Personalized Activity Streams: Sifting through the "River of News"

Ido Guy, Inbal Ronen
IBM Research Lab
Haifa 31905, Israel
{ido,inbal}@il.ibm.com

Ariel Raviv
Department of Computer Science
Technion, Haifa 32000, Israel
arielr@cs.technion.ac.il

ABSTRACT

Activity streams have emerged as a means to syndicate updates about a user or a group of users within a social network site or a set of sites. As the flood of updates becomes highly intensive and noisy, users are faced with a "needle in a haystack" challenge when they wish to read the news most interesting to them. In this work, we study activity stream personalization as a means of coping with this challenge. We experiment with an enterprise activity stream that includes status updates and news across a variety of social media applications. We examine an entity-based user profile and a stream-based profile across three dimensions: people, terms, and places, and provide a rich set of results through a user study that combines direct rating of the objects in the profile with rating of the news items it produces.

Categories and Subject Descriptors: H.3.3 [Information Search and Retrieval]: *information filtering*

General Terms: Experimentation, Human Factors.

Keywords: Activity Streams, News Feed, Personalization, Real-time Web, Recommender Systems, Social Media, Social Networks, Social Streams, Twitter.

1. INTRODUCTION

The emergence of Twitter and Facebook marked a new phase in the evolution of the web, which is characterized by streams of updates and news. Millions of users who share their activities with their friends and followers end up creating a 'fire hose' of data that reflects hot topics and trends and allows real-time news spreading [19]. This evolvement is often referred to as the *real-time web* [11].

Twitter, the leading microblogging service, allows users to publish short messages (often called "status updates" or "tweets"), describing their activities and opinions or pointing at interesting content. The Twitter stream is homogenous in the sense that it consists solely of status updates. In contrast, typical *activity streams* are heterogeneous and syndicate various types of news and activities within a social network site (SNS) [4].

RecSys'11, October 23–27, 2011, Chicago, Illinois, USA.

The most well-known activity stream is the Facebook News Feed, whose introduction marked a major change on the site [20]. The News Feed occupies the central part of each user's Facebook homepage, showing friends' recent activities, including status updates, friend additions, group joining, page "liking", profile changes, photo sharing or tagging, and more.

Following Facebook, other leading SNSs, like MySpace and LinkedIn have added activity streams of their own. In addition, third-party services such as FriendFeed[1] started aggregating activity streams from various social media websites. Activity streams have also been featured by enterprise SNSs, such as Yammer[1] and Chatter[1]. This proliferation has recently led to the creation of the Activity Streams format[2] for syndicating social activities around the web, which has been adopted by leading sites, such as Facebook and MySpace.

The flood of news updates within activity streams poses new challenges in terms of filtering and personalization. Bernstein et al. [3] interviewed users of Twitter and found that they struggle to balance the promise of interesting content with the sheer volume of incoming updates. Currently, the default filtering of Twitter and Facebook is based on the list of individuals the user chooses to follow ("followees") and the list of the user's friends, respectively. However, this filtering approach is often insufficient, as some friends or followees may produce many non-interesting news items or dominate the stream, while interesting updates may also come from sources outside the circle of friends or followees. Facebook has recently introduced a default view to the News Feed, called "Top News", which presents the user with the presumably most interesting recent news [24]. However, little is known about the methods used for selecting the top news and, to the best of our knowledge, no research has been published to evaluate them. In this work, we examine the challenge of personalizing the activity stream. We refer to it as a recommendation task, aimed at suggesting relevant news items to the user from the overall stream. We refer to a *news item* as the basic unit of which the activity stream is composed. A news item may refer to a network activity (e.g., adding a friend), an activity over an 'entity' (e.g., editing a wiki page, "liking" a file), or a status update.

Several studies have examined the recommendation of social media *entities*, such as blogs, wikis, or communities [1,8,14,29]. Recommendations are typically based on the inter-relationships between users and entities: the content-based (CB) approach recommends entities that have similar content to ones the user

[1] www.{friendfeed, yammer, chatter}.com

[2] http://activitystrea.ms

has 'preferred' in the past, thus building on user-entity and entity-entity relationships; collaborative filtering suggests entities that individuals who are similar to the user preferred, thus building on user-user and user-entity relationships. Some of the algorithms take advantage of the new forms of metadata that social media introduces, such as direct user-user relationships as reflected in SNSs, or tags that are used to annotate content. Ultimately, the user profile used for recommendation is built based on the relationships among users, entities, and their corresponding attributes. We refer to this kind of user profile as *entity-based profile* (*EBP*). Effectiveness of social media entity recommendation using EBPs has been found to be high [1,8,15].

Recommending **news items** differs from recommending **social media entities** in many ways. Most of the differences stem from the higher importance of item freshness in the news item recommendation task. News items are typically considered interesting for only a short period of time after being published, and hence are more exposed to the cold start problem of new items [27]. Furthermore, the occurrence frequency of news items is much higher than of entities. For example, after a shared file is created, users can edit it, "like" it, or comment on it multiple times. Thus, news items are likely to be noisier, in addition to being sparser in content as compared to entities. Entity and news item recommendations also differ in their usage scenario. While entity recommendations expose the user to content and thus increase engagement on the site [9], news item recommendations are targeted at informing users about activities or events, thus increasing social awareness [23].

