ABSTRACT

Traditional social networking sites allow users to enter responses to a set of predefined fields when populating their personal profiles. In the system discussed in this work, freeform ‘About You’ entries allow users to craft their own questions/topics. We found that this kind of flexibility often leads to low content contributions and infrequent updates. The ‘About You’ recommender system described in this paper differs from many recommender systems in that it recommends content for users to create, rather than consume. We present empirical data from an experiment with 2,000 users of a social networking site during a one month period. Our findings suggest that users who receive recommendations create more entries and update them more over time. Further, using articulated social network information for recommendations performed better than content-based matching.

Categories and Subject Descriptors
H.3.3 [Information Search and Retrieval]: Information filtering.
H.3.4 [Systems and Software]: User profiles and alert services.

General Term

Keywords
User profiles, recommendation, social networking, self description, impression formation.

1. INTRODUCTION

Online profiles on social networking sites allow users to create self representations using various text fields such as “Hobbies”, “Favorite Books”, or “School”. Research has shown that these profiles are important for the formation of personality impressions, and that users intentionally craft their profiles to convey a desired impression (e.g. [6], [7], and [25]). Individual fields in a profile can have a positive or negative influence on the perception of an online user [6] and the presence of certain fields is positively associated with the number of online connections [17]. Despite those benefits, the freedom to describe oneself in a diverse way is limited in existing social networking sites which usually support only a limited set of predefined fields.

An alternative, more flexible way for users to richly express information about themselves, is to enter not only responses but to contribute their own questions or topics as well. We implemented these free-form ‘About You’ descriptions as a feature on Beehive, IBM’s internal enterprise social networking site [10]. Our data show that this freedom in describing oneself creates diverse profiles that help in the formation of online connections [5]. Also, users get inspired by others and reuse existing topics for their own profiles, through an explicit “reuse” feature. However, we observed that the average number of About You entries among users with entries is relatively low (mean 3.33, median 2) and only 25.9% of all Beehive users had About You entries in their profiles. Also, our log data and interviews showed that users rarely changed their About You entries once they had created an initial set, i.e. profiles become stale. Some users, when interviewed, indicated that they would be more likely to change those descriptions if prompted to do so.

This paper describes our approach to address the above issues by proactively recommending personalized About You topics to Beehive users. Recommending About You topics is novel and differs from traditional content recommendations in taste-related domains, such as books, movies, etc., in a number of ways: Firstly, we are recommending that users create content, in order to increase participation; whereas in taste-related domains, recommendations are usually made for consumption, i.e. to support the discovery of content matching the user’s tastes. Therefore, users need to put in additional effort to craft an answer when accepting our recommendations. Secondly, the recommended About You topic not only affects the profile owner but the topic and its corresponding answer is consumed by all users viewing the profile as well; whereas a movie or book recommendation is targeted at a single user. Users carefully craft profiles [6] and are likely to be selective in how they choose to describe themselves. Ultimately About You topic recommendations need to predict the personality of a user and the impression a user is trying to create when building a profile. And taste is only one aspect of personality. Finally, although theoretically users can create an unlimited number of About You entries in their profile, they are probably less likely to create a new entry if they already have a significant number of entries. In contrast, a movie or book recommendation is independent of the number of movies watched before.

These characteristics highlight some of the difficulties in recommending this kind of content. The goal of this paper is to gain a basic understanding of the value of these kinds of recommendations in the social media space. We are also trying to understand how characteristics of the social space, such as articulated relationships, i.e. the friends list of a user, can be used
for recommendations in this domain. In particular, we hypothesized that

- recommending About You topics will increase the number of About You entries on personal profiles and the number of users with entries in their profiles,
- articulated social network information can be effectively used to make recommendations for content creation, compared to content-based recommendations.

In order to answer these questions, we designed and implemented an About You recommender system for the Beehive social networking site, offering both social network-based and content-based recommendations. In a preliminary evaluation we asked 100 selected Beehive users to explicitly rate About You recommendations on the site and answer a short survey by email. In a controlled study with 2,000 participants, we deployed an About You recommendation feature on the site to half of the users for a period of 30 days.

The remainder of the paper is structured as follows: After discussing some related work, we introduce the Beehive About You entries in Section 3. Section 4 describes our recommender system in Beehive focusing on the user interface and algorithms used. Section 5 outlines our experimental approach and presents results from the two studies. We conclude with a discussion and directions for future work.

