Proposed Title

Using Twitter as a Curation Filter to Overcome Information Overload

Grant Summary

For our funding, we will to apply for the Social-Computational Systems program grant (10-600) (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503406). In applying for the grant, we hope to realize the National Science Foundation's goal of creating "socially intelligent computing" by designing a system that better controls the overwhelming number of tweets one sees in his/her Twitter stream. Our proposal will try to solves the National Science Foundation's appeal for a system that allows people to more easily comprehend the noise of today's information environment and more specifically the Twitter's environment. In performing such research, we hope to better understand the impact that filtering will have on a person's interaction with Twitter and its news feed.

Problem

As readily available information has increased exponentially in each of the last five decades, the problem of information overload has increasingly worsened (Carlson, 2003). Information overload is described as "a subjective experience of the insufficiency of time needed to make effective use of information resources available in specific situations" (Savolainen, 2007). Ho and Tang (2001) argue that there are serious implications of information overload since it can affect an employee's performance and therefore stunt a business's productivity. In extreme cases, information overload can even cause health damage and "technostress" (Bawden & Robinson, 2009; Carlson, 2003).

Although the problem of information overload has existed for some time now, as modern technology continues to increasingly influence our lives, such information "has aggravated the problem to unprecedented levels" (Francisci et al., 2012). Too much information can cause a person to feel overwhelmed, thus such person might avoid certain services altogether or greatly limit his or her service usage. Himma argues, "(t)he amount of online content to which people in reasonably affluent nations have access is increasing at an astonishing – even alarming – rate" (Himma, 2007). Such implications can be serious because Internet users in particular experience information overload every day from email, to spam, to social networks. It seems as though the problem is only getting worse.

Neylon (2011) suggests that perhaps the problem is not one of having too much information, but rather not being able to filter it effectively. This is interesting because it points to the core of

what information overload is. Often times what we refer to as "too much information" is in reality a failure of the medium or information source to prioritize or highlight the most relevant information.

The problem of information overload exists in abundance on social networks and is well recognized within the research community. For instance, Cherubini et al. (2010) argue that "As a result of the continuous growth of shared online content, most of the users of social networks are overloaded by constantly changing feeds and they struggle to keep up with the content." The introduction of the newsfeed, or rather real-time data streams filled with user-generated content, have introduced an entirely new level of information overload.

Facebook realized that a service's popularity could cause a strain on its potential success if it does not find a means to curtail information noise because users might have difficulty finding relevant information. Consequently, it implemented various algorithms that filter a user's newsfeed for the most important news items based on factors such as the user's interaction with friends. In a similar fashion, Erdos et al. strive to coordinate a solution that utilizes "nodes [that] are responsible for the removal (or significant reduction) of the redundant data items relayed through them" (Erdös, Ishakian, Lapets, Terzi, & Bestavros, 2012, p. 1). As a particular tweet becomes more prominent in a person's nodes, it creates a greater chance of such a message being filtered for a user to see.

Liu (2010) describes the problem as one of "curational overload":

> Ubiquitous online social technologies generate a plethora of user-generated content publicly accessible via the web. As we increasingly create personal collections online, we each face the problem of ―curatorial overload: too much information, too difficult to organize and retrieval... Designing curatorial tools can help people sift through and make sense of these massive content streams in real-time (Liu, 2010)

Kobsa & Stephanidis (1999) describe how tools can enhance a user's ability to find important information in a limited amount of time. There are currently a number of such tools available, Phelan et al. (2009) argue that "Content-based and collaborative filtering techniques have been used to good effect and the recent growth of services such as Digg (http://www.digg.com) demonstrate the value of recommendation techniques when it comes to deliver a more relevant and compelling news service" (Phelan et al., 2009).

Our inability to filter information effectively is most pronounced when looking at the realm of social media. Tsagkias et al. (2011) describe the relationship between online news and social media such as micro-blogging as a "symbiotic" one, where they both now rely on each other to

be effective. This has certainly become the case with Twitter, the leading micro-blogging service. Golbeck describes that currently, as of May 2012, Twitter does not afford its users the ability to avoid a particular topic or event when searching for Tweets (Golbeck, 2012). Such is an example of the inability of its user to filter out the noise of events of non-interest.