In this work, we experiment with the activity stream of an enterprise social media application suite. The stream consists of status updates as well as activities across the different social media applications, which include social bookmarking, file sharing, blogging, communities, wikis, and an SNS. For example, a news item can be: "*John Smith edited the wiki page Design Principles in the Cloud Computing wiki*". Previous work has examined the recommendation of mixed entities across these applications (bookmarks, files, blog entries, communities, and wikis). An EBP that included the user's related people and tags was shown highly effective for recommending such entities [15].

We set out to explore whether the stream data, dynamic and noisy as is, can be used for building a user profile for the news item recommendation task. We hypothesized that in spite of these drawbacks, the high frequency and freshness of the news items composing the stream can be beneficial for inferring the user's most recent interests. We refer to a profile built based on the stream itself as a *stream-based profile (SBP).* In addition, we observe that for the recommendation of news items, the entities themselves may be part of the user profile, since multiple news items can relate to the same entity. We therefore examine a third dimension as part of the user profile; in addition to people and terms, we take into account *places*, which include entities such as blogs, wikis, communities and files. We hypothesized that a user who is related to a place may be interested in further news items that originate from it (e.g., multiple updates, comments, and/or "likes" of a file she created).

Our evaluation is based on a user study with 126 participants. We compared the three profile dimensions: people, terms, and places, while experimenting with both an EBP and an SBP. The study included two phases: the first phase directly evaluated user profiles and the second evaluated news items that were produced by these profiles. We also inspected the relation between participants' input for each of the phases.

We evaluated the effectiveness of a user profile for the purpose of news item recommendation using two measurements. *Accuracy* measures the percentage of items rated interesting out of all items recommended to the user (averaged over all participants). While accuracy is often used as a sole measurement for the success of recommender systems, it has been argued that on its own it is not sufficient for evaluating effectiveness [22]. In the case of news item recommendation, freshness of the items becomes highly significant. Hence, a user profile that can generate very accurate news may still not be very effective if the news items appear at low frequencies. Our second measurement is thus *throughput*, defined as the number of news items the user profile produces over a given period of time (averaged over all participants). Higher throughput allows more flexibility in satisfying news consumption rates, yet if accuracy is low, the profile is not useful either. An effective profile would therefore combine high throughput with high accuracy and balance these two, naturally trading-off, qualities. In a sense, accuracy and throughout are analogous to the well-known information retrieval measures precision and recall [21].

The main contributions of this work are: (1) Distinguishing the challenge of personalized news item recommendation from "regular" recommendation of social media entities and providing a comprehensive evaluation over a heterogeneous activity stream; (2) Suggesting *throughput* as a measure for the effectiveness of news item recommendation, which should be used in conjunction with accuracy, (3) Showing that a stream-based profile is effective for personalizing the activity stream, and suggesting future enhancements, (4) Showing that the inclusion of 'places' in a user profile can be highly effective for news item recommendation, and (5) Characterizing the differences among people, terms, and places, in terms of accuracy and throughput.

2. RELATED WORK

The rising popularity of social media has established the need for personalization and recommendation systems in this domain. Systems for recommending people [13], communities [8], blogs [1], bookmarks [29], questions within a forum [26], and mixed social media items [14,15], have been studied in recent works. These recommender systems suggest social media entities to the user, typically based on a user profile built upon social media data, such as community membership, tags, or social relationships within an SNS. In this work, we focus on the recommendation of news items, rather than social media entities, and use an entity-based profile as a baseline for our evaluation.

The emergence of Twitter introduced microblogging as a new form of social media that allows users to write status updates. The highly intensive stream of updates poses new challenges in terms of filtering and personalization, due to its typical noisiness and content sparsity. Recently, a number of studies have suggested ways to cope with these challenges. For example, Bernstein et al. [3] point at three factors that drive satisfaction from reading individual tweets: topic relevance, tie strength, and serendipity. They introduce Eddie, an alternate Twitter interface that empowers topic-based browsing. Chen et al. [7] study URL

recommendation on Twitter as a means to better direct user attention. They evaluate several algorithms through a controlled field study and conclude that both topic relevance and social voting are useful for recommendations. Several works have examined the use of the Twitter stream for user profiling. Phelan et al. [25] suggest a method that promotes news stories from a user's favorite RSS feeds based on Twitter activity. They apply a CB recommendation technique by mining terms from both the RSS feeds and the Twitter messages. Esparza et al. [11] examine CB movie recommendation by harnessing the messages from a Twitter-like service that allows users to express their views on a variety of product types. They find that microblogging messages can serve as a useful recommendation signal. Finally, Hannon et al. [16] focus on suggesting people to follow on Twitter. Their "Twittomender" is based on common followers, followees, and topics extracted from the content of users' own messages, as well as their followers' and followees' messages.

These studies all focus on Twitter, which is a homogenous stream consisting of status updates only. Facebook's News Feed, on the other hand, is a heterogeneous stream of different news item types [17,28]. Other heterogeneous activity streams have also emerged, for example, FriendFeed, which consists of friends' activities on other social media websites, such as blogs and microblogs, social bookmarking systems, photo and video sharing sites, and SNSs [6,12]. The literature on these heterogeneous activity streams is sparser than on Twitter and little is known about their filtering and personalization systems.

