

# Personalised Network Activity Feeds: Finding Needles in the Haystacks

Shlomo Berkovsky<sup>(✉)</sup> and Jill Freyne

CSIRO, Epping, Australia  
{shlomo.berkovsky,jill.freyne}@csiro.au

**Abstract.** Social networks have evolved over the last decade into an omni-popular phenomenon that revolutionised both the online and offline interactions between people. The volume of user generated content for discovery on social networks is overwhelming and ever growing, and while time spend on social networking sites has increased, the flood of incoming information still greatly exceeds the capacity of information that any one user can deal with. Personalisation of social network activity news feeds is proposed as the solution that highlights and promotes items of a particular interest and relevance, in order to prioritise attention and maximise discovery for the user. In this chapter, we survey and examine the various research approaches for the personalisation of social network news feeds and identify the synergies and challenges faced by research in this space.

## 1 Introduction

Growth of the Web is relentless and set to continue, even accelerate, as the Web continues to evolve and accommodate new forms of user-generated content [30]. Social networking sites have experienced unprecedented popularity in the past decade and have contributed significantly to increased levels of user generated content that have been fuelling this growth. Social networks, designed to allow anyone to create and distribute content for others to consume, have become rich and diverse sources of information that compete with and complement traditional search engines in the diffusion of information.

Social networks allow users to hook up to streams of information, from trusted individuals, and so act as personal filters for online content. In essence, users hand pick the information sources, whose contributions make up their personal information channel or *news feed*. This methodology, where contributed items or actions are conveniently combined and presented in reverse chronological order worked well, allowing individuals to quickly discover updates and content of interest [5]. The popularity of social networks and the ease of sharing content has, however, swamped the simple news feed aggregation mechanism, as contributions from an increasing numbers of friends and connections flood the feed. The social network structure, which once delivered hand picked content, has become a victim of its own success, much to the frustration of users.

News feeds of many social networks do not support keyword based search, and users cannot easily find posts relating to topics of their interest. Users can of course unfriend, unfollow, hide, or mute undesired connections in their online social circles who flood their news feed, but the creation of rules and filters requires time and effort and only provides rigid options that turn off the posts from certain users. Add to this potential personal and social unease at unfriending or unfollowing online connections, as the connections are informed of this or this may even become public in some social networks, and we see barriers being raised that practically preclude users from actively curating their own friend list.

Automatic re-organisation of the news feed, aimed at filtering out irrelevant or uninteresting posts and, vice versa, highlighting posts of particular importance and relevance, offers a solid alternative to the manual rules and filters. Thus, more and more researchers<sup>1</sup> in the areas of data mining, machine learning, natural language processing, and social sciences have looked at the social network news feed filtering problem [3]. Several directions were studied under this broad umbrella: what factors make social network posts valuable [19,21], how can the feeds be ranked in a generic manner [12,18,27], and what semantic Web approaches can alleviate the ranking task [6]. However, many of these works stumbled upon a major obstacle – posts that are interesting for one user are not interesting for another user. That is, importance and relevance are user dependent, such that the feeds need to be filtered in a *personalised* manner.

Confirmation of the need for personalisation was found by Alonso et al. in a study that solicited opinions on how interesting or uninteresting a set of tweets from reputable organisations (BBC, New York Time, Reuters, and so forth) [1]. The authors asked five raters to indicate interestingness of more than 2,000 tweets and, unsurprisingly, the overall inter-rater agreement was close to 0. That is, the perceived interestingness of tweets was found to be subjective and user-dependent. To learn more, Deuker and Albers investigated user priorities for news feed content in a series of structured interviews aimed at uncovering the factors that determine content attractiveness in social networks [11]. They concluded that content attractiveness stems from two major factors: recipient’s *interest* in the topic of a content item and in the *author* (or poster) of the item. Clearly, both factors are user-dependent – a user may be interested in certain topics more than in others and appreciate content posted by some users more than by others. The design space of news feed personalisation was studied by Chen et al. in [8]. The authors focussed on Twitter and, in line with the findings of [11], outlined three considerations which were shown to correspond to user interest or satisfaction. The *source* of the tweets and the recipient’s *topics of interest* were highly relevant; in addition, the *popularity* of the tweets was also a driving factor.

<sup>1</sup> Notwithstanding, the under-water part of this iceberg includes applied research done by large-scale social networks, such as Facebook, LinkedIn, and Twitter. They put much effort into feed filtering and develop proprietary solutions, but these are most often not disclosed due to the commercial sensitivity and competitiveness.