2. RELATED WORK

Social and collaborative filtering techniques for recommendations are becoming increasingly popular. These approaches have in common that they leverage user, community, or user relationship information to make recommendations as opposed to more traditional content-based techniques. Collaborative filtering systems (e.g. [1], [15]), recommend items based on the similarities of users and their preferences. One drawback of collaborative filtering is that users are required to provide preferences in the form of item ratings or similar feedback. Since the cost of obtaining user feedback is high, leveraging implicit user feedback [3] for recommendations has become popular. For example, ASSIST [9] is a social recommender framework that uses interaction history and browsing patterns to make content recommendations. The AnnotatEd [8] system supports social navigation by recommending links based on page visits and annotations. Song et al [24] leverage access patterns to model information flow between users in a social network and use the model to make personalized document recommendations. In contrast to these approaches, our work leverages explicit social information by using the articulated social network of a user on a social networking site.

There is an increasing body of research that uses articulated social network information similar to our work. Seth and Zhang [22] use real-world social network information from Orkut to make personalized recommendations of messages in communities. Their user model is grounded in news media research and focuses on news recommendations, taking into account simplification and diversity features. Groh and Emig [11] use articulated social networks for neighborhood generation in collaborative filtering and show that it outperforms traditional collaborative filtering. Bonhard et al. [2] show that rating overlap and profile similarity can be a powerful source of information for judging the appropriateness of a recommendation. They suggest integrating functionality from social systems with recommender systems. Sinha and Swearingen [23] compare recommendations made by online systems and friends in a lab experiment showing that users prefer friends over online systems.

Existing work can also be organized by the type of content being recommended. Much research has been done in taste-related domains to support finding and discovery of books [18], movies[16], news [22], alerts [21], web sites [19], or links [8]. There is little research on recommendations for creating content for online profiles. Facebook.com has recently launched a people recommender that leverages social networking information but there is no usage data available. Guy et al. [13] introduce a social network analysis service called Sonar. This service aggregates social networking information from various sources including articulated social networks, i.e. friend lists. The service has been recently used in IBM to recommend people to add to your friends list similar to the one in Facebook.

As described in the introduction, recommending topics for self descriptions in online profiles differs from traditional taste recommendations. The primary goal is to increase participation and not to solve an information discovery problem, i.e. we are recommending content for users to create that will later be consumed by multiple people. Similarly, Harper et al. [14] studied personalized invitations containing content to increase participation in online conversations. Cosley et al. [4] show that personalizing recommended editing tasks in Wikipedia increases the number of edits over random recommendations. However, predicting topics for self-description is difficult because we need to predict the personality and intent of a user rather than taste. While research has shown that similarities in users are associated with taste preferences [2], it is unclear whether this also holds true for self-descriptions on profile pages, where aspects of impression management play an important role [7] and users are very selective in the way they describe themselves to a large audience.

3. ‘ABOUT YOU’ PROFILE ENTRIES

Profiles in Beehive [10] include user-created content, a social network of a user (“friends”), comments, and user activity. User-created content includes the freeform ‘About You’ question and answer pairs (called ‘About You’ entries) as well as photos and lists. While many of the About You entries were questions, others had more topical titles, such as “Favorite Book”. Figure 1 shows a detail of an About You section, with real entries taken from multiple users and a fictitious owner.

Users have two ways to create new entries. First, they can add them directly on their profile page. When users add a new entry, they are given two text fields to provide their own question as well as their response to it. Entries are added incrementally and the total number that a user can create is not constrained. Users are able to edit, reorder, and delete entries. The second way of creating new About You entries is to copy existing questions found on other users’ profiles to your profile but provide your own answer. Users can also reuse questions found on list pages with questions others have used.
about Nick
[CICK ON ANY OF NICK'S QUESTIONS TO REUSE IT FOR YOUR PROFILE]
What do I do at IBM?
I'm a software engineer working on a Cognos product.
My first IBM project?
I procured 160 KB floppy disks for the original PC and asked, "What are you going to do with all that space?"
What do I do with my spare time?
[deliberately claim]
My most successful contract negotiation
Convincing my wife to sign that marriage certificate.

Figure 1. Example 'About You' section on a user profile.

3.1 Problems with Freeform Entries
We discovered two problems associated with our freeform profile entries: not enough people created them and those who did tend not to have very many and updated them infrequently.