Freyne et al. [10] present an initial study over a heterogeneous stream. They suggest narrowing the feed of the SocialBlue enterprise SNS based on person relevance and action relevance, inferred based on users' browsing behavior on the site. In this work, we consider further factors that affect relevance, such as the content of the news item and the place it refers to.

3. SYSTEM DESCRIPTION

3.1 Lotus Connections and the River of News

In our work, we use the activity stream of Lotus Connections (LC) [18] – a social media application suite for organizations that includes various social media applications: an enterprise SNS that allows employees to tag and connect to each other; a microblogging service that allows posting of status updates and replying through board messages; a bookmark application that enables employees to store, share, and tag intranet and internet pages; a blogging application that facilitates the creation of blogs; communities that contain forums for creating and replying to discussion threads; a system for file sharing; and a wiki system that allows co-editing of pages. Commenting and "liking" can be performed on blogs, files, and wikis.

LC publishes an activity stream of all public actions occurring in its applications, called the "River of News" (RoN). Each news item includes a textual description with the action, actor(s) and entity(ies) involved, and occasionally a short excerpt of the text. Figure 1 illustrates the user interface of the RoN. News items are displayed in reverse-chronological order with an indication of their freshness. Each item includes a picture of its actor and an icon indicating the originating LC application. Each underlined entity within the news item content is a link to its corresponding LC page. Metadata of the news item includes the unique LC ID of each of these entities.

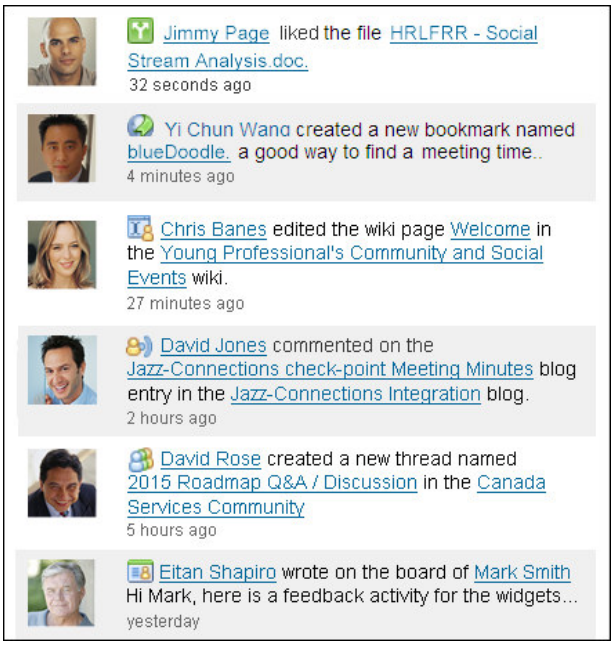

Figure 1. The "River of News" User Interface.

Table 1 details the different types of news items, including all possible actions, and their occurrence frequency during one month of data collection within our organization (percentage of items out of the entire activity stream). Overall, the stream included 139,691 news items during that month, which average out to 4656 news items per day.

Table 1. News items included in the River of News

News Item Type	Actions	% Occurrence
Status updates	create	9.33
Status replies	reply on own board, reply on other board	7.57
Network activities	add to network, tag another person	13.58
Wikis	create, edit, delete, comment, like	33.15
Files	create, edit, delete, comment, like	12.1
Blogs	create, edit, delete, comment, like	6.8
Bookmarks	create, delete	7.12
Communities	create, delete, reply	10.35

3.2 Entity-based Profile (EBP)

The EBP extends the Social Networks & Discovery (SaND) user profile that has been successfully used for recommending entities within LC, such as blogs, bookmarks, communities, files, or wikis. The SaND profile is described in previous work [15] and we provide a short summary in this subsection.

The SaND profile is based on the SaND social aggregation system, which models relationships among people, entities, and tags, as extracted from the different LC applications. SaND builds a relationship matrix that maps a person, an entity, or a tag to all related people, entities, and tags, weighted according to their respective relationship strength. For example, a person

can be directly related to an entity as an author, member, tagger, or commenter; to another person as a friend, tagger or manager; and to a tag as a user of or tagged by this tag. In addition, indirect relations can be inferred through mutual direct relations to a common person, entity, or tag. For example, two persons are indirectly related if they share a common manager or are co-editors of a wiki. The SaND model is based on data accumulated over the last three years leading up to the study.

Using SaND, a user profile is built that consists of the user's top related people and tags. The list of related people is retrieved by considering both direct and indirect person-person relations, scoring them, and aggregating them into a single person-person relationship strength. The list of related tags is retrieved based on direct person-tag relations only, i.e., tags used by the user and/or applied on her by others. A recommender that is based on this hybrid user profile has been shown to yield a nearly 70:30 interest ratio for the top 16 recommended entities [15].

For the EBP, we extended the SaND profile with a third dimension, called 'places'. As noted before, these are entities in which activity can generate multiple news items. The list of places related to a user includes all places to which she is directly related, including blogs, communities, wikis, and files. We calculate the user-place relationship strength by summing the strengths of all direct relations between them. For example, a person can be related to a file since she both shared it and commented on it.

3.3 Stream-based Profile (SBP)

The SBP is built through the user's own activity stream – all news items that include the user as an actor. As opposed to the EBP, it takes into account every news event or activity and is thus more intensive and dynamic, and can accumulate more data over a shorter period of time. For the SBP, we considered a three-month old stream, i.e., all news items from the last three months leading up to the study.