Chen et al. (2010) contend that along with the abundance of information comes the scarcity of attention, which introduces two core user needs: filtering and discovery. Filtering helps users sift through all the information they are already receiving, while discovery present them with new information from outside of their own information streams that they may find useful. Wu et al. (2010) describe a solution that automatically creates unique tags from the words of a user's tweets to better identity the user's interest. This solution eases to create a more effective filtering mechanism by successfully labeling the content of each message.

Our proposed solution focuses on the need to filter the abundance of information on Twitter by employing collaborative filtering, using social chatter and the implicit indicators of interest already present within Twitter, to suggest and recommend relevant and popular news content to users.

Background

The microblogging service Twitter has become the de facto news source for many users, though it originally was not designed for the purpose of discovering and consuming online content. Over half (56%) of Twitter users utilize the service to share information with 22% of that user population sharing news and business content (Westman & Freund, 2010). With approximately 200 million users, that amounts to roughly 24.6 million users actively using Twitter to share such content. According to Nagpal et al. (2012), as of 2010, 25% of the approximately 90 million tweets per day contained links. Some studies have found even more significant evidence of the emergence of Twitter as a news source, Tsagkias et al. (2011) found 85% of Twitter statuses to be news-related.

Twitter is one of many web services that support a wide-range of communities with different needs to collaboratively discover, share, and discuss web-based content. Twitter does afford hashtags (#) and @ symbols to allow for easier tagging and searching, but such metadata is still not reliable or fast enough to control information overload; for instance people may use the same hashtags for entirely different topics. Twitter also has a tool called "Twitter Lists" which is essentially a way to view self-curated Twitter timelines based on a specific set of Twitter users, without the need to individually follow each Twitter user. While Twitter lists is useful in certain use-cases, it is not useful as a means to help consume the plethora of news articles users see in their timelines, and in fact could exacerbate the problem by providing even more sources for news.

There are serious challenges with filtering all this content that users see on social networks such as Twitter. Some argue that the lack of an effective filtering mechanism can actually cause people to share less:

> Content shared through social networks may give rise to adverse selection. In other words, without a proper mechanism to distinguish good quality from low quality content, owners of good content may be reluctant to publish their content thus lowering the average quality of all shared content. (Cherubini et al., 2010)

The two main techniques that have emerged for management of social media streams are collaborative filtering and content-based techniques. De Francisci et al. (2012), argue that collaborative filtering is the most successful recommendation paradigm, which seeks to suggest content based on user ratings and their interaction with the system. Based on these findings and past research, we believe a tool that quickly lets users read the most relevant and popular news articles from their Twitter timeline can help solve the problem of information overload for users looking for news content on Twitter. If we provide users who consume a significant amount of content on Twitter with a collaborative filtering tool that enables them to quickly and easily find the most relevant and popular articles from Twitter, users will be able to effectively discover and consume content.

Similar tools do exist, Phelan et al. (2011) created a tool called Buzzer that served as a sort of new RSS reader that used Twitter data as well as RSS to recommend articles to users. Buzzer adopted a content-based recommendation technique, by mining content terms from RSS and Twitter feeds as the basis for article ranking. In essence, if an article appeared in the RSS feed as well as their Twitter feed, it would placed into a recommended queue and given a specific score and presented to the user. There are a number of other web services that attempt to address information overload for news by curating the content for you and telling you what you should be reading, based on a set of algorithms. Services like Zite, Flipboard, News.me, Float, and Smartr all give you content based on your social stream. Although none of the services we have come across focus solely on Twitter.

Our tool will be in the form of a web application that will use the Twitter REST API to communicate with Twitter. Our solution will allow Twitter users to log in with his or her Twitter credentials and immediately be presented with an interface that filters his or her Twitter timeline to display only the articles that have been tweeted by users he or she follows. This would drastically reduce the amount of information users have to process, by presenting them with only the actual content from their timeline rather than every tweet and update from the users they are following. So rather than seeiing an update from their bestfriend about the Chpitole meal they just had, users will only see articles. We will do this by filtering tweets according to their tweet entity “URL,” which is provided through the Twitter REST API.