Our previous research [5] showed that only 25.9% of Beehive users created About You entries, with an average of 3.34 per user (median 2, max 42). We also did some additional analysis before deploying our About You recommender system. For those members who created About You entries, we looked at the amount of time that passed between the first and last entry they created. The average time between the first and last entry was 7.18 days. However, the median was 3.312 minutes. This suggests that many people created most of their About You entries within a short period of time after they signed up for Beehive. In fact, only 12% of our users updated their profile after 1 week of first making it; only 17.9% updated their profile after 24 hours of creating the first entry. In interviews, when one user was asked if he would change his About You entries over time, he said "[there's] nothing prompting me [...] no reason to go change anything [...] I might if it was more exposed."

4. 'ABOUT YOU' RECOMMENDER
To address the above issues, we built a recommender system for About You questions. Our system consists of determining a set of About You questions to recommend to users, the algorithms used to compute recommendations for users, and the user interface used to present the recommendations to users.

4.1 'About You' Candidates
While there were 19,308 About You entries created by a total of 5,788 members of Beehive before the start of the experiment, many of these entries represented duplicate questions. To avoid making duplicate recommendations, we reduced the entries to a set of distinct questions by applying various string manipulation techniques. We converted questions into a set of terms, applied the Porter Stemming algorithm [20], and removed a small list of stop words. Based on this technique, we identified 2,861 distinct questions for the algorithms to recommend. We call these distinct questions About You candidates. Each candidate maps to one or more real About You questions.

4.2 Algorithms
For comparisons, we deployed social Network-based and Content-based recommendations. We also recommended Excluded items that were neither Content nor Network-based. All three algorithms took into consideration the About You entries already created by a given user, and did not recommend the candidates those entries mapped to.1

4.2.1 Content
To determine candidates similar in content, we mined each user’s associated content to create a set of terms. This included

- From Beehive: the status message a user has set, text from their existing About You entries, and the title, description, tags, and any text-based content associated with a user’s shared lists and photos
- From the corporate directory: job title & city the user is located in2.

Similar to the technique we used to create unique About You candidates, we also removed stop words and stemmed terms.

We compute candidate relevance scores $S(u, c)$ for all users with content terms using a modified binary overlap coefficient formula. Let $T_u$ and $T_c$ be the sets of terms for user $u$ and candidate $c$:

$$S(u, c) = \frac{|T_u \cap T_c|}{\text{Min}(|T_u|, |T_c|)} \in [0,1]$$

Because the average number of terms for candidates was much less than the average number of terms per user (3.58 compared to 32.14), we simplified the equation to be:

$$S(u, c) = \frac{|T_u \cap T_c|}{|T_c|} \in [0,1]$$

Of the 22,353 users on Beehive before the start of the experiment, we were able to create terms for 99%. The average number of terms per user for those with terms was 32.14. Content-based recommendations could be computed for 96% of those all users.

4.2.2 Network
The Network algorithm scores a candidate $c$ based on whether or not anyone in the social network of user $u$ created an instance of that candidate. The social network was the subset of Beehive users who user $u$ had explicitly listed as a friend in Beehive. Matching candidates were then shown to the user in random order. Let $N_u$ be the set of unique candidates that are available in the social network of user $u$, then the score $S(u, c)$ is calculated as:

$$S(u, c) = \begin{cases} 
0 & \text{if } c \not\in N_u \\
1 & \text{if } c \in N_u 
\end{cases}$$

Since we did not know ahead of time the possible factors that affect whether or not a user is more likely to accept a

---

1 The recommendation sets created by the three algorithms were similar in word length and recency. However, one interesting effect of the algorithms used was that the Network item set contained more popular items and the Excluded set contained less popular items.

2 In a test with 1,000 randomly selected users, the location city matched at least one candidate for 62 users. State, country, and other location information did not produce any matches.
recommendation, we refrained from creating a more complex score for Network recommendations. Various possible factors include the popularity of the candidate in the user’s network, the last time the candidate was used, or the relationship strength between the user and the person who created the candidate.

Not all Beehive users have friends with About You entries or even an articulated social network. As such, the coverage of the network algorithm is low. In the data set used we were able to create Network recommendations for only 48.4% of all users.

4.2.3 Excluded Recommendations
In order to test our hypothesis we also wanted to see how users responded to items that were not recommended by the Network-based and Content-based algorithms. We recommended Excluded items by removing the candidates that were recommended by either of the other two algorithms, and randomly returned a candidate from the remaining set. This set of recommendations was the largest on average per user compared to Content and Network-based recommendations.