People. As in the EBP, we extracted people based on both direct and indirect relations. Direct relations consider people who co-occur with the user on the same news item (e.g., adding a friend or writing on someone's board). Indirect relations are inferred based on common places, i.e., places that appear on the user's and the other person's stream. For example, people who replied to the same forum thread the user replied to, or people who liked a file on which the user commented. The overall relationship strength with a person considers all direct and indirect relations between the user and that person.

Terms. The RoN does not include an inherent concept of tags. We thus extracted terms for the user profile from the content of the stream itself. Activity streams are challenging for term extraction since they are quite sparse in content. We used the Kullback-Leibler divergence (KL) [2], which is a non-symmetric distance measure between two given distributions. Our goal was to identify a set of terms that maximizes the KL divergence between the language model of the news item's content and the language model of the entire stream. On top of the KL statistical measure we applied a tag-boost (TB) [5], which promotes keywords that are likely to appear as tags, based on a given well-tagged folksonomy. We used the LC bookmark application folksonomy for this purpose. This method has been found to be highly effective for term extraction in non-tagged or sparsely-tagged domains [5]. The weighted list of a user's related terms was generated by applying KL+TB on the content of that user's stream, after filtering out people's names and reserved keywords (such as 'wiki' or 'create') and stemming.

Places. Places were extracted directly from the user's activity stream. Most news items include a reference to one or more places, however there are several types of news that do not relate to a place (e.g., friend addition, people tagging, or status updates). The relationship strength between the user and the place was determined based on the number of occurrences of the place within that user's stream.

As opposed to the well-established EBP, many of the choices made for building the SBP can be optimized to further enhance its performance. Nevertheless, our goal was to create a good baseline for a stream-based profile over the three dimensions, in order to test whether a profile based on stream data can be productive for personalizing the stream.

The main differences between SBP and EBP stem from the data on which they rely. While both profiles change over time, news items occur much more frequently, which refreshes the SBP more rapidly and renders it more dynamic. For example, consider a user who makes multiple edits to a wiki page she created a few months ago, and then also tags it and "likes" it. These changes barely affect the EBP, but influence the SBP because the user's stream now includes various news items related to the wiki. Therefore, each of these actions connect the user more tightly to other editors of the wiki, the topic(s) of the wiki, and the wiki as a place. Another key difference is the time window. The SBP is built on news from the last three months and therefore reflects the user's short-term interests and relationships. The EBP is built on data from three years, reflecting longer-term interests and relationships.

4. EXPERIMENTAL SETUP

Our evaluation is based on a user survey that included two phases. In the first phase, participants evaluated lists of *profile objects*, i.e., lists of people, terms, and places that make up the profile itself. In the second phase, participants evaluated news items that originated from those different profiles.

For both EBP and SBP, we retrieved the top 10 people, terms, and places, using the methods described in the previous section. The participants of the survey were active users of LC, who had at least 10 objects in each dimension of both profiles (overall at least 10 objects in 6 lists). We sent an email invitation with a link to the survey to each of the 615 identified users and 126 completed the survey. They originated from 23 countries, spanning different divisions: 28% sales, 24% software, 21% services, 20% headquarters, 2% systems, 2% research, and 3% others. We note that this sample does not represent the entire population of our organization's employees, but rather active users of LC, who are the target population for our system.

In the first phase of the survey, each participant was presented with 3 lists of up to 20 people, terms, and places, respectively. Each list was a union of the 10 objects from the EBP and the 10 from the SBP, in randomized order. If an object was included in both profiles, it appeared only once on the list, and was marked as related to both for analysis purposes. Participants were asked to go through each list and choose up to 10 objects for which

they would like to get news. Only after the participants marked between 1 to 10 persons, terms, and places, could they move on to the second phase.

In the second phase, we asked participants to evaluate news items. For each list of 10 objects, corresponding to a profile type (EBP or SBP) and a dimension (people, terms, or places), we produced up to 5 news items in the following manner: We first iterated over the 10 objects according to their order in the user profile (i.e., their relationship strength to the user). For each object, we selected the most recent news item from the last month that did not include the participant herself as an actor (if one existed). We finished the iteration when 5 news items were selected or when the list of 10 objects was exhausted. Thus, it was possible that less than 5 news items were produced per profile. Finally, we unified the 6 lists of up to 5 news items, merging duplicate items (while keeping all their origins for analysis purposes). The final list of up to 30 news items was shuffled to create a randomized order of the items and was presented through the user interface illustrated in Figure 1. Participants were asked to rate each news item as 'very interesting', 'interesting', or 'not interesting'.

5. RESULTS

5.1 Profile Evaluation

We first inspect the overlap between the EBP and the SBP for our survey's participants, over each of the three dimensions. Table 2 shows the statistics referring to the ratio of identical objects within the lists of length 10. It can be seen that the overlap between EBP and SBP is generally very low, with a median of 0.1 for all three dimensions (i.e., 1 identical object out of 10). The average for people is slightly higher than for terms and places. These numbers indicate that the activity stream yields a very different profile from the one built based on entities and further motivate the comparison between the two.