So, what information can facilitate relevance judgements and scoring needed for personalisation? Let us name a few. A post or news feed item typically has an *author*, some *textual content*, and has often experienced *social approval* judgements, such as ‘likes’, ‘shares’, ‘retweets’, or ‘replies’. Posts also often contain some additional *content*, e.g., hashtags, URLs, pointers to other users, and so forth. Moving away from individual posts, social networks, by their very nature, have an underlying *network structure* that reflects explicit friendships, follower/followee relationships, or other articulated connections established on the network. Finally, there is a moderate amount of *interaction* data that, if captured and mined properly, can provide a rich source of information for user model creation. For example, content viewing can imply user interest, whereas viewing user profiles of others and direct communication with others can provide implicit indicators to supplement the explicit connection information. Finally, in social network user profiles often include unverifiable *background information* related to user demographics, location, preferences, skills, interests, and many other facets, which can be leveraged to inform the feed personalisation. Aggregating all these into a robust and accurate personalisation mechanism is not a straightforward task.

In this chapter we survey and bring together the state-of-the-art approaches for news feed personalisation. Our survey uncovers three dominant themes, under which the research fits. The first theme focuses primarily on those users who contribute posts and content, those who potentially see their posts, and the links between the two. User-to-user tie strength research examines how two users have interacted in the past to determine if one user’s posts should be given high priority in the news feed of the other. The second theme deals with the actual content of the posts and will sound familiar for readers knowledgeable about content based data mining and information retrieval, where the content of items (typically, text included in or linked to by a post) is examined to determine correlation to user interests or a query. The third theme details a set of works that look at the graph and structure of the underlying social network of both users and their posts, to determine similarity and relevance as criteria for inclusion of posts in the news feed. We also briefly touch upon other considerations, such as temporal information, use of latent factor models, and use of mobile apps and devices. Upon surveying the state-of-the-art works, we synthesise them, highlight popular motives coming through, raise emerging topics and research questions, and outline promising directions for future research.

## 2 Feed Personalisation – The Current State of Play

Let us start with some formalisation of the feed personalisation problem. Feed personalisation can be naturally considered as either a top- $K$  recommendation or a re-ranking problem. Let us denote by  $N$  the set of candidate items that can potentially be included in the feed, e.g., all the activities carried out on the social network by the user’s friends or all the tweets posted/retweeted by the user’s followees. With no personalisation applied to the feed, these are typically shown in

a reverse chronological order. Personalisation implies selecting a subset  $K \in N$ , such that  $|K| \ll |N|$ , which correspond to items of a higher importance for the recipient of the the feed. Given this formalisation of the personalisation problem, the recommendation task entails scoring the  $|N|$  candidate items and selecting  $|K|$  top-scoring items on behalf of the user. Alternatively, the re-ranking task entails re-ordering the  $|N|$  chronologically ordered candidates, i.e., scoring all the candidate items, keeping the  $|K|$  top-scoring items on top of the list, and removing the remaining  $N \setminus K$  items. The distinction between the recommendation and re-ranking solutions is quite blurry, such that in the rest of this chapter we will consider works that belongs to both without explicitly classifying them.

The central role of accurate news feed item scoring mechanism in the personalisation process is evident. In the following subsections, we survey several approaches for feed item scoring. We first elaborate on the user-to-user relationships; then, we examine works that incorporate text and content factors; then, social network and graph representation related considerations; and, finally, we introduce additional factors such as temporal information and constraints posed by the mobile device used by a feed recipient.

## 2.1 User-to-User Relationships

One of the pivotal considerations in identifying and scoring items of relevance look at the relationships between the user who performed the action or posted the content, and the recipient of the feed. Several works looked into the quantification of the online *tie-strength* between two social network users.

The trail-blazing work in this area was performed by Gilbert and Karaholios [15]. They proposed seven dimensions, that represented the strength of the relationship, or tie-strength, between pairs of Facebook users: *intensity* - amount of communication exchanged between the two; *intimacy* - use of intimacy and familiarity language in the communication; *duration* - period of time since the two established the online ties; *reciprocal* - resources, apps, and information shared between the two; *structural* - common groups and networks, or shared interests; *emotional* - gifts or congratulations exchanged between the two; and *distance* - similarity of religion, education, or political views. Using these dimensions, the authors derived 70 features and populated them using the observable online communication between the two users. The overall tie-strength score was computed as a linear combinations of these features. The tie-strength model was trained and the weights of individual features were determined using more than 2,000 explicit judgements provided by 35 participants (questions like “how strong is your relationship with X?”). It was found that the intimacy dimension accounted for more than 30% of the tie-strength score, whereas the most highly correlating individual features mirrored the duration of relationship: days since first and last communication. An offline study achieved predictive error smaller than 10%, showing the validity of the developed model.