4.3 User Interface
In our experiment, recommendations were shown in three places: a) on the home page after a user logged in, b) on the profile page in the About You view, and c) in periodic email digest messages. The recommendation interface looked identical on the home and profile page, where users were given the About You candidate and two options: “Answer it” (which provided a text entry box for the user’s answer when selected) and “I don’t like it, show me another.” The interface is shown in Figure 2. Upon selecting either choice, the page was refreshed and a new recommendation was shown. In the email messages sent to users as part of their normal daily/weekly updates, we included the personalized recommendation as well as a link that would take them to their home page and show that recommendation, allowing them to create or decline it.

5. EVALUATION
We conducted two experiments: a preliminary evaluation with a smaller group of users and a controlled study in which we deployed About You recommendations as a real feature to a larger group of users on the site.

5.1 Preliminary Evaluation
The goal of the first experiment was to get some preliminary data on the effectiveness of About You recommendations. We wanted to know how Beehive users would rate recommendations from the different algorithms and what they generally think about recommendations of this kind.

5.1.1 Method
We randomly selected 100 active Beehive users based on two criteria. First, they had logged in during the 4-week period preceding this study. Second, we were able to generate at least 12 Network, 12 Content, and 12 Excluded recommendations for each user. We invited each user to rate between 36 and 60 About You recommendations on the site on a customized page. Similar to the recommendation interface described above, the rating page showed one About You recommendation at a time. Every time an item was rated, the page was refreshed and the next item was shown. We conducted a within subject study, i.e. each user received recommendations from all three algorithms. To control for order effects, we presented items in a modified Latin square sequence.

With this group of users, as opposed to our second experiment, we explicitly invited users to participate and these users had an expanded set of rating options. They could rate recommendations as “Good for me” or “Not good for me”. If they checked off “Good for me”, they were also given the option to create an About You entry of that topic on their profile. They were given the option to indicate whether or not they already had a similar item (“Have it”). After a two-week period, we sent an email survey with 16 questions related to About You recommendations.

5.1.2 Results
Of the 100 users selected, 89 logged in during the 2-week period and 76 participated in rating recommendations on the site. 54 users rated 9 or more recommendations, with a total of 2,333 recommendations across all 54 users. Later, 10 of the 76 participating users responded to our follow-up survey by email.

5.1.2.1 Ratings
Figure 3 shows the results for the 2,333 recommendations rated by 54 users. The ratings data suggests that Network-based recommendations were liked most (44.1%) followed by Content-based (36.2%). Only 33.4% of the Excluded recommendations were rated positively.

Figure 3. Acceptance rates in the preliminary evaluation.

These numbers can be considered an upper bound of recommendation acceptance for our algorithms if deployed as a real feature on the site because in this experiment we allowed users to rate items as ‘good’ without having to create real About You entries.

3 We also mirrored the Latin square sequence and randomized the algorithm used to start the sequence across users.
On the other hand, since we asked our users to complete their ratings in a relatively short period of time, we did not expect them to actually create many real About You entries during this experiment. Hence, the number they created for each algorithm can be considered a lower bound for accepted recommendations for a real recommender. 13.2% of all Network, 10.5% of all Content, and 8.6% of all Excluded recommendation resulted in created About You profile entries, as shown in Figure 3.

Users indicated that 27.6% of all Content recommendations were similar to entries they already had, followed by 15.1% for Network, and 7.3% for Excluded recommendations. These numbers reflect how well our question reduction technique (see Section 4) worked for the different algorithms. Note that we did not recommend About You topics that a person already had on their profile. However, while the reduction technique used stemming and stop words, it does not catch similarities such as “My current role at IBM” versus “What do you do at IBM?” Although these two questions have a slightly different meaning, a user with an entry very similar to the one recommended is not likely to use the recommendation. The Content-based algorithm is likely more prone to these errors because it uses word matching of a user’s content (including About You entries) as a technique to find recommendations.

5.1.2.2 Survey
Overall, the users who responded to our survey seemed to enjoy recommendations, describing the feature as “fun,” “useful,” and “helpful.” Only one user did not find the feature useful, as he considered his About You section complete. For those who found the feature beneficial, their reasons ranged from liking the questions “preformatted questions makes things simpler.” to “What do you do at IBM?” Although these two questions have a slightly different meaning, a user with an entry very similar to the one recommended is not likely to use the recommendation. The Content-based algorithm is likely more prone to these errors because it uses word matching of a user’s content (including About You entries) as a technique to find recommendations.