Our web application will then rank the articles found in users' timelines by their popularity across Twitter; calculated by summing the total number of tweets across Twitter that contain the URL for that specific article. This tweet count will not only include retweets, but also the raw number of tweets that contain the article's URL. This is important because the official retweet function that Twitter often tracks and pushes users to use, is not often an accurate metric to use for the popularity of a given tweet. Many times, users will tweet an article from the publisher's page or from another application that auto-generates a tweet. Our tweet count will include retweets and any tweets that contain the URL, which we believe will provide a much more accurate measure of how popular an article truly is on Twitter. We will obtain this tweet count by using the Twitter Search API, which allows you to search the entire Twitter firehouse for any query including URLs. The result of the above will be a ranked article view of the user's timeline that allows the user to quickly see what the most popular articles are based on the Twitter users they follow.

One critical aspect of our solution is the tweet count upon which we base the rankings of articles, so it is worthwhile to explore our design decision in detail. Nagpal et al. (2012) state, "People often find useful content on the web via social media. However, it is difficult for users to aggregate the information and recommendations embedded in a torrent of social feeds like email and Twitter." The authors suggest using social chatter as a tool to curate content, though they state, "personalized and implicitly curated social recommendations of web content already exist in the user's personal communications (such as email) and social networks (such as Twitter)." We are applying this theory by using the number of tweets across Twitter as an indicator of popularity for the content. If a tweet contains an article, the Twitter user who tweeted the article is implicitly recommending it. Furthermore, we consider retweeting an article to be a strong indicator of interest by a user. Uysal and Croft (2011) explain that "Since the practice of retweeting includes reading the tweet, deciding that it is worthwhile to share, and then acting on it; we can use retweet (RT) behavior as an explicit signal that the user considers the tweet to contain useful or interesting information." For this reason, we believe a retweet or tweet is an excellent indicator of the usefulness of an article, essentially acting as an implicit curation filter. Since we are aggregating this implicit indicator from many different users, the result is a crowdsourced filter which reflects how popular articles are amongst the crowd of Twitter users.

Another aspect that is important to highlight is the concept of time and how that plays a role in our solution. In general, the longer the duration of time an article is present on Twitter, the more it will be retweeted and tweeted. This can be seen by simply viewing the details of a tweet from Washington Post, which displays the number of retweets, waiting two hours and viewing the count again. As more users discover the tweet and the article, the more that tend to share it. This could prove to be an issue if we are presenting users with a ranked view of articles based on the number of tweets, it may result in an article staying ranked at the top for days. To address this, we will implement exponential decay into our ranking algorithm. This means although an article

may continue to be tweeted more and more, as time goes on we will decrease the article's ranking, so as to always show the user the "freshest" news. This way no one article will remain at the top of the ranked view for an extended period of time. We will need to do extensive testing with users to see what the optimal decay value is, but we are confident incorporating this will improve the user experience.

Research question(s)

RQ 1. How pervasive and impactful is the perception of information overload amongst the population of Twitter users?

- RQ1A. What is the threshold at which Twitter users find themselves feeling overwhelmed by the amount of content they are exposed to?
- RQ1B. Do feelings of information overload change in the various contexts and times in which Twitter is used?

RQ 2. Are any of the currently available Twitter clients effective in reducing feelings of information overload compared to the standard Twitter interface?

RQ 3. Is the prototype developed by our team effective in reducing feelings of information overload?

Research Design & Methods

An evaluative approach will be utilized for our research design and methods where there are three research study sets, each designed to provide a holistic insight to the problem being identified, potentially involving complex social issues. This research design examines the possible complex interplay of the many variables involved, as our research problem involves a broad scope of users and many possible use cases. In our evaluative approach, each research study seeks to offer unique answers that will provide necessary insight and validate our proposition.

Research Study 1: Interviews and Surveys (Qualitative & Quantitative)

An interview and survey will be conducted to produce qualitative and quantitative data that will be synthesized into findings to determine the validity of the proposed solution.

Participants will be recruited from a call for research using Twitter. The participants will be recruited by selecting a group of 10 representative users where representative users are defined as Twitter users who utilize Twitter for the purpose of aggregating information from any source. Interviews will be conducted in person and recorded by a note-taker.

