

Understanding Library User Engagement Strategies through Large-Scale Twitter Analysis

Hongbo Zou^{*}, Hsuanwei Michelle Chen[§], Sharmistha Dey^{*}

^{*}Information Systems School, Queensland University of Technology, Brisbane, AU

[§]School of Information, San José State University, San José, US

Email: hongbo.zou@hdr.qut.edu.au; hsuanwei.chen@sjsu.edu; s.dey@qut.edu.au

Abstract—Public libraries are increasingly using social media to connect their users in an innovative way. Librarians make use of social media with the purpose of “being part of their communities”, or promoting libraries’ services and events. Social media has become a significant platform for libraries to create their own participatory services emphasizing engagement with users. However, there has been little empirical investigation into the success of social media use by libraries. In this paper, we study the role of a recently popular social media, Twitter, in engaging users with a focus on public libraries. We use topic-modeling techniques to classify the library user engagement strategies into four categories – literature exhibits, engaging topic, community building, and library showcasing. These four engagement strategies are re-examined with the tweets collected from ten public libraries over three months. The tweets topic distribution of every library is discussed in the paper. Finally, the impacts of every strategy on user engagement have been evaluated by users feedback on every tweet. Through the data mining of public libraries’ tweets, we aim to explore how user engagement strategies are used by the libraries on Twitter and suggest the best practices for libraries on social media initiatives to engage their users effectively.

Keywords—Big Data; User Engagement; Social Media; Data Mining;

I. INTRODUCTION

Over the last few years, social media has become one of the main ways for users to communicate with their friends and share the information with the public online. Millions of active users post and view hundreds of messages through various social media everyday to keep up with topics of their interests. Such online community naturally grows into one of the most effective ways of reaching out to customers. Many businesses, cultural organizations and social institutions are leveraging social media to achieve their strategic goals. According to research that assessed the social media activity of the top 100 most valuable global brands, those that were most socially active saw an 18% increase in revenue during the previous year, while the least active experienced a 6% revenue decrease [10]. With this trend, libraries are also using social media to connect with their users in an innovative way [21]. According to a study published in November 2013 by the Library Research Service (LRS), 93% U.S. public libraries in all the population categories had social media accounts in 2012 [30][37]. Librarians make use of social media with the purpose of “being part of their communities” [8] or “promoting libraries

services and events” [6]. Such trend perfectly matches the new proposed concept of the participatory library, which suggests that the library should engage in conversations with its community and that these conversations should inform how the library operates [13]. Therefore, social media has become a significant platform for libraries to create their own participatory services emphasizing engagement with users.

By having social media channels that are always open and participating in conversation with users, the library is able to constantly and effectively evaluate and refine its programs, products and services to ensure that the users are getting what they need [2]. Libraries can take advantage of such social media channels to invite participation, with active rather than passive participation being the goal [23]. Passive participation is when the library provides excellent content and simply asks the user to comment, while active participation involves the library inviting its users to create a community with the library and to help in shaping its direction, co-authoring content and engaging with other users to form a vocal community of users [23]. Social media creates an innovative method for library participatory service, which provides an effective way of connecting library users. However, how to apply different user engagement strategies to building better channels and engaging more users is still being questioned.

Among many social media platforms, Twitter has rapidly grown to a popular social network in recent years and provides a large number of real-time messages for users. Twitter becomes the most popular microblogging service and an important social network with over 500 million registered users as of July 2014. It allows users to share information with their friends or the public by posting text messages of up to 140 characters, which are called tweets. Most users update their Twitter messages frequently and over 500 million tweets are generated per day. Users can access a filtered timeline of tweets from specific persons by explicitly following them. The retweet mechanism allows users to share information with their followers and accelerate the spread of information in the social network. On Twitter, users follow celebrities, friends or anyone else they are interested in knowing and benefit greatly from the fresh information of these followees.