A similar vein of research, aimed at predicting professional and personal closeness of an enterprise social network users, was done by Wu et al. [34]. They derived 60 features predicting user-to-user closeness and split these into five categories:

*subject user* - activity of the user who performed the action; *target user* - activity of the recipient of the feed; *direct interaction*: intensity of direct interaction between the subject and the target user; *indirect interaction* - intensity of indirect interaction between the two through common friends; and *corporate* - distance between the two in the organizational structure. Also here the overall closeness score was computed as a linear combination of the individual features. For the model training, the authors collected more than 4,000 explicit professional and personal closeness scores (“how closely are you currently working with X?” and “how likely are you to talk with X about your non-work life?”. Interestingly, the most important group of features was found to be the direct interaction between the two users: it accounted for close to 40% of weight for the professional and 48% of weight for the personal closeness. Predictive accuracy was evaluated, and the error was 18% for the professional and 22% for the personal closeness.

Another perspective on user-to-user relationship was taken by Jacovi et al. in [20]. The authors focussed on the interest in a user, i.e., curiosity in the activities done by that user. The interest reflects a directional asymmetric relationship, which may differ from closeness and tie-strength. They proposed four implicit indicators that may signal interest in a user: directly *following* the user, *tagging* the user in a people-tagging service, *viewing content* contributed by the user, and *commenting* on the user’s posts. Close to 120 participants were presented with lists of their online acquaintances and asked to select users of interest. Out of the above four indicators, tagging users was the most strongly correlated with interest, followed by direct following, and then by viewing accessing and commenting that were comparable. It should be noted that the observed correlation between the two explicit signals (tagging and following) and the interest level was almost double the correlation of the two implicit signals (viewing content and commenting).

A model that combines the features of [15] with the interaction-based weighting of [34] was proposed by Berkovsky et al. in [5]. In addition to the tie-strength score, the model also incorporated user preferences towards certain social network actions, such as posting/viewing content, commenting on posts of others, uploading images, and so forth. The underlying social network was an experimental portal of people engaged in a healthy living program. As the portal was fully controlled, the authors trained the model against the observed feed clicks as implicit interest indicators, and conducted an online inter-group study with a subset of the participants being exposed to personalised feeds. More than 500 feeds with clicks were reconstructed and analysed, and it was found that personalising the feeds increases user interactions, extends the duration of portal sessions, and boosts the contribution of user-generated content.

## 2.2 Text and Content Factors

Predicting the tie-strength of the user who posted a content item or performed a network activity is only one facet of the overall importance of the item in the news feed [11]. Other things that should be taken into consideration include the content of the posted item. The term ‘content’ embraces both the immediate

text included in the posts and other information, such as URLs, pointers to other users, tags, and more.

Paek et al. collected Facebook data pertaining to the perceived importance of individual news feed entries (“how important do you feel this feed item is for you?”) [24]. 24 users provided close to 5,000 explicit feed item importance judgements, which were discretised into binary important/unimportant labels. Using the observable Facebook logs, the authors mined and populated 50 predictive features across three groups: *social media* - metadata, number of comments, views, and likes, inclusion of URLs, and temporal information about posts and users; *text* - processed content of the post and previous communication between the two users, such as  $n$ -grams and  $tf \times idf$  vectors; and *background information* - static information about the users’ location, education, activities, interests, mined from their public Facebook profiles. An SVM classifier was trained and an ensemble model of all the features achieved close to 70% accuracy, which dropped to 63% when textual features were removed from the predictive model. This drop highlights the importance of the textual content of the posts/activities in the feed scoring model.

A similar model for ranking of tweets on Twitter was developed and evaluated by Uysal and Croft in [32]. They aimed specifically at the tweet ranking task and derived a suite of features that were split into four categories: *poster* - reputation, popularity, and activity of the person who posted the tweet; *content* - inclusion of hashtags, URLs, user mentions, and other emotional signals in the tweet; *text* - novelty and language model of the tweet content; and *recipient* - relation and past interactions between the recipient and the poster of the tweet. These categories of features were used individually as well as in combination, and evaluated offline using a corpus of more than 2,500 previously observed retweets. The best classification accuracy was achieved by the combined model (F-measure of 0.72). Content features were the top-performing category (F-measure of 0.5), while the performance of the pure textual features was surprisingly poor (F-measure of 0.04), perhaps due to the noisy nature of the text included in tweets. The dominance of the content features was re-affirmed in a tweet ranking accuracy evaluation.