The set of questions designed will provide insight of how subjects interpret their own behavior to qualitatively answer the question of Research Question 1, “How pervasive and impactful is the perception of information overload amongst the population of Twitter users?” Research Question 1A “What is the threshold at which Twitter users find themselves feeling overwhelmed by the amount of content they are exposed to?”, and Research Question 1B, “Do feelings of information overload change in the various contexts and times in which Twitter is used?”

- Questions will include:
    - Name, age, gender, occupation, education, location, devices used?
    - What devices to you use Twitter on?
    - How often do you use Twitter?
    - What information do you seek when using Twitter?
    - Why do you like Twitter?
    - What do you not like about Twitter?
    - Do you ever feel overwhelmed when using Twitter? If so, explain.
    - Do you use any tools with Twitter, such as TweetDeck or applications to streamline your news source?
    - Amongst the currently available Twitter clients, are any of these applications effective in reducing feelings of information overload compared to the standard Twitter interface?
    - How do you avoid information overload on social networks such as Twitter and find the most relevant news and articles
    - How often do you Tweet?
    - What do you Tweet about
    - Do you use Twitter for personal or business related purposes?
    - How many followers and people do you follow?
    - Who do these followers include
    - What is your handle?
    - Rate your satisfaction of using Twitter on a scale of 1-10, 10 being very satisfied
    - Rate your level of stress about how cluttered your Twitter feed is on a scale of 1-10, with 10 being not cluttered at all.

Data will be gathered as quantitative data from the Twitter handle of each subject that will include the amount of followers, amount of people following, a rating of satisfaction of using Twitter, and the level of stress of Twitter being cluttered.

The synthesis and analysis of the qualitative data will be conducted by forming groups of personas and identifying pain-points. This data will support the quantitative data from this research study as well as the following research studies, two and three.

The synthesis and analysis of the quantitative data will be conducted by carrying out a statistical analysis. We will conduct a multivariate regression analysis to describe how the variables, thus being amount of followers, amount of people following, a rating of satisfaction of using Twitter, and the level of stress of Twitter being cluttered for each subject respond simultaneously to changes in others. This analysis seeks to provide insight to each variable and how they relate to each other. The correlation of the two data sets will examine the relationship between the two concepts, being how much information is involved (followers and people being followed) with the user rated level of effectiveness (satisfaction to being overwhelmed). This will identify a threshold of how the users are effectively using Twitter and should be used in conjunction with the qualitative data gathered in this study. It will also inform the design of the questions for Research Study 2, "User Testing of Prototype (Qualitative)."

Research Study 2: User Testing of Prototype (Qualitative)

A series of prototypes will be built and tested. The prototypes designed and tested will provide insight of how subjects feel about their ability to effectively engage with and utilize the tools we present them with and the findings will enable us to gauge effectiveness and answer the question of Research Question 3, "Is the prototype developed by our team effective in reducing feelings of information overload?"

The series of prototypes will test various interfaces for consuming content from Twitter that are designed specifically to account for information overload. The data gathered in testing will determine if an interface for consuming content from Twitter that is designed specifically to account for information overload can effectively reduce this problem. The prototypes will be of a tool that will ideally enable users to identify and consume the most relevant and popular news articles from their Twitter timeline in a manner that will help solve the problem of information overload for users looking for news content.

For the first round of testing, the team will develop a minimum of three prototypes, each exploring a different interface design approach.

A minimum of thirty participants will be recruited via Twitter through the extended networks of the team. Each of the three prototypes will be tested on ten of the research participants. Their activity during the tests will be collected via screen capture software, as well as through video and audio recording tools.

The set of tasks the user will carry out will be designed to provide insight on the level of effectiveness each of the prototypes demonstrates in reducing information overload in simulated

situations. Users will be instructed to think aloud by talking through their thoughts at each decision. Tasks will be defined based on the user, which will be manipulated to gauge effects. Questions will precede tasks that are informed by Research Study 1.