The ability to discover users’ interests can greatly help libraries to engage users through more strategic interactions. Twitter provides a bi-directional way to connect the libraries with their users. For example, libraries could post a short

message of newly arrived collections and then collect the messages posted on the twitter, such as favorites and retweets, to identify the popularity of a book theme. The library could then follow up by posting more topics under the same theme to attract more users. Users, in response, could follow the library's tweets and comment on the tweets with feedback, such as "retweet" and "favorite". For example, New York Public Library (NYPL) could organize a weekend reading event related with the novel "Pride and Prejudice" after they received over 47 favorites and 66 retweets with the tweet "Behold: two centuries of Pride and Prejudice covers! Which one is your favorite". Such user feedback can be considered a significant resource for libraries to customize their services to better engage their users. Therefore, using Twitter is considered a way of active participation for library user engagement, in which the library can enter into the space of the user and better understand the needs of the user, rather than waiting for the user to come to it. While Twitter provides another powerful avenue for libraries to interact with their users, there are, however, many challenges that need to be addressed for social media initiatives and active participation in the libraries. For instance, how would libraries post the new tweets with intelligently designed themes to help libraries better understand user interests? How would libraries create the tweet with well-chosen words to easily transform the idea to users? Is there a way for libraries to gain more knowledge about their users from mining online user behavior and accordingly enhance libraries' services and user experience? These challenges motivated us to propose this research, aiming to explore how user engagement strategies are used by the libraries on Twitter and suggest strategic implications for social media initiatives in libraries.

The remainder of this paper is organized as follows. Related work is discussed in Section II. Section III introduces the collected tweets data set used in our study. In section IV, the tweets are classified into four categories for user engagement strategies evaluation. Section V presents the analysis performed on the collected data set and findings derived from the results. We conclude the paper in Section VI.

II. RELATED WORK

While libraries have traditionally been user focused, the participatory library expands on the radical trust and gives the users more ability to guide the direction of the library service [20]. The public library of the future involves close contact between the library and its users. This participatory library is engaged in conversation with users [25]. By engaging in conversation with users, the library develops knowledge about them that can inform development and delivery of services and collections [14]. This conversational idea also supports the notion of user-driven change, which is often cited as one of the core principles of the future library [2][35]. Social technologies can support the key ideas that underpin the idea of a participatory library service: user-centred change, participation from users in developing service, and continual re-evaluation of services [2]. Social media also allows the library to enter into the space of the user, rather than waiting for the user to come to it. The library then begins actively seeking out conversations and participation and is able to speak with people it may otherwise not reach [14][39]. The information

and feedback that users provide is the "single best tool" that public libraries have to ensure that they remain relevant [2]. If the key role of the librarian is to "improve society through facilitating knowledge creation in their communities" [13], then librarians must come to understand that a participatory environment is key to facilitating knowledge creation. Social media provides a ready-made communication channel that the library can use to create user engagement and move towards a participatory service [11]. However, there is still a gap between libraries and users through social media communication. The gap is that the topics created by libraries on social media don't necessarily match users' topics of interest in most cases, so the question is: how are libraries using social media to create participatory networks that foster knowledge? [14].

Using social media for library management is an emerging topic, which has gained increased attention in both academia and practice in recent years. Social media can facilitate communications and engagement on library collections and services. Rutherford [24] and Tiffen and England [26] suggest that some libraries are using social media to develop communities and to personalize interactions between the library and users. Tools such as Facebook and Pinterest have been used to build relationships and rapport with client groups [17][22][40], to promote libraries [33], and to provide better information services [16]. The use of social media tools to communicate and to increase engagement can have powerful and positive effects on repeated library visits, rapport-building, referrals or positive feedback [26]. Twitter provides another new Internet venue to market a library's online brand and impression. There are many libraries that have already created their Twitter communities to connect with their users. While Twitter provides a great avenue for sharing and promotion, it does have its words limits. The library can only post short messages, images or videos, and it is not conducive to detailed discussions. How to understand users solely from responses to social media is a challenging issue and requires an in-depth, innovative data analysis approach.

III. DATA COLLECTION

To study what user engagement strategies are used by libraries, we observe ten public libraries in the US. These ten libraries are chosen with a goal of being geographically distributed to different locations. In Table 1., we first present the descriptive data of the 10 selected libraries, including their names, locations, join date, total tweets, and followers. The data has been collected on Oct 21, 2014. The dataset collected from the libraries contains over 10K tweets in a time span ranging from several months to several years.

