Evaluating User Activity in Enterprise Social Media

Tad Hogg

January 4, 2010

Abstract
Organizations are deploying social media, such as blogs, to improve internal collaboration. Observing how users react to organizational work practices and company restructuring can help identify factors promoting effective use of social media within enterprises. This possibility is illustrated by two examples involving social media use at a large technology company, to identify how people attend to content and the robustness of communities formed through social media.

1 Introduction
Social media, such as blogs and wikis, used within organizations allow authors to circumvent organizational hierarchies and discover other employees with similar interests and complementary skills. This discovery process can lead to collaboration and knowledge sharing within the organization [5]. Such enterprise social media are especially helpful in connecting people who are geographically or organizationally far apart and thus may otherwise be unlikely to find each other. Thus deploying social media within organizations has the potential to enable diverse collaborations and “crowd sourcing”.

Success with social media requires widespread and continued user participation. A person’s choice to participate reflects a mix of intrinsic motivation [9] and positive reinforcement from the user community [6, 14]. To help identify these motivating factors, enterprise social media provide ready access to user attributes, e.g., job function and location in the organization through an online employee directory [1, 12]. These media also operate in a more trusted environment than the public internet. For instance, the employee directory allows contacting users who stop participating for survey-based studies. For enterprise social media, surveys [17] and aggregate behaviors [1] suggest visible feedback, e.g., in the form of comments from readers, motivate contributions, especially diverse feedback from across the organizations business divisions and geographic locations.

A major challenge for interpreting these observations is determining which relationships arise from causal influences. Knowledge of factors influencing participation can suggest designs and incentives to improve social media effectiveness. The ideal approach to identifying causal relationships is randomized intervention experiments [10]. Such experiments randomly assign users to different settings and observe their social media use. Unfortunately, such experiments are usually not feasible.

A more feasible approach to studying causal influences on technology use arises from “natural experiments” where some change in the organization, unrelated to the technology in question, approximates randomized selection [13]. In particular, the use of social media in organizations offers opportunities for natural experiments, e.g. due to changes in work schedules, reorganization of departments, mergers and layoffs.

We illustrate such experiments with social media use at a large global technology enterprise, Hewlett-Packard (HP). We categorized contributions made between February 2006 and June 2009 into the venues listed in Table 1. We cross-referenced the posts with the employee directory, which indicates where authors work, their job function, and who they report to. Almost all of this content is only accessible inside the firewall and so not visible to those outside the company. Participation in these enterprise social media varies significantly across geographies and job functions [1].
Table 1: Observed participation in various social media venues at HP during the study period. The authors are the number of distinct users contributing to each venue. The last column gives the number of months over which data was collected in each venue.

<table>
<thead>
<tr>
<th>Venue</th>
<th>Authors</th>
<th>Posts</th>
<th>Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blogs</td>
<td>1462</td>
<td>18299</td>
<td>34</td>
</tr>
<tr>
<td>Blog comments</td>
<td>1692</td>
<td>7395</td>
<td>15</td>
</tr>
<tr>
<td>Discussion forums</td>
<td>14625</td>
<td>117807</td>
<td>28</td>
</tr>
<tr>
<td>Wiki pages</td>
<td>525</td>
<td>2456</td>
<td>17</td>
</tr>
</tbody>
</table>

Figure 1: Number of comments as a function of (a) time and (b) subsequent posts since the blog post, for posts made on Mondays (circles) and Fridays (squares). Each point is the number of comments made on a total of 2009 or 2072 posts, for Monday and Friday posts, respectively. The plots show the number of comments on a logarithmic scale.

2 Natural Experiments with Enterprise Social Media

Using the HP data set, we consider situations that indicate how people attend to content and how well social media promote the formation of communities within the organization, especially across geographic regions. The remainder of this section presents two illustrative examples from a portion of the data set.

2.1 Attention Decay: Time vs. Age

As people post new content to social media, older postings receive less attention. This decay arises for two reasons. First, older content is harder to find: content is typically shown in reverse chronological order, so older content appears on subsequent web pages, which users are less likely to visit [4]. Second, older content tends to be less interesting to users as “old news”.

With respect to this second factor, i.e., the decay in novelty, a question for enterprise social media is the extent to which posts are internally focused on the organization vs. externally focused on, say, current news events outside the organization. By comparison with sites such as Digg, a focus on external events is likely to have novelty decay depend primarily on time since the post [15], whereas an internal focus is more likely to be related to its age [2], i.e., number of subsequent posts, which reflects the activity rate within the organization.

In aggregate, these two behaviors are difficult to distinguish: the number of comments on posts shows similar power-law distributions as functions of either time since the post or the age of the post. For enterprise social media, the variation in activity between the work week and weekends can help distinguish these cases. That is, most activity occurs during work days so the less active weekends provide a recurring separation between subsequent number of posts and the elapsed time.

As an example of using this variation in activity, Fig. 1 compares how attention, measured by number of comments, decays for posts made just before and just after weekends. The days with especially low activity
correspond to weekends\textsuperscript{1}. The figure shows the activity in response to Friday posts on the following Monday is more similar to the next day activity for Monday posts than to three days after the Monday posts (i.e., on Thursdays). By contrast, the decrease in comments as a function of age is similar for posts on these two days. The differences as a function of time and similarity as a function of subsequent posts continues beyond the ranges shown in the figures.