Shen et al. proposed a method for a personalised interest-based reordering of tweets of a user’s followees [29]. User interests were determined by analysing the tweets published and consumed by the user, and modelling the topics of these tweets. The reordering incorporated five feature models: *temporal* - freshness of the tweet; *influence* - authority of the poster: number of poster’s followers and followees, number of lists on which the poster appears, and age and verification of the poster’s account; *quality* - length, URL, and hashtags of the tweet, as well as the number of retweets; *match* - match of the tweet to the interests of the recipient; and *social* - number of retweets and replies between the poster and the recipient. An ensemble model incorporating the above features was built and trained, considering the tweets that were retweeted or replied as interesting and aiming to prioritise and position these at the top of the tweet list. The reordering model was found to outperform the non-personalised and time-based

models with respect to several evaluation metrics. Interestingly, the most important features were the freshness of the tweet, the number of retweets, and the number of poster’s followers.

Rather than sorting and filtering tweets, they can be grouped into lists, each bearing a degree of relevance to various topics of interest of the target user. This method was studied by Burgess et al. in [7]. The set of topics and users posting tweets related to these topics was extracted by analysing the established follower-followee links, clustering users in dense sub-graphs, and mining the textual content of tweets posted by these users. The method was evaluated against manually created lists containing several hundreds of users. The automatically extracted lists of users and topics resembled the manual ones, with the observed F1 scores hovering between 0.7 and 0.8. The authors also developed several heuristics for assessing the relevance of individual tweets to the extracted lists. One heuristic was underpinned solely by the textual content of the tweets (unigrams and  $tf \times idf$  vectors), whereas the other considered also the included hashtags. The heuristics demonstrated a comparable degree of accuracy, with a slight preference toward those based on the textual content. The heuristics were also shown to be robust to noise, such that their accuracy only degraded slightly when the level of noise was as high as 50%.

### 2.3 Network and Graph Structure Factors

The value of textual features in microblogs is significantly lower than in typical social media, as the characteristics such as the length of posts, presence of acronyms, and high dynamicity of topics make text analytics difficult. Chen et al. in [9] proposed that the extracted features needed to be augmented with information mirroring the structure of the network. The authors devised a personalised tweet ranking model, based on the observable retweets as implicit interest indicators. The model encapsulated a suite of features categorised into four groups: *relation* - friendship between the poster and recipient, overlapping of their followees, and number of mentions in previous tweets; *content relevance* - relevance of the tweet content to the recipients’ status, retweet, or hashtag histories; *content* - length, hashtags, and URLs of the tweet; and *poster authority* - number of poster’s followers, followees, mentions, and status updates. These features were fed into a latent factor model that was evaluated using a corpus of more than 100,000 retweets. It was found that the model was consistently superior to several baselines and achieved average precision of 0.76. Also, the combined model substantially outperformed the individual models underpinned by single groups of features.

The work of Feng and Wang used the graph-based model of Twitter to rank tweets [13]. The nodes of the graph encapsulated the users (both the tweet poster and the recipient) and the tweets themselves, whereas the edges expressed the poster-recipient and recipient-tweet relationships. Additional features about the tweets (hashtags, URLs, age, popularity), users (similarity, mentions, reputation, probability to retweet and be retweeted), as well as user-tweet relationships (user profile vs tweet content similarity, mentions, hashtags) were mined. These features were used to train a factorisation model, aimed at predicting the retweet

probability for a given author, recipient, and tweet. The model was deployed to rank tweets according to their predicted retweeting probability and was evaluated against a corpus of more than 2.1 million retweets done by more than 28,000 users. The average precision of the combined model was around 42%, whereas a comparison of individual features showed that the author-recipient and recipient-tweet edges dominated the user and tweet nodes. This large-scale evaluation highlights the value encapsulated in the Twitter graph structure.