Their responses supported our finding that users seldom change their profiles. When asked how often they added to their About You section before recommendations, 7 out of 10 users gave answers such as “Never”, “Hardly ever,” and “Only when I created the profile.” Only one user responded that he has changed it “several times.” However, the majority thought that the feature would help them add more About You entries overall. One user said, “It’s a feature which clearly would make me change it more often.” Two others felt recommendations provided inspiration; one said: “It helped spark some ideas. Once I got going I came up with a few on my own.”

We received feedback both on recommendations users liked as well as those they did not. The questions users liked were described as “unique”, “not typical”, “a little different or quirky”, “creative – unusual, but still generally applicable,” and “fun, eclectic.” Users also mentioned that these sometimes went beyond what they normally would have thought of by themselves: “that I wouldn’t normally think about sharing or people wouldn’t normally know about me,” and “Provided additional information that I had not thought of.” Questions not liked were generally described as too personal, referring both to things that some users may be less comfortable sharing: “They were more personal than I felt comfortable sharing with the larger IBM audience,” as well as those too personalized for an individual: “The worst were one-offs that meant nothing except to that person and their friends.” Some did not make sense to the users or were considered poorly written.

One user mentioned that he disliked receiving what he considered duplicate recommendations.

Another common theme among subjects was a strong preference towards sharing either personal or professional information in their About You entries. Those interviewed brought up this preference both when discussing recommendations they would not accept and possible “ideal” recommendations. While we did not categorize or recommend questions in this way, this highlights another important aspect in our users’ preferences and could influence the recommender in the future.

We also tried to determine whether or not users felt that freeform About You entries were difficult because they required a user to come up with both the question and the answer, as compared to traditional profiles where users need only provide answers. The subjects were nearly evenly divided on this: 5 said it was more difficult, 4 said it was not difficult. However, both users who said it was difficult, and those who said it was not, answered that they created their profiles with “help” from others: “Yes, it was [difficult]. I ‘borrowed’ a few from other users when I was building my profile.” And “Not very difficult, as I was invited by others, and I viewed his profile so I have [a] reference.” That some users originally looked to those in their network reinforces the finding that the Network recommender performed better in the preliminary results.

While a user’s social network may act as a good filter for About You questions, we did not find strong evidence that knowing whether a question was from one’s network would influence a user’s decision to create an About You. When we explicitly asked subjects whether they would rather receive About You recommendations from within their network of friends (as opposed to those their network does not have), only 1 person said yes. 7 other users had answers ranging from “Both” to “Either” and “Don’t care.” Further, when users were asked whether an explanation of why we recommended certain questions to them might influence their decisions, the responses leaned slightly towards explanations not influencing decisions. Only 2 users said explanations would influence them, 3 said they might be influenced by an explanation, and 4 users said explanations would not influence them. Those who said it would not influence them felt very strongly about it: “Probably not. I see my page on Beehive as being a personal space for me to customize. I will include content because it’s meaningful to me.” And “I think that’s a waste of someone’s time. I’m capable of making that decision alone.” As the biggest factors that influence whether or not a user would create an About You from a recommendation, subjects said, “If I think it represents me,” “Is it different and I feel that answering that item might reveal something about me,” and “Whether or not I believe the information would benefit IBMers who view my page.” We discuss this in further detail in the Discussion Section and hypothesize possible reasons for these results.

5.2 Controlled Study
Encouraged by the results of the preliminary evaluation, we decided to study About You recommendations as a real feature on the site in a controlled experiment with a larger number of users. The study ran for a period of 1 month between April 2, 2008 and May 2, 2008. As outlined in our hypotheses in Section 1, we wanted to study the impact on the number of About You entries and the effectiveness of the recommendation algorithms.
5.2.1 Method
To carry out the study we assigned 2,000 randomly selected users to two groups: control and experimental. When assigning the groups, we controlled for the length of Beehive membership to make sure that the average time of using Beehive in both groups was similar. Like in our first experiment, another condition was that we were able to create at least 12 recommendations from each algorithm for each user.