Tasks may include:

- Gathering information in a way that is typical to the user's current Twitter usage
    - Some examples may include: identifying and consuming articles relevant to the users interest or highlighting relevant Twitter posts from other users that the user will seek to take action on, such as retweet, take action on a third-party article, or to network, interact with, or follow a new user.
- Finding the most interesting and relevant items from a large set of a available content
    - Some examples may include: finding job postings through a specific network of recruiters active on Twitter or finding relevant news articles based on a particular set of interests of the user such as technology, politics, culture, etc.

At the end of each testing session, a survey will be administered to the participants. Survey questions may include:

- Do you feel this prototype effectively solved the problem of information overload?
- Do you feel tasks were easier to perform?
- Is there anything you particularly liked about this solution?
- Is there anything you disliked about this solution?

Data will be synthesized from both the prototype testing and the subsequent survey to inform a second round of prototype development. The second iteration of the prototype will then be tested on an additional thirty research participants.

Participants for the second round of testing will again be recruited via Twitter through the extended networks of the team. The nature of the testing is anticipated to be similar to the methods used in the first round of prototype testing; though the team anticipates that refinements to the questions and methodologies utilized will naturally occur as insights into the product are gained throughout the project.

Research Study 3: Contextual Inquiries (Qualitative)

Contextual inquiry will be used to gather descriptive research data to inform our knowledge of existing solutions and to qualitatively answer the question of Research Question 2, "Are any of the currently available Twitter clients effective in reducing feelings of information overload compared to the standard Twitter interface?" The findings of the research study will provide insight on existing relevant forms of interfaces, technology, and solutions that will be synthesized into descriptive findings informing our approach to the proposed solution.

The contextual inquiry will consist of an interview, an observation period, a summarization, and an analysis. The interviews, observations, and summarizations will be conducted with the subject to gather raw data. The data gathered will be later synthesized.

Participants will be recruited from a call for research using Twitter. The participants will be recruited by selecting 10 representative users where representative users are defined as Twitter users who utilize Twitter for the purpose of aggregating information from any source.

Interviews will be conducted in person and recorded by a note-taker. Each subject will be interviewed individually in their own natural environment, when carrying out their tasks, with the least amount of interference from the facilitator as possible. The interview aims to gather an understanding of these said applications in terms of their effectiveness in reducing feelings of information overload compared to the standard Twitter interface. This will be conducted by analyzing existing social media feedback content that exists in the form of ratings, articles, user generated feedback, forums, etc

Potential interview questions will include:

- What are some currently available Twitter clients and news aggregation services that use Twitter as an input?
- What design patterns, key findings, and heuristic best practices are you aware of?
- Of these, what problems and pain points have you recognized?
- How do you feel about [said applications] in terms of their effectiveness in reducing feelings of information overload compared to the standard Twitter interface?

The observation period will consist of the facilitator and note taker observing the subject interactation with the systems identified in the interview.

The data gathered will then be synthesized by constructing a contextual inquiry model of using affinity diagrams to call out relevant information and key findings. This will aim to reveal associations from data to validate our hypothesis informing our solution.

Timeline

The goal of the timeline is to identify the tasks and length of tasks necessary to conduct the research in this proposal.

These are the essential tasks and a time estimate it will take in order to complete all tasks necessary to carry out the research in this proposal:

Develop Prototypes ● Initial Prototype ● Iterative Prototype One ● Iterative Prototype Two	12 months ● 4 months ● 4 months ● 4 months
Recruit and schedule participants	1 month
Set up research study environments	1 month
Carry out interviews and surveys	1 month
Synthesize data	2 months
Make any changes to prototype	1 month
Formalize research findings	2 months
Total Time	**20 months***

*Time estimate includes total length of separate tracks of including development of prototype, recruiting participants, setting up environment, and all other necessary steps with the assumption the four members of our group will be carrying out the tasks.

Budget

The total budget for this research project is estimated to be $128,836 based on the following breakdown of activities and expenses.