By separating posts based on their relation to the weekends, we have a simple natural experiment indicating age is a more significant factor than time in decreasing attention. The enterprise setting also allows more complex comparisons. For example, the employee directory identifies the country each employee works in and company holidays for each country. This information allows comparing the number of comments over a holiday weekend in the US by employees with and without that holiday. Our current data set has too few instances of such cases for meaningful comparison, but as additional data is collected, this comparison will provide a fine-grained evaluation of factors affecting attention decay. With additional data, we could also discriminate between decrease in comments due to reduced visibility of a post vs. loss of interest in that post by examining how often users comment on content given that they view it. Data on views is available from server activity logs, which are readily accessible for social media used within a single enterprise.

2.2 Community Robustness

Social media sites typically have a relatively small portion of the user community contributing most of the content \cite{8, 14}. In particular, major users play an important role in helping people discover others in the organization with similar interests and expertise relevant to their jobs. This discovery can be direct, e.g., through finding an employee who posts on a relevant topic. Alternatively, the discovery process can be indirect through the development of communities among users who find each other by regularly commenting on the same set of posts, such as a popular person’s blog. In this latter case, a significant question for social media is whether these groups become sustainable communities of shared interests, or are the group members mainly participating only to interact with the popular user?

One approach to this question is the persistence of interactions among community members after the popular user stops participating. If the popular user is the main motivation, we’d expect community interactions to decrease significantly after the user stops participating. However, a user who voluntarily chooses to stop participating may reflect, rather than cause, a decline of interest among the readers in that user’s posts \cite{6, 14}. An additional challenge is identifying when a person intends to quit: typically people do not delete their accounts but rather just stop contributing. Gaps in contributions can arise for reasons that do not reflect an intention to quit, e.g., vacations or approaching work project deadlines. These difficulties are avoided when organizational changes result in people leaving the company. Provided such changes are unrelated to users’ participation in social media, the changes act as a natural experiment to examine how feedback maintains contributions and the robustness of communities formed around active users.

An illustration of this technique is when a major user ("Alice") leaves the company due to restructuring. Fig. 2 shows interactions among users who were active both before and after the restructuring, who commented on Alice’s posts and had considerably more direct interactions with Alice (i.e., commenting on her posts) than interactions with each other prior to her departure. In this case, the network defined by interactions remains similar after Alice leaves, including the diversity across geographic regions, which is a major correlate of continued participation in enterprise social media \cite{1}. The similarity of the network before and after Alice leaves suggests the community that initially developed around her contributions has become self-sustaining: once the users involved “met” each other through their common interest in Alice’s posts, they maintained their interactions.

Combining this example with other cases of major participants leaving the company can provide useful insights into the persistence of communities that develop initially through common interest in the activities of a major user. A larger set of examples will allow testing for other correlates, such as overall activity rate in the user community and individual organizational context (e.g., activity by the person’s manager) that could affect community activity independently of the influence of the major users who left. It would also be interesting

\textsuperscript{1}HP employees work in many countries. We use GMT time as an average work time for employees. This time zone usefully distinguishes 24-hour periods when most employees work or are off work for the weekend.
Figure 2: Interactions, in the form of at least one comment on a post, among users who commented at least once on a major user’s posts prior to that user leaving the company due to organizational restructuring. The diagram groups these users by region of their work location: Asia, North America and Europe. In this case, the pattern of activity is similar in the two 3-month periods.

to compare cases where the community around the major user existed for different amounts of time or with differing activity levels. This comparison could identify when a community initially focused on the popular user generates enough other mutual interests to sustain the community even without continued participation by that popular user. A temporal analysis could also identify changes in behavior as the community grows, e.g., through mergers of companies with pre-existing social media covering similar topics.

3 Conclusion

Enterprise social media are a useful tool for investigating user contribution to web communities. The extensive organizational data available on the users (and non-users) allows detailed studies and interventions [16, 17]. These possibilities contrast with social media on the public internet where such specific information on users is often not readily available or restricted due to privacy concerns [7, 11].

Ongoing organizational changes and operational scheduling provide opportunities for natural experiments to help identify causal relationships among participation, feedback, network topology and roles in the organization. One caveat on observations from a single large change to the organization (e.g., large mergers or layoffs) is there may be repercussions on activity beyond those employees directly affected. Thus it is important to examine many such organizational changes, both large and small, for reliable conclusions.

In terms of applications of insights gained from enterprise social media, highlighting information on attention from the community, such as diversity of readership, that correlates with continued participation [1] could help strengthen user communities. Moreover, relating social media use to the organization’s performance measures [16] can identify practical benefits to the organization from encouraging the use of enterprise social media as an informal collaboration tool. More generally, identifying factors influencing attention and the formation of communities of employees with similar interests and skills could improve theoretical models of the performance of communities of practice [3].
Acknowledgments

I thank Michael J. Brzozowski and Thomas Sandholm for helpful discussions and data collection.

References