In [35], Yan et al. proposed a graph-theoretic model for personalised tweet recommendations. The recommender leverages a heterogeneous graph model consisting of a graph of users and a graph of tweets. In both sub-graphs, the nodes represent the users and the tweets, respectively, while the edges reflect the degree of their similarity. The user-to-user similarity is established based on the commonality of their followers, while the similarity of tweets is computed using their semantic content. Additionally, edges that connect the user and tweet sub-graphs indicate the original poster and the retweeters of a tweet. The nodes of the two sub-graphs are initially scored using the personalised PageRank algorithm, and then co-ranked, such that tweet score correspond to the scores of its poster and retweeters and, vice versa, user score correspond to the scores of the tweets they posted and retweeted. The model was applied for the tweet ranking task and evaluated using a corpus of more than 55 million retweets. The results demonstrated good ranking (nDCG greater than 0.5 for various sizes of the list) and classification (precision close to 62%) accuracy and outperformed several personalised ranking competitors.

## 2.4 Other Considerations

The challenge of ranking social updates on LinkedIn using click stream data was studied by Hong et al. in [17]. The authors evaluated three families of predictive models. The family of *linear* models included a feature-based model (features of the source user, recipient, and the update were used), a bias model (source user, recipient, and update category bias considered), and a temporal model. The family of *latent factor* models included a matrix and a tensor factorisation models, as well as their variants capitalising on a suite of manually crafted user features, such as seniority, connectedness, frequency and recency of visits, and so on. Lastly, since the above two families are optimised against different loss functions, they were combined using a *pairwise learning* model. The models were evaluated offline using LinkedIn’s interaction logs. Out of the linear models, the bias model achieved the highest precision, 0.53–0.60, for different training/testing splits. Tensor factorisation model with features was the top-performing latent model, with precision at the range of 0.59–0.65. Pairwise learning managed to combine the strengths of the two models and demonstrated precision scores hovering between 0.62 and 0.66.

Another clue to the importance of feed items lies in the temporal information, as user interests may drift over time. Two temporal dependencies – in performing social network activities and estimating user-to-user relevance – were studied by Freyne et al. in [14]. The authors exploited an offline dataset of user interactions



with the news feed of an enterprise social network and evaluated several short-term, long-term, and combined temporal models. Although the combined model demonstrated the best performance, it was found that short-term models were predictive of user-to-user relevance, while long-term models were found suitable for assessing the relevance of network actions for users. In other words, implicit user-to-user tie-strength score were found more volatile than the observed behaviour and interaction patterns of users.

The use cases of social networks are becoming increasingly mobile, such that the specific question of personalising the feed on a mobile device becomes relevant. This question was addressed by Cui and Honkala in [10]. They developed several content-based approaches (collaborative approaches cannot run on the client, as no access to information of others is available) that score feed entries and predict future clicking probabilities according to the items clicks observed in the past. A personalised PageRank predictor, a Bayesian predictor, and an ensemble model were evaluated in a 4-week live user study involving 40 participants. The Bayesian predictor was found to outperform the PageRank predictor individually, while the highest accuracy was achieved by the ensemble model. Furthermore, it was found that incorporating the time dimension in the model substantially improved the accuracy of the obtained results.

### 3 Discussion and Emerging Topics

Having surveyed a range of works on personalisation of social network activity feeds, we would like to summarise them in a concise manner. Table 1 presents the key contributions grouped by their underlying social environment, the predictive features that were used by the personalisation mechanism, the data against which the mechanism was trained, and the evaluation metrics used.

As can be seen from the table, most feed personalisation work published to date had conducted their evaluation on public social networks and only several used proprietary networks or ad-hoc communities established for the evaluation purposes. We would like to highlight the prevalence of Twitter as the chosen evaluation platform. We posit that this is attributed to two reasons. Firstly, the sheer volume of tweets faced by Twitter users means that Twitter is the “poster boy” for feed personalisation. Initial works in personalisation on Twitter were for the followee recommendation functionality and ways of expanding your network, but this has quickly been followed by works that support filtering through personalisation. Twitter is an attractive platform on which to carry out evaluations also due to the availability of data and API for easy crawling [22, 23]. We note also the strong dominance of implicitly provided training data, e.g., feed clicks, retweets, and replies, over explicitly labelled data. Indeed, it is unreasonable to expect users to explicitly annotate their network activity feed items, unless this is rewarded or directly related to their mainstream social network interactions.