Table 2. Overlap between EBP and SBP for all participants

	People	Terms	Places
avg (stdev)	0.157 (0.119)	0.118 (0.126)	0.11 (0.122)
median	0.1	0.1	0.1

In the first phase of the survey, participants were asked to select up to 10 people, terms, and places, from lists of up to 20. On average, each participant selected 7.4 people (std: 3, med: 8), 8 terms (std: 2.5, med: 9), and 6 places (std: 2.7, med: 6). Figure 2 shows the average number of selected people, terms, and places that originated from the EBP versus those from the SBP. The results are quite different across the dimensions. For places, each participant selected on average 4.02 SBP places and only 2.99 EBP places, which is a significant difference (one-tailed, paired t-test, $p<0.001$). For people, the difference is smaller: on average 4.64 SBP people versus 4.1 EBP people were selected, which is still a significant finding, by a small margin (one-tailed, paired t-test, $p=0.044$). For terms, the results are almost identical, with a slight insignificant difference for EBP over SBP – 4.86 versus 4.8 (one-tailed, paired t-test, $p>0.05$). These results give a first indication that the stream-based method is able to generate a useful profile, which is comparable to the existing entity-based method for extracting terms, and even outperforms it for people. We were somewhat surprised by the people result, since we had assumed that the well-established SaND network, which relies on long-term relationships, would yield more interesting people than the stream-based method, which is based on interaction data of three months only. It appears that the freshness of the data compensates for its shorter range. Participants might have also taken into account the activity level of the person, which is likely to be higher for people with whom they had recent intense interaction. For places, SBP clearly outperforms EBP. As opposed to terms and people, entity-based places have not been evaluated before. The difference could be attributed to the freshness of the stream-based profile, which is an especially important factor for places.

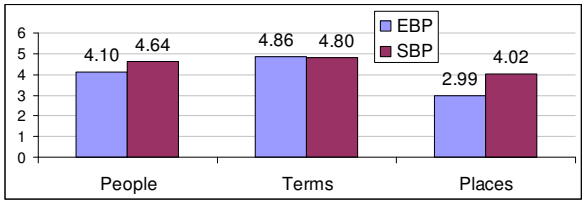

Figure 2. Number of profile objects seclected in phase 1.

Places can be further classified according to the type of social media application from which they originate. Table 3 focuses on a breakdown of the places by their type. The middle columns show the distribution of types (in percentages) across the EBP and SBP of all survey participants (1,260 places per each profile type). The distribution of EBP places sharply favors files and communities, while SBP places are distributed much more evenly. Wikis occupy over 40% of the stream-based places, which can be attributed to the dominance of wiki-related news in the overall stream (Table 1). The sharp difference between wiki occurrence in the EBP and SBP can be explained by the fact that for each saving of a wiki edit a news item is produced.

Table 3. Places occurrence and selection by type

Place	EBP dist.	SBP dist.	% Selected
Blogs	3.73	17.93	33.46
Communities	42.57	24.41	49.03
Files	50.75	17.05	14.54
Wikis	2.95	40.61	28.07

The rightmost column of Table 3 shows the percentage of places selected in phase 1 by place type. Communities are chosen much more frequently than others. Nearly 50% of the communities were selected as interesting for news. About 33% of the blogs and 28% of the wikis were selected, while files have the lowest selection rate – 14.5%. It could be that it was harder to grasp a shared file as a place, even though it allows multiple edits, comments, and likes.

5.2 News Item Evaluation

In the second phase of the survey, participants rated news items that originated from the different profiles. Table 4 compares the ratings of the EBP and SBP across the three profile dimensions. These ratings reflect the accuracy of the profile, i.e., the percentage of interesting items it produces. The differences between EBP and SBP are quite small, despite the fact that they hardly overlap. For terms, the results are almost identical, which

is consistent with the way participants rated them in phase 1. Although SBP people were rated higher in phase 1, the differences in news item ratings were very small. The only statistically significant difference is for the places dimension: SBP places yielded an impressive 30.56% of very interesting items and only 21% non-interesting items, compared to 20.99% and 26.54%, respectively, for the EBP places (one-tailed, unpaired t-test, p<.0001).

Table 4. News item rating results across the 3 dimensions

Dimension	Profile	Not Int.	Int.	Very Int.
People	EBP	41.94	44.25	13.81
	SBP	43.6	40.72	15.68
Terms	EBP	54.25	33.33	12.42
	SBP	55.24	32.77	11.99
Places	EBP	26.54	52.47	20.99
	SBP	20.96	48.48	30.56

Table 4 indicates that there are considerable differences among the rating results of news items based on their respective profile dimension. Figure 3 highlights the rating results across the three dimensions over all news items (regardless of the profile type). Evidently, places yielded the most interesting news items, with 76.6% rated as interesting (including very interesting). In contrast, terms yielded the least interesting items, with less than half (45.2%) rated as interesting. People are in-between terms and places, with 57.2% rated as interesting. One-way ANOVA indicates that ratings across the three dimensions are significantly different ($F(2,2378)=79.73, p<.0001$). Tukey post-hoc comparisons of the three dimensions indicate that all differences are significant.