Personnel	Role	% Effort	Months	Base Salary	Fringe 24%	Salary Dollars
Scott Abromowitz	Principal Investigator	50	20	$65,000	$13,000	$67,167
Yonas Beshawred	Coinvestigator	25	12	$65,000	$3,900	$20,150
Laura Rogers	Coinvestigator	25	9	$65,000	$2,950	$15,113
Kevin Vigneault	Coinvestigator	25	9	$65,000	$2,950	$15,113
					Subtotal	**$117,542**

Expense	Rate	Quantity	Total Cost
Server Hosting	$50 / month	20 months	$1,000
Recruiting Tools	$49 / month	6 months	$294
Statistician	$75 / hour	100 hours	$7,500
Participant Incentives	$50 / participant	50 participants	$2,500
		Subtotal	**$11,294**

Feasibility

To successfully conduct the proposed research, the research team must have adequate skills to both conduct the initial research and build the proposed solution. Beyond these skills, a number of additional factors must be accounted for.

Twitter, the product our research focuses on, has been in public use for over six years and it is highly likely that it will continue to be operational throughout the duration of the study. The product has amassed an extremely large user base; currently known to be over 140 million people. Additionally, our research does not focus on a particular audience or demographic. Therefore, finding viable research candidates is very achievable.

The research methods we propose do not require specialized equipment or space. The team will employ free and low-cost methods for conducting and recording the sessions. For the focus groups and prototype testing, we intend to use space available to team at the University of

Maryland's Human-Computer Interaction Lab. The contextual inquiries can be conducted in a variety of public spaces at no cost.

The prototype will be developed using freely available, open-source programming languages and tools. Twitter's free and publicly available API, which the prototype will be built upon, has a long history of third-party use and it is highly likely that it will continue to be available throughout the duration of the study. The prototype will be hosted using standard web servers that can be rented on a low-cost, monthly basis with no long-term commitment. Additionally, the team has an ample amount of collective experience designing and building software prototypes. These combined factors lead us to believe to that building and maintaining the prototype is very achievable.

Scott Abromowitz, Yonas Beshawred, Laura Rogers, and Kevin Vigneault all have experience using Twitter for content consumption. Every member of the team has experienced “information overload” particularly while using Twitter, so the research team is intimately familiar with and thoroughly researched the problem this research project will attempt to address. All four researchers are currently pursuing a Masters of Science in Human-Computer Interaction at the University of Maryland, College Park; therefore, the team is familiar with best practices for usability and interface design. The team also analyzed dozens of peer-reviewed journal articles to better understand current research pertinent to solving the problem.

Dissemination plan

The primary audiences for the dissemination of the project’s results are researchers within the following fields: human computer interaction, information studies, and computer science. We will reach this audience by presenting our results to faculty and students at the University of Maryland, College Park; posting the report to the Human-Computer Interaction Lab’s website; and submitting a paper to the ACM SIGCHI Conference on Human Factors in Computing Systems. The secondary audience is application developers who build similar products aimed at automatically curating content via social networks similar to Twitter. To reach this audience, we intend to submit an abstract of our findings to popular blogs within the application development community. Additionally, dissemination may occur through presentations at conferences beyond the ACM SIGCHI Conference and through articles published in peer-reviewed journals.

Key Person Profiles

Scott Abromowitz: Scott is a full-time student in the Human-Computer Interaction Master's Program at the University of Maryland, College Park. He has experience conducting usability testing on consumer products. He earned a Bachelor of Arts in Political Science from the University of Wisconsin, Madison.

Yonas Beshawred: Yonas is a full-time student in the Human-Computer Interaction Master's Program at the University of Maryland, College Park. He has experience developing web applications using Twitter's API. Prior to entering the program, Yonas worked as a Business Analyst & Strategic Consultant at Accenture. He earned a Bachelor of Science in Information Systems from the University of Maryland, College Park.

Laura Rogers: Laura is currently employed full-time as an Information Architect at Sapient Government Services. She is a part-time student in the Human-Computer Interaction Master's Program at the University of Maryland, College Park. Laura has experience conducting user interviews and developing surveys. She earned a Bachelor of Fine Arts in Digital Arts from George Mason University.

Kevin Vigneault: Kevin is currently employed full-time as User Experience Designer at Viget Labs. He is a part-time student in the Human-Computer Interaction Master's Program at the University of Maryland, College Park. Kevin has experience conducting contextual inquiries, as well as designing and developing software prototypes. He earned a Bachelor of Science in Finance and Economics from the University of Maryland, College Park.