Considering the groups of used features, we highlight the large number of works leveraging the network structure in the prediction process. This is not surprising, given that relationships and links established on a social network inherently reflect user interest in other users and/or in the content they contribute [16]. Of direct relevance to this is the reputation or authority of content

**Table 1.** Summary of feed personalisation works. Predictive features include user *activity* (actions of the poster and recipient), observed *interactions* (messages, tagging, replies), *network* structure (relationships between users, common friends), *graph* features (PageRank score, node connectedness), *textual* content of the posts, other *content* (hashtags, URLs, statuses), *reputation* (number of followees, followers, mentions), *temporal* information (posting frequency, account age), *similarity* (between the post and user interests, between the interests of users), and *static* information (location, gender, interests). Evaluation metrics include click-through rate (CTR), classification accuracy (CA), prediction accuracy (PA), and ranking accuracy (RA) metrics, usability questionnaire, computation time, and contributed content [28].

Work	Network	Features	Training	Metrics
Berkovsky et al. [5]	Health-related community	Action, activity, interaction	Feed clicks	CTR, CA, RA, contribution
Burgess et al. [7]	Twitter	Network, text	Explicit lists	CA, PA, novelty
Chen et al. [9]	Twitter	Text, network, content, similarity, reputation	Retweets	CA, CTR
Cui and Honkala [10]	Ad-hoc user community	Graph, interaction, temporal	Feed clicks	Usability, PA
Feng and Wang [13]	Twitter	Text, content, network, reputation, activity, similarity, temporal	Retweets	CA, time
Freyne et al. [14]	Enterprise social network	Activity, temporal, interaction	Feed clicks	CTR, RA
Hong et al. [17]	LinkedIn	Reputation, temporal, graph, network, static	Feed clicks	CA
Paek et al. [24]	Facebook	Text, network, static, content	Explicit judgments	CA, PA
Shen et al. [29]	Twitter	Temporal, reputation, content, similarity, activity, network	Retweets, replies	RA, CA
Uysal and Croft [32]	Twitter	Reputation, activity, network, interaction, content	Retweets	CA, RA
Yan et al. [35]	Twitter	Graph, network, text	Retweets	RA, CA

posters, which serves as an indicator of their salience on the network and is exploited in a number of works. It is important to note that the network structure can be pre-computed ahead of time, with relationship scores determined by previously observed interactions, and a simple lookup will determine an item’s expected relevance. Then, we note the use of content features – both the textual content and other content, such as hashtags and URLs. The content may serve as a direct predictor for whether the post or tweet will be relevant to the interests of the recipient. There is often a computational requirement involved in this process, where content, not included in a post needs to be fetched and examined

before a relevancy judgement can be made. In similar to various predictive tasks, temporal information is important and, as expected, these features are exploited in a number of approaches.

Two groups of features, increasingly leveraged in recent works, should be discussed. The first refers to features extracted from a graph-based representation of the social network in hand. These features supplement network features and encapsulate graph-based metrics derived from representing the data as a graph [31]. The potential of these features has been shown in other domains, such that their use for filtering and ranking of network feeds is natural and timely. Also in this case, the social network graph is often known and can be pre-computed and referenced to achieve personalisation in real time. In addition, the graph allows for content discovery, as relationships with users who are non-direct friends, but to whom the user is close in a network graph, can be seen to provide relevant feed items for consumption, facilitating serendipitous people and content discovery.

The second group of features deals with user-to-user interactions. Although some insight can be obtained from the established network links, a fine-grained quantification of user-to-user relationships should be derived from their mutual interactions, e.g., viewing the contributed content, mentioning each other, sending direct messages or retweeting, or even interacting with the same group of users [33]. We conjecture that features reflecting observable network interactions will gain an increasing popularity. Indeed the work of Berkovsky et al. [5] showed how relationships between different pairs of users can vary. Users may have friends whose photos they like to see, or whose blog articles they like to read. Being able to extrapolate strength and context of relationships between users is a valuable means for not only filtering out users but filtering out posts, in order to satisfy the differences identified in what users find interesting.

The vast majority of the work on feed personalisation use classification accuracy metrics. These highlight the requirement to simply predict whether a post will be of interest for the recipient or not, rather than determining the exact level of interest. Understanding the performance of an algorithm in a classification task is most suitable when the system aims to filter items from the news feed, rather than making explicit recommendations for items to consume. The second most-frequently used metric is ranking accuracy. This is natural, considering that the size of the feed shown to users is typically limited, and the importance of correctly ranking items in the feed is paramount [28]. If the exact level of interest in an item can be predicted, short lists of recommended items can further reduce the effort required to discover the most interesting items. Metrics incorporating click-through-rates or predicting the scores of feed items are also used, but less popular than classification and ranking metrics.