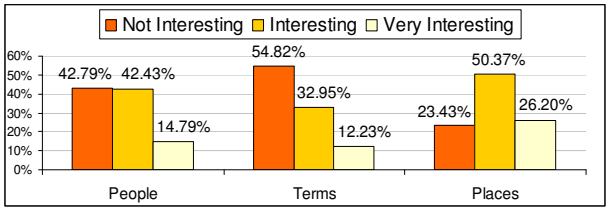

Figure 3. News item rating results by dimension.

As previously noted, interest ratings are not the sole measure for the effectiveness of a user profile. A profile that yields one interesting news item per month, even if it is an interesting one, is not likely to be considered very effective for stream personalization. We thus set out to explore the throughput of the different user profiles, measured as the number of news items they produced in a month. Table 5 depicts the average and median throughput for EBP and SBP across the three profile dimensions for our survey's participants, as measured during the month of experimentation. An average throughput of 100, for example, implies that the 10 profile objects produced together 100 news items that month for the average participant. The rightmost column shows the *throughput diversity*, which we define as the ratio of objects in the profile that yielded at least one news item during the month. For example, a throughput diversity of 0.5 implies that, on average, for each participant, 5 out of the 10 objects in the profile produced at least 1 news item per month. As could be expected, the SBP had a significantly higher throughput than the EBP for all three dimensions (one-tailed, unpaired t-test, p<.0001). When inspecting the accuracy and throughput results, it can be seen that the higher throughput of the SBP does not come at the expense of lower accuracy. In general, the SBP produced significantly more news items, while keeping the same level of accuracy as the EBP. For places, SBP had both higher accuracy and higher throughput.

Table 5. Throughput across the 3 dimensions

Dimension	Profile	Thrpt Avg (std)	Thrpt Median	Thrpt Diversity
People	EBP	193 (419)	42	0.53
	SBP	1043 (986)	740	0.84
Terms	EBP	2837 (2051)	2667	0.82
	SBP	5280 (1482)	5288	0.98
Places	EBP	17 (31)	4	0.26
	SBP	122 (114)	88	0.63

Comparing the throughput results across the three dimensions reveals the natural trade-off between accuracy and throughput. While terms have the lowest accuracy, they also have the highest throughput by a large margin over people and places. On average, for a given participant within a month, the terms in the EBP produced 2837 news items, while the SBP terms produced as many as 5280 news items. Throughput diversity indicates that almost every term in the SBP produced at least 1 news item per month, while for EBP, 82% of the terms produced news (on average). People's throughput is substantially lower than terms and demonstrates a big difference between the two profile types: 193 for the EBP and 1043 for the SBP. At least 1 news item was produced by 84% of the people in the SBP and 53% in the EBP. Throughput of places, which had the highest accuracy, was the lowest, with 122 news items from the SBP and as few as 17 places (median 4) for the EBP. On average, only 63% of the SBP and 26% of the EBP places produced at least 1 news item. The low throughput, as well as the large difference between EBP and SBP, may be explained by the fact that places often have a short "shelf life": they are active for a certain period and then become idle. The high throughput of terms may also result from ambiguity, while people and places are uniquely identified.

Table 6. News Item Rating Results by Selection in Phase1

Dimension	Selected Phase1?	Not Int.	Int.	Very Int.
People	yes	32.2	45.51	22.29
	no	50.3	38.25	11.45
Terms	yes	42.43	38.91	18.66
	no	65.85	27.41	6.74
Places	yes	15.98	49.23	34.79
	no	30.09	50	19.91

We next explore the relation between the two phases by inspecting how the selection of profile objects in phase 1 was reflected in the selection of news items in phase 2. Table 6 depicts the interest ratings for news items that originated from people, terms, or places that were selected in phase 1 compared to those not selected in phase 1. For all three dimensions, the

differences are significant (one-tailed unpaired t-test, p<.0001), indicating that users indeed selected people, terms, and places that yielded more interesting news items for them. The participants' ability to recognize interesting profile objects suggests that allowing them to edit their profile may be a desirable feature in personalized activity stream applications, to complement and fine-tune the automatically-inferred profile.

Conducting a similar comparison for throughput reveals that the average throughput of people selected in phase 1 was higher than of non-selected people at 73.5 versus 49.8. We also found similar differences for places, with 12.3 for selected versus 4.7 for non-selected. For terms, however, the throughput was very similar with 407.6 versus 404.5. These results indicate that participants may have indeed considered throughput when choosing people and places; they were more likely to choose those that would generate more news (in addition to interest considerations). When choosing terms, however, throughput did not seem to have an effect. Because terms generally have a high throughput, participants may have not considered it a critical factor, and might have even preferred to exclude very common and general terms with very high throughput.

Finally, we examine the rating results by type of the news item. Figure 4 shows the rating results for each of the item types listed in Table 1. The percentage in parentheses indicates the portion of the item types out of the entire recommended news items in our survey. It can be seen that the most accurate results are for news items that relate to entities: news items about blogs were the most interesting (68.6% rated interesting or very interesting), followed by communities and bookmarks (61.7% each), files (60%), and finally wikis (51.8%). The lower percentage of wikis can be explained by the fact that wiki is the most common type of news item, with many incremental updates on wiki edits.