In the experimental group, we presented recommendations one at a time, from all 3 algorithms, in the same Latin square sequence as in the first experiment. We computed a minimum of 36 and a maximum of 60 recommendations per user. When computing the recommendations, we also found that the overlap of Content and Network recommendations was very small in practice. While it is possible for a candidate to have a non-zero score for both algorithms, less than 0.01% did. Overlapping recommendations were presented to a user only once and in order of the next algorithm in the Latin square sequence.

The control group did not see any recommendations. To control for the extra attention that About You entries get by occupying prime real estate in the experimental group, we advertised About You entries to the control group in the same places in the user interface and via email without making actual recommendations.

5.2.2 Results
Of the 1,000 users in the experimental group, 764 users logged in since the start of the experiment. This is nearly equal to the number of users who logged in to the site from the control group – 755. Of the 764 who logged in, 690 saw the recommendations made to them.4 For the 690 users who saw the recommendations, 288, or 41.7% participated by accepting or declining recommendations.

5.2.2.1 Impact on number of ‘About You’ entries
To understand how recommendations influenced the number of About You entries on users’ profiles and the number of users with entries, we compared the activity related to About You entries in the experimental group with the control group. In the experimental group, 22% of the users added 663 total About You entries during the experiment. In the control group, only 8.4% added new entries, with a total of 169 entries created. We compared the number of About You entries before and after the experiment for those participants with at least one entry. As shown in Figure 4, the experimental group increased the average number of About You entries from 3.9 (median 3) to 4.7 (median 4), while the control group had only a slight increase from 3.9 (median 3) to 4.0 (median 3). The reliable interaction effect (\(F \ [1,1219] = 41.2, p < .001\) indicates that providing About You recommendations was successful in increasing the average number of items.

Another problem was that About You entries are often added by users all in one session and become stale later on. The time between users’ first About You entry and last entry was on average 7.18 days, but the median was 3.312 minutes. We computed this number for our control and experimental groups before and after the experiment to see if recommendations could also be useful in encouraging About You additions over time. Before the experiment, the control group had, on average, a time of 12.86 days between first and last entry, and the experimental group had 11.22 days. However, after the experiment, both groups showed an increase in the number of days between first and last About You (see Figure 5). The control group had, on average, only 16.05 days between first and last About You created whereas in the experimental group, users kept creating their About you entries with an average of 29.71 days between first and last creation. The significant interaction effect (\(F \ [1,1222] = 42.6, p < .001\) suggests that, while the advertisement of About You entries in the control group may have encouraged people to initially add entries, actual recommendations in the experimental group can lead to more entries being added not just in a single session, but over time as well, creating less stale profiles.

5.2.2.2 Effectiveness of recommendation algorithms
One goal of this research was to understand how effectively articulated social network information, i.e. the friends list of a user, can be used to make recommendations for content creation compared to a Content-based approach. Our preliminary evaluation suggested only lower and upper boundaries for acceptance rates.

We examined the recommendation acceptance for those users who responded to at least 9 recommendations, since we wanted at least 3 responses for each algorithm as a minimum. This included 91 users, who responded to a total of 2,125 recommendations. Figure 6 shows the mean acceptance rates per algorithm in our experimental group when deploying these algorithms in a real world setting. Note again,

---

4 It was possible for a user to log in and not see the recommendations made as they were only shown on a user’s home page and the user’s own profile page.
that users had only two possible responses to a recommendation: “Answer it” or “I don’t like it, show me another” and acceptance denotes a user creating an About You entry from the recommendation. The Network-based recommendation algorithm was most effective among the three algorithms. New About You entries were created from 27.5% of our Network-based recommendations on average across 91 users, followed by 20.7% for Content-based and 15.8% for Excluded recommendations \((F[2,180] = 13.3, P < .001)\). Post-hoc comparison (LSD) showed that the results for each condition were significantly different from each other \((p < .05)\).

Figure 6. Mean percent acceptance of recommendations from the three algorithms.

6. DISCUSSION

The data from our controlled study confirms both of our hypotheses. First, recommending About You topics increases the average number of entries per user and the number of users with entries. A time-based analysis also shows that those with recommendations are more likely to update their profile over time.

Although the observed differences are significant, the average increase is rather moderate compared to what we had originally expected. There are a number of possible reasons, for example, the quality of our recommendations could have been a problem, or, given that we make recommendations for content creation, this extra effort might have kept many users from actively participating. On the other hand, considering that asymmetric reading and creating patterns on social networking sites are common \([10]\), a 7.8% increase in users who have About You entries and more frequent profile updates are encouraging results.