We are seeing increased diversity in the formats and media shared on social networks, and yet the research that we are seeing is primarily limited to a few commercial products. While news feeds are considered very natural and established features of widely popular social networks such as Facebook and Twitter, there is a gap in personalisation for social networks such as Pinterest, Instagram or Flickr, on which photo sharing is the aim but network, but hashtags and user tie-strengths could also be applied with ease. In addition, we need to prepare

for the arrival of new types of networks and media and more work is needed to uncover effective ways to merge the news feeds from multiple sources and networks to simplify things even further. Some progress into the aggregation of news feeds across multiple social network has been made by Summify<sup>2</sup> and several other commercial sites which link network accounts, monitor the target user's social graph, and email the digest of few most relevant stories per day.

Much of the work that was overviewed in our survey focuses on the mechanics and technicalities of using data mining and machine learning to identify interesting posts. Little work was uncovered that discussed the user interface needs associated with personalised news feeds. The metrics mentioned above lend themselves to different interface types, as discussed. Users have become increasingly familiar with the chronological lists that include all posts, interesting and boring together. Thus, there is a great need for work into novel interfaces, visualisations, and control mechanisms (consider tag clouds, multimedia plugins, and approaches recently proposed by Twingly<sup>3</sup> and Fidgt<sup>4</sup>) desired by users so that the feed personalisation is recognised as the time saving, productivity tool that it is originally aimed to be.

A number of years ago Facebook started to filter their news feeds, by removing content from users that it deemed the individual was less interested in. They faced an incredible push-back from the users, who were far from impressed by the lack of control and rigidity of the new filtering features. In essence, the users felt that this filtering contributed to the so-called *filter bubble* [25], a situation where the network decides on behalf of the users what feed items they are interested in, such that the users become isolated from cultural or ideological bubbles different to their opinions and viewpoints. This situation is unacceptable from ethical perspective and the social networks should find the gold spot between personalising the feed and limiting information exploration at the same time. Likewise, social feeds may pose a privacy threat, as they expose potentially sensitive information about activities in the user's close social circles that can accessed by untrusted parties or used inappropriately. Hence, privacy considerations should also be taken into account when filtering news feed items [4].

All in all, we feel that the research into social network activity feed personalisation is relatively in its infancy. Several solid algorithmic techniques were developed and thoroughly evaluated so far. Having said that, social network designers should keep in mind that their networks are user-facing systems. As such, much attention should be devoted to user aspects of personalisation: what do users find interesting and valuable [19]; how should the feeds be visualised and presented [26]; how do they prefer to interact with the feed [33]; does the feed answer their needs in the most encompassing and unobtrusive way [2]. We conjecture that these topics will have an increased exposure in the coming years and encourage researchers to consider these questions in their work.

<sup>2</sup> <http://www.summify.com>.

<sup>3</sup> <http://www.twingly.com/screensaver>.

<sup>4</sup> <http://sourceforge.net/projects/fidgtvisual/>.