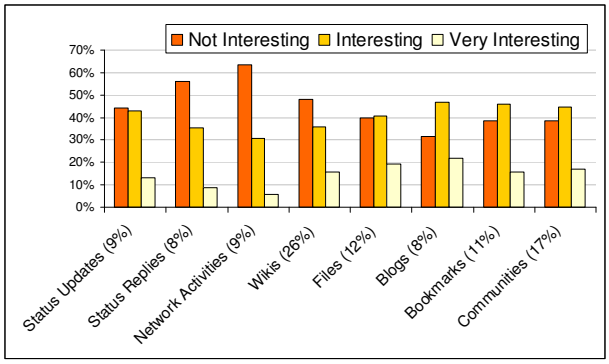

Figure 4. News item rating results by item type.

Status updates also received relatively accurate results, with 55.6% rated as interesting or very interesting (higher than wikis, but lower than all other entities). Interestingly, status replies, both on own board and on others, are considered much less interesting, with only 44% rated interesting or very interesting. It appears that replies are perceived as an interaction between two other individuals, and are of less interest to the public. Network activities were rated as least interesting, with only 36.4% interest. These news items also refer to interactions between two other individuals, but with even less content than the status replies.

6. DISCUSSION AND FUTURE WORK

The results of our study indicate that building a profile based on the activity stream is a productive method for personalizing the stream and demonstrates various advantages over a traditional EBP. This result is not trivial. The stream itself is noisy, dynamic, sparse in content, and reflects shorter-term interests and relationships. It was not obvious that the network and terms extracted from three months of stream data could equal the well-established network built by SaND or the manual tags people use or are tagged with — both of which have previously shown to effectively represent the user's interests and relationships [14,15]. Yet, the SBP we examined was found to have a level of accuracy similar to the EBP, with a significantly higher throughput. Moreover, the diversity of the places dimension and the accuracy of its results were significantly higher for the SBP.

The addition of places to a user profile has shown to be highly beneficial for the news item recommendation task. Places yield very high accuracy — nearly 80% for the SBP. While producing high accuracy, places are also shown to have the lowest throughput (only 122 items a month for the SBP). At the other extreme, terms produce noisy items with low accuracy and high throughput. People are the middle dimension in terms of both measurements. Hybridization of profile dimensions is likely to improve the results even further and allow more flexibility when adapting to the user's needs. Future research should examine hybridization techniques based on the characteristics shown here for each of the profile dimensions.

The accuracy results across the various types of news items are quite different. Network additions and person tagging are of substantially lower interest than status updates and replies, while entity-related news have the highest interest rates. These results indicate that stream personalization methods should account for the type of news, as has been proposed by Freyne et al. [10].

We suggested using throughput together with accuracy, when evaluating user profile effectiveness for stream personalization. The trade-off between accuracy and throughput is well reflected in our results. In addition, the selection of people and places in phase 1 was found to be associated with throughput — users tend to choose people and places that would generate more news. Furthermore, places with higher throughput, like communities, are chosen more often as compared to other types of places with lower throughput, such as files. When composing a user profile, throughput should be taken into account and balanced. For example, terms with extremely high throughput are likely to represent a very general, common, or ambiguous topic, and may thus be less desirable. On the other hand, many items with very low throughput may inflate the profile without adding much to its overall productivity. It should also be noted that different users may prefer very different throughput rates; hence throughput needs to be controlled on a personal basis.

In this work, we focused on recommending an initial set of up to 30 news items, based on a preliminary SBP that was built using 3 months of stream data. We did not examine temporal aspects, such as the dynamics of the SBP and how it changes over time. Places seem to become stale more quickly than people and terms. As part of profile maintenance over time, it may be desirable to remove places that become inactive. Generally, since activity streams are very dynamic, with bursts of places, topics, and people over time, it is essential to keep refreshing the

user profile, removing stale objects and adding emerging ones. This can be achieved by a combination of automatic stream analysis over time with manual profile updates by the user.

One of the interesting challenges with activity streams is the aggregation of several news items into a single one. Aggregation can take place with respect to people or actions, e.g., '*P1 and P2 commented on a blog*', or '*P1 created, edited, and liked a wiki*'. Aggregation may enhance personalization quality, as it decreases redundant news items. Additionally, an aggregated news item may be considered more important than a single one. For example, two people in your network adding someone as a colleague within close time proximity renders a more interesting news item. Our future plans include evaluating aggregation methods and their effect on stream personalization.

We conducted our study on an enterprise activity stream that is built over a social media application suite. We believe that the applications we experimented with constitute a good sample of the most popular social media applications, thus our results are applicable for other large organizations. Since variants of most of these applications exist outside the firewall, our study's implications may span beyond the enterprise. Yet, further research should be conducted on activity streams outside the firewall, to validate and extend the findings of this work. This will also help address challenges such as the higher frequency of news updates and the multitude of activities that may be available through the stream on the web.

7. CONCLUSION

We explored activity stream personalization through experimentation with an enterprise activity stream, which includes a mixture of news items about social media activity. We focused on three user profile dimensions: people, terms, and places. Our results indicate that building a profile based on the user's stream data can be effective for the personalization task, considering both accuracy and throughput. News items that stem from places demonstrate an especially high accuracy, suggesting that the addition of places to the user profile is of high value. We suggest several directions for future research, including hybridization techniques, adaptation over time, and news item aggregation. The personalization task is likely to become even more challenging as the volumes and types of news continue to grow, and further flood the streams and their users.