References

Bawden, D., & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. *J. Inf. Sci.*, *35*(2), 180–191. doi:10.1177/0165551508095781

Carlson, C. N. (2003). Information overload, retrieval strategies and Internet user empowerment. Conference Paper. Retrieved February 18, 2012, from http://eprints.rclis.org/handle/10760/5432#.T0AX0nJSRG4

Cataldi, M., Di Caro, L., & Schifanella, C. (2010). Emerging topic detection on Twitter based on temporal and social terms evaluation. *Proceedings of the Tenth International Workshop on Multimedia Data Mining*, MDMKDD '10 (pp. 4:1–4:10). New York, NY, USA: ACM. doi:10.1145/1814245.1814249

Chen, J., Nairn, R., Nelson, L., Bernstein, M., & Chi, E. (2010). Short and tweet: experiments on recommending content from information streams. *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 1185–1194).

Cherubini, M., Gutierrez, A., de Oliveira, R., & Oliver, N. (2010). Social tagging revamped: supporting the users' need of self-promotion through persuasive techniques. Proceedings of the 28th international conference on Human factors in computing systems, CHI '10 (pp. 985–994). New York, NY, USA: ACM. doi:10.1145/1753326.1753473

De Francisci Morales, G., Gionis, A., & Lucchese, C. (2012). From chatter to headlines: harnessing the real-time web for personalized news recommendation. *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12 (pp. 153–162). New York, NY, USA: ACM. doi:10.1145/2124295.2124315

Erdös, D., Ishakian, V., Lapets, A., Terzi, E., & Bestavros, A. (2012). The filter-placement problem and its application to minimizing information multiplicity. *Proc. VLDB Endow.*, *5*(5), 418–429.

Golbeck, J. (2012). The twitter mute button: a web filtering challenge. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12 (pp. 2755–2758). New York, NY, USA: ACM. doi:10.1145/2208636.2208673

Himma, K. E. (2007). The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition. *Ethics and Inf. Technol.*, *9*(4), 259–272. doi:10.1007/s10676-007-9140-8

Ho, J., & Tang, R. (2001). Towards an optimal resolution` to information overload: an infomediary approach. *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '01 (pp. 91–96). New York, NY, USA: ACM. doi:10.1145/500286.500302

Kobsa, A., & Stephanidis, C. (1999). SELECT: Social and Collaborative Filtering of Web Documents and News. *Proceedings of the 5th ERCIM Workshop on User Interfaces for All: User-Tailored Information Environments* (pp. 23–37).

Liu, S. B. (2010). The Rise of Curated Crisis Content. *Proceedings of the 7th International ISCRAM Conference–Seattle* (Vol. 1).

Nagpal, A., Hangal, S., Joyee, R. R., & Lam, M. S. (2012). Friends, romans, countrymen: lend me your URLs. using social chatter to personalize web search. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 461–470).

Neylon, C. (2011). It's not filter failure, it's a discovery deficit. *Serials: The Journal for the Serials Community*, *24*(1), 21–25.

Phelan, O., McCarthy, K., & Smyth, B. (2009). Using twitter to recommend real-time topical news. *Proceedings of the third ACM conference on Recommender systems*, RecSys '09 (pp. 385–388). New York, NY, USA: ACM. doi:10.1145/1639714.1639794

Savolainen, R. (2007). Filtering and Withdrawing: Strategies for Coping with Information Overload in Everyday Contexts. *Journal of Information Science*, *33*(5), 611–621. doi:10.1177/0165551506077418

Tsagkias, M., de Rijke, M., & Weerkamp, W. (2011). Linking online news and social media. *Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 565–574).

Uysal, I., & Croft, W. B. (2011). User oriented tweet ranking: a filtering approach to microblogs. *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11 (pp. 2261–2264). New York, NY, USA: ACM. doi:10.1145/2063576.2063941

Westman, S., & Freund, L. (2010). Information interaction in 140 characters or less: genres on twitter. *Proceedings of the third symposium on Information interaction in context*, IIiX '10 (pp. 323–328). New York, NY, USA: ACM. doi:10.1145/1840784.1840833

Wu, W., Zhang, B., & Ostendorf, M. (2010). Automatic generation of personalized annotation tags for Twitter users. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, HLT '10 (pp. 689–692). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1857999.1858100