## References

1. Alonso, O., Marshall, C.C., Najork, M.: Are some tweets more interesting than others? #hardquestion. In: HCIR (2013)
2. Bao, J., Mokbel, M.F., Chow, C.: Geofeed: a location aware news feed system. In: ICDE, pp. 54–65 (2012)
3. Berkovsky, S.: Network activity feed: finding needles in a haystack. In: MSM, p. 1 (2013)
4. Berkovsky, S., Borisov, N., Eytani, Y., Kuflik, T., Ricci, F.: Examining users' attitude towards privacy preserving collaborative filtering. In: DM.UM (2007)
5. Berkovsky, S., Freyne, J., Smith, G.: Personalized network updates: increasing social interactions and contributions in social networks. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 1–13. Springer, Heidelberg (2012)
6. Bontcheva, K., Rout, D.P.: Making sense of social media streams through semantics: a survey. *Semant. Web* **5**(5), 373–403 (2014)
7. Burgess, M., Mazzia, A., Adar, E., Cafarella, M.J.: Leveraging noisy lists for social feed ranking. In: ICWSM (2013)
8. Chen, J., Nairn, R., Nelson, L., Bernstein, M.S., Chi, E.H.: Short and tweet: experiments on recommending content from information streams. In: CHI, pp. 1185–1194 (2010)
9. Chen, K., Chen, T., Zheng, G., Jin, O., Yao, E., Yu, Y.: Collaborative personalized tweet recommendation. In: SIGIR, pp. 661–670 (2012)
10. Cui, Y., Honkala, M.: A novel mobile device user interface with integrated social networking services. *Int. J. Hum.-Comput. Stud.* **71**(9), 919–932 (2013)
11. Deuker, A., Albers, A.: Who cares? content sharing on social networking sites: a grounded theory study. In: PACIS, p. 156 (2012)
12. Duan, Y., Jiang, L., Qin, T., Zhou, M., Shum, H.: An empirical study on learning to rank of tweets. In: COLING (2010)
13. Feng, W., Wang, J.: Retweet or not?: personalized tweet re-ranking. In: WSDM, pp. 577–586 (2013)
14. Freyne, J., Berkovsky, S., Daly, E.M., Geyer, W.: Social networking feeds: recommending items of interest. In: RecSys, pp. 277–280 (2010)
15. Gilbert, E., Karahalios, K.: Predicting tie strength with social media. In: CHI, pp. 211–220 (2009)
16. Guy, I., Zwerdling, N., Ronen, I., Carmel, D., Uziel, E.: Social media recommendation based on people and tags. In: SIGIR, pp. 194–201 (2010)
17. Hong, L., Bekkerman, R., Adler, J., Davison, B.D.: Learning to rank social update streams. In: SIGIR, pp. 651–660 (2012)
18. Huang, H., Zubiaga, A., Ji, H., Deng, H., Wang, D., Le, H.K., Abdelzaher, T.F., Han, J., Leung, A., Hancock, J.P., Voss, C.R.: Tweet ranking based on heterogeneous networks. In: COLING, pp. 1239–1256 (2012)
19. Hurlock, J., Wilson, M.L.: Searching twitter: separating the tweet from the chaff. In: ICWSM (2011)
20. Jacovi, M., Guy, I., Ronen, I., Perer, A., Uziel, E., Maslenko, M.: Digital traces of interest: deriving interest relationships from social media interactions. In: ECSCW, pp. 21–40 (2011)
21. Lage, R., Denoyer, L., Gallinari, P., Dolog, P.: Choosing which message to publish on social networks: a contextual bandit approach. In: ASONAM, pp. 620–627 (2013)

22. McCormick, T.H., Lee, H., Cesare, N., Shojaie, A.: Using twitter for demographic and social science research: tools for data collection. *Sociological Methods and Research* (2013)
23. Morstatter, F., Pfeffer, J., Liu, H., Carley, K.M.: Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In: *ICWSM* (2013)
24. Paek, T., Gamon, M., Counts, S., Chickering, D.M., Dhesi, A.: Predicting the importance of newsfeed posts and social network friends. In: *AAAI* (2010)
25. Pariser, E.: *The Filter Bubble: What the Internet is Hiding from You*. Penguin UK, London (2011)
26. Parra, D., Brusilovsky, P., Trattner, C.: See what you want to see: visual user-driven approach for hybrid recommendation. In: *IUI*, pp. 235–240 (2014)
27. Sarma, A.D., Sarma, A.D., Gollapudi, S., Panigrahy, R.: Ranking mechanisms in twitter-like forums. In: *WSDM*, pp. 21–30 (2010)
28. Shani, G., Gunawardana, A.: Evaluating recommendation systems. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) *Recommender Systems Handbook*, pp. 257–297. Springer, New York (2011)
29. Shen, K., Wu, J., Zhang, Y., Han, Y., Yang, X., Song, L., Gu, X.: Reorder user's tweets. *ACM TIST* **4**(1), 6 (2013)
30. Susarla, A., Oh, J., Tan, Y.: Social networks and the diffusion of user-generated content: evidence from youtube. *Inf. Syst. Res.* **23**(1), 23–41 (2012)
31. Tiroshi, A., Berkovsky, S., Kâafar, M.A., Vallet, D., Chen, T., Kuflik, T.: Improving business rating predictions using graph based features. In: *IUI*, pp. 17–26 (2014)
32. Uysal, I., Croft, W.B.: User oriented tweet ranking: a filtering approach to microblogs. In: *CIKM*, pp. 2261–2264 (2011)
33. Wilson, C., Boe, B., Sala, A., Puttaswamy, K.P.N., Zhao, B.Y.: User interactions in social networks and their implications. In: *EuroSys*, pp. 205–218 (2009)
34. Wu, A., DiMicco, J.M., Millen, D.R.: Detecting professional versus personal closeness using an enterprise social network site. In: *CHI*, pp. 1955–1964 (2010)
35. Yan, R., Lapata, M., Li, X.: Tweet recommendation with graph co-ranking. In: *ACL*, vol. 1, pp. 516–525 (2012)