Our second hypothesis was that an articulated social network can be effectively used to make recommendations to create content. Our data show that users liked to use About You topics more from within their network compared to Content-based or Excluded recommendations, i.e. they liked to describe themselves similarly to their friends. Bonhard et al. \([2]\) showed that similarity between user profiles influences the choice of a recommendation when information about the recommenders is provided. Similarly, Golbeck \([12]\) found that trust in social networks produces better ratings compared to collaborative filtering approaches. However, it is interesting to note that our experimental recommender interface showed neither the algorithm nor the source of a Network recommendation. Nonetheless we saw a similar effect in that users prefer topics that their friends have used to describe themselves. This may be because they are similar to their friends and are more likely to describe themselves in similar ways.

Interestingly, our survey in the preliminary evaluation did not disclose this particular, possibly subconscious preference when choosing a recommendation. This may be related to the particular domain of About You profile entries. For example, when receiving a movie recommendation from a friend, your opinion of the recommendation might depend on which friend recommended the movie and their taste in movies. By contrast, the About You recommendations contain the complete information, and a user can make a decision based on that alone. Indeed, users felt it was sufficient – citing whether or not the question best represented themselves as the biggest factor in influencing their decisions. Given these findings, it is unclear whether or not showing information about the source of the recommendation could further improve acceptance rates of our Network-based algorithm.

We believe that the acceptance rates of the Network and Content-based algorithms are relatively high given the obvious difficulties of making recommendations for self-descriptions. These difficulties are less technological but rather domain-specific as illustrated above. However, our recommendation algorithms are relatively simple and there is room for improvement. For example, reducing the number of duplicate recommendations by improving our question reduction technique, e.g. by using a thesaurus, could have a positive impact, in particular on the Content-based recommendations where 27.6% of all recommendations were similar to existing About You entries. Manual inspection of the recommendations and interviews revealed that some About You candidates were also overly personal, e.g. containing person names etc. that would lead to bad recommendations.

The algorithms can also possibly be improved by using more sophisticated ranking techniques. Our Content-based algorithm uses a simple binary overlap coefficient which does not take into account the frequency of terms. Our Network algorithm does not use any ranking at all, i.e. a topic just has to be used once within the network to qualify. Some articulated social networks in Beehive are very large (max 1,761). Not every friend in a network has the same meaning and relationship strength. Prioritizing About You topics from closer friends may improve the Network-based algorithm. Also, using the argument of information “completeness” \([22]\), the inclusion of weak ties could possibly further improve performance. To that end, we are currently conducting social network analysis of our data and we are in the process of designing a new Network-based algorithm that takes various factors into account for ranking recommended topics.

Additional work is also needed to better understand how our recommendation algorithms interact with the popularity of About You topics. As described above, the algorithms used in creating the recommendation sets resulted in more popular About You topics in the Network set and less popular topics in the Excluded set. It seems natural that popular items would emerge from the social Network-based algorithm, but this may vary based on factors such as network homogeneity. A challenge also remains to find ways to present less popular items, which may be both distinctive and valuable, in a manner that will encourage use.

A limitation of the Network-based algorithm is that it requires a user to have a social network with friends that have About You entries. As indicated earlier, Network recommendations could be computed for only 48.4% of all users. To solve this problem, we are currently investigating a hybrid algorithm that combines both Content and Network-based recommendations.
7. CONCLUSION
We have demonstrated that social networking information can be effectively used to make recommendations to create content and outperform Content-based recommendations. Unlike traditional recommender work in taste-related domains where the goal is to reduce information overload, our primary goal is to increase participation. Our data show that, About You recommendations significantly increase the amount of profile content created by users and lead to more users contributing to their profile. These results are encouraging as users are highly selective when it comes to managing their online identity.

In this paper, we focused on users who are already active on the site, i.e. they already have an established social network that can be leveraged to make recommendations. Future work will support newcomers to the site, as we have observed that many users sign-up then do not come back and participate. We believe that providing some guidance during the sign-up process, through actively recommending content to create and people to connect to, could address this issue. Making recommendations to newcomers who do not have any content or a social network remains a challenge. We are currently investigating hybrid algorithms that combine randomness but also leverage data external to Beehive.

8. ACKNOWLEDGMENTS
We thank Beth Brownholtz and Joan DiMicco for their support and the fruitful discussions. We also thank all beehive users who participated in our study.

9. REFERENCES