8. REFERENCES

[1] Arguello, J., Elsas, J., Callan, J., & Carbonell J. 2008. Document representation and query expansion models for blog recommendation. *Proc. ICWSM '08*.

[2] Berger, A. & Lafferty, J. 1999. Information retrieval as statistical translation. *Proc. SIGIR '99*, 222-229.

[3] Bernstein, M. S., Suh, B., Hong, L., Chen, J., Kairam, S., & Chi, E. H. 2010. Eddi: interactive topic-based browsing of social status streams. *Proc. UIST '10*, 303-312.

[4] Boyd, D.M. & Ellison N.B. Social network sites: Definition, history, and scholarship. *J.CMC 13, 1* (Oct. 2008), 210-230.

[5] Carmel, D., Uziel, E., Guy, I., Mass Y., & Roitman H. Folksonomy-based term extraction for word cloud generation. *Proc. CIKM '11*.

[6] Celi, F., Di Lascio, F.M.L., Magnani, M., Pacelli, B., & Rossi, L. 2010. Social network data and practices: the case of Friendfeed**.** *Advances in Social computing*, 346-353.

[7] Chen, J., Nairn, R., Nelson, L., Bernstein, M., & Chi, E. 2010. Short and tweet: experiments on recommending content from information streams. *Proc. CHI '10*, 1185-1194.

[8] Chen, W., Chu, J., Luan, J., Bai, H., Wang, Y., & Chang, E. Y. 2009. Collaborative filtering for Orkut communities: discovery of user latent behavior. *Proc. WWW '09*, 681-690.

[9] Freyne J., Jacovi, M., Guy I., & Geyer W. 2009. Increasing Engagement through Early Recommender Intervention. *Proc RecSys'09*, 85-92.

[10] Freyne J., Berkovsky S., Daly E.M., and Geyer W. 2010. Social networking feeds: recommending items of interest. *Proc. RecSys '10*, 277-280.

[11] Garcia Esparza, S., O'Mahony, M. P., & Smyth, B. 2010. On the real-time web as a source of recommendation knowledge. *Proc. RecSys '10*, 305-308.

[12] Gupta, T., Garg, S., Carlsson, N., Mahanti, A., & Arlitt, M. 2009. Characterization of Friendfeed: A web based social aggregation service. *Proc. ICWSM '09*.

[13] Guy, I., Ronen, I., & Wilcox, E. 2009. Do you know? Recommending people to invite into your social network. *Proc. IUI'09*, 77-86.

[14] Guy, I., Zwerdling, N., Carmel, D., Ronen, I., Uziel, E., Yogev, S., & Ofek-Koifman, S. 2009. Personalized recommendation of social software items based on social relations. *Proc. RecSys '09*, 53-60.

[15] Guy, I., Zwerdling, N., Ronen, I., Carmel, D., & Uziel, E. 2010. Social media recommendation based on people and tags. *Proc. SIGIR '10*, 194-201.

[16] Hannon, J., Bennett, M., & Smyth, B. 2010. Recommending twitter users to follow using content and collaborative filtering approaches. *Proc. RecSys '10*, 199-206.

[17] Hoadley, C.M., Xu, H., Lee, J.J, & Rosson, M.B. 2010. Privacy as information access and illusory control: The case of the Facebook News Feed privacy outcry. *E-Commerce Research and Applications 9, 1 (Jan. 2010)*, 50-60.

[18] IBM Social Software for Business – Lotus Connections: http://www.ibm.com/software/lotus/products/connections.

[19] Kwak, H., Lee, C., Park, H., & Moon, S. 2010. What is Twitter, a social network or a news media?. *Proc. WWW '10, 591-600*.

[20] Lampe, C., Ellison, N. B., & Steinfield, C. 2008. Changes in use and perception of Facebook. *Proc. CSCW '08*, 721-730.

[21] Manning, C., Raghavan, P., & Schtze, H. 2008. *Introduction to Information Retrieval*. Cambridge University Press.

[22] McNee, S. M., Riedl, J., & Konstan, J. A. 2006. Being accurate is not enough: how accuracy metrics have hurt recommender systems. *Proc. CHI '06 Extended Abstracts*, 1097-1101

[23] Naaman, M., Boase, J., & Lai, C. 2010. Is it really about me?: message content in social awareness streams. *Proc. CSCW '10*, 189-192.

[24] Official Facebook Blog: New Views for Your Home Page http://blog.facebook.com/blog.php?post=162536657130

[25] Phelan, O., McCarthy, K., & Smyth, B. 2009. Using twitter to recommend real-time topical news. *Proc. RecSys '09*, 385-388.

[26] Qu, M., Qiu, G., He, X., Zhang, C., Wu, H., Bu, J., & Chen, C. 2009. Probabilistic question recommendation for question answering communities. *Proc. WWW '09*, 1229-1230.

[27] Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. 2002. Methods and Metrics for Cold-start Recommendations. *Proc. SIGIR '02*, 253-260.

[28] Sun, E., Rosenn, I., Marlow, C., & Lento, T. 2009. Gesundheit! Modeling contagion through Facebook News Feed. *Proc. ICWSM*.

[29] Vatturi, P. K., Geyer, W., Dugan, C., Muller, M., & Brownholtz, B. 2008. Tag-based filtering for personalized bookmark recommendations. *Proc. CIKM '08*, 1395-1396.