relates to human sexuality, with point of view 0 generally taking a more conservative stance. As Wikipedia as a whole became more popular, this point of view became increasingly popular relative to others on the page, perhaps reflecting an early demographic shift among editors.

dia as users picking strategies, and the resulting social interaction as being indicative of a reward. Having received a reward, users change their future behavior according to a model of human game playing first studied in economics[4]. With this understanding, we can predict which community a user will pick next.

Which types of social feedback are most motivating? Figure 1 concerns one type of social feedback, comment replies. Another common type of feedback in social media is voting: Other users help to determine a score for each contribution, which is displayed prominently and determines that contribution's visibility. In the game model of social media, which type of feedback yields a greater reward? Exactly how motivating is each? This problem is related to inverse reinforcement learning, but here agents are observed learning from feedback rather than acting optimally to obtain it.

Model parameters are inferred from the contribution histories of several thousand reddit users. These include the learning model itself — the speed of learning and the impact of a user's initial propensities — as well as the relative importance of different types of social feedback. Evaluation is with respect to contributions by these users made after the training set. The model's predictions are probabilistic, and so I estimate the divergence function of a proper scoring rule in order to make principled performance comparisons.

With an accurate model of individual human behavior, what can we say about large groups interacting with and learning from each other? Simulations show a strong herding effect: communities with more participants provide more social feedback to their users, making them more attractive to users and creating a rich-get-richer effect. This effect may extend to competition between social media sites, helping to explain the flow of users between sites as well as between communities within a site.

3. LEARNING FROM INTERACTIONS

Social interactions are a fundamental part of collective intelligence. They mold the behavior of participants, but they can also provide valuable insights about content. By examining social interactions in a rich framework, we are able to learn from online discourse at a large scale[1]. Based on observed actions and disagreements, we can automatically infer the topics, and points of view within each topic, which are being discussed, and the preferences of individual users. Figure 2 shows the points of view inferred on a Wikipedia page. They provide a high-level understanding of the discourse on controversial topics, and can even be used to define controversy. These inferences allow us to find pages on controversial topics which are dominated by a single point of view, and even to suggest editors who may bring a different perspective.

This investigation is enabled by the use of Web-scale data: the full revision history of every English Wikipedia article. This wealth of data creates opportunities for more abstract inferences, but also poses methodological challenges: Scaling simultaneous inferences about millions of Wikipedia users requires an efficient and parallelizable algorithm, which I develop.

Evaluating inferences about topics and points of view at scale is another new challenge. I build a dataset with pairs of users who have antagonistic relationships, using models inferred from user interactions to differentiate these pairs from users who have positive or neutral relationships, a form of domain adaptation. Wikipedia contains structured discourse and rich metadata which facilitate fitting and evaluating such models, but transferring these models to unstructured discourse in social media seems a promising direction.

4. ONGOING RESEARCH

The next question I am interested in combines two lines of my recent research: what role does social motivation play in the points of view we choose to discuss in social media? Rather than choosing a community and receiving feedback in that community, participants choose a topic and point of view and receive social feedback in response to that choice. Behavior changes in response to social feedback are again interesting from a human psychology perspective, but may also provide a richer understanding of group communication dynamics and the evolution of discourse in social networks.

This research will require additional empirical investigation: Is the effect of social feedback on our discourse analogous to its effect on our choice of community? If so, there are important questions concerning the evolution of discourse in social networks, best answered through a combination of theory and simulation with empirical grounding. What happens to minority points of view, and what is the dependence on network structure? How does the evolution of discourse affect the formation of social ties?

5. REFERENCES

[1] S. Das and A. Lavoie. Automated inference of point of view from user interactions in collective intelligence venues. In *Social Computing Workshop (at EC '13)*, 2013.

[2] S. Das and A. Lavoie. The effects of feedback on human behavior in social media: An inverse reinforcement learning model. In *Proceedings of the Thirteenth International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '14, 2014. To appear.

[3] S. Das, A. Lavoie, and M. Magdon-Ismail. Manipulation among the arbiters of collective intelligence: How Wikipedia administrators mold public opinion. In *Proceedings of the Twenty-Second ACM Conference of Information and Knowledge Management*, CIKM '13, pages 1097–1106, 2013.

[4] I. Erev and A. E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Amer. Econ. Rev.*, 88(4):848–81, September 1998.