In the table, New York Public Library (NYPL) serves as far more than a repository of bookish delights. It also happens to be a wonderful museum of American culture. The NYPL has posted 16.6K tweets on Twitter since Nov. 2008. On average, NYPL posts 7.6 tweets every day (excluding holidays closings). Los Angeles Public Library (LAPL), Birmingham Public Library (BMPL) and Seattle Public Library (SEPL) all have posted over 11K tweets since join date. For LAPL, in particular, the average posts reach up to 8.5 tweets per day. San José Public Library (SJPL), Houston Public Library (HTPL), San Francisco Public Library (SFPL), California State Library

(CAPL), Columbus Library (OHPL), and North Dakota State Library (NDPL) have posted less than 3.2 tweets per day on average. Both SFPL and HTPL have over 6K tweets and over 8.8K followers since join date. OHPL has 4,637 tweets and 15.9K followers since Dec. 2008. Each of the total number of tweets of SJPL, CAPL, and NDPL is less than 2,500. However, the tweets have attracted more than 1,500 followers, with a

maximum of 2,814 followers. Twitter site features items of interest from the SJPL, including the tweets that provide a glimpse into their local, regional, and California history collection, and their collection of historic and rare children's books. The tweets of SJPL interest and benefit almost 2,039 followers.

TABLE I. DATA COLLECTED FROM 10 SELECTED LIBRARIES

Library Name	Location	Join Date	Total tweets	Followers
New York Public Library (NYPL)	New York City, NY	Nov. 2008	16.6K	352K
San José Public Library (SJPL)	San José, CA	Oct. 2009	1,075	2,039
San Francisco Public Library (SFPL)	San Francisco, CA	Mar. 2009	6,201	8,847
Los Angeles Public Library (LAPL)	Los Angeles, CA	Apr. 2009	16.6K	9,989
Birmingham Public Library (BMPL)	Birmingham, AL	Mar. 2008	13.6K	10.7K
California State Library (CAPL)	Sacramento, CA	May 2009	2,393	2,814
Seattle Public Library (SEPL)	Seattle, WA	Jun. 2010	11K	10.9K
Houston Public Library (HTPL)	Houston, TX	May 2007	7,719	12.9K
Columbus Library (OHPL)	Columbus, OH	Dec. 2008	4,637	15.9K
North Dakota State Library (NDPL)	Bismarck, ND	Oct. 2009	2,018	1,462

^a. The data has been collected on Oct. 21, 201

In summary, the “followers” of half of the collected public libraries exceed 10K people. Moreover the “followers” of the world-class library, such as NYPL, reach almost 352K people. All of these Twitter feature numbers are showing a trend that Twitter is being increasingly used by public libraries to exhibit their resource and service and engage their users. Such trend is prompting the libraries to potentially tailor their user engagement strategies based on the information received from Twitter.

IV. USERS ENGAGEMENT STRATEGIES CLASSIFICATION

While Twitter gives the libraries a free reign to select and organize the content they posted on the Twitter site, a well-studied users engagement strategy can help the libraries provide many high-quality participatory services online to better understand the needs of their users and respond with more interesting content. Planning a good users engagement strategy should include an in-depth study on how to use Twitter for the library users engagement.

Although there has been much discussion of the use of social media in the literature, there has been little empirical research undertaken related to library user engagement strategies [2][24]. To understand how public libraries are using Twitter to generate conversations and move towards a participatory service, we categorize the libraries' user engagement methods into four categories: *literature exhibits*, *engaging topic*, *community building*, and *library showcasing*. Such categorization is based on the engagement purpose, rather than the posted content. “*Literature exhibits*” mainly shares book covers, historic archives, precious literature, video recodes, and any documentation can be found in the library. “*Literature exhibits*” services on the traditional library purpose,

in which the library delivers the information to the users, and users passively consume library posts. “*Engaging topic*” needs the librarians to create appealing topics and contents for users. In the meantime, users can actively retweet/comment their favorite contents in response to improve the topic. “*Community building*” provides a way for libraries to directly interact with the users in their community. Libraries aren't just delivering the information to patrons, they're also using Twitter to create a virtual club or forum and launch the special topic for discussion. Any library community concerned topic could be launched and discussed between users and library. “*Library showcasing*” aims to keep the users staying up-to-date about what their local library is doing, what new programs have been launched by the library, etc. The libraries can use “*library showcasing*” to collect ideas and suggestions on their local services and events. This classification has been proposed and discussed in our previous study on “quantitative analytics for library user engagement strategies through social media” [38]. Below we present more details about these four categories.

A. Literature Exhibits

The purpose of “*Literature exhibits*” is to share book covers, historic archives, precious literature, voice recodes, and any documentation that can be found in the library online. This engagement category is a new legacy of the traditional mission of the library, in which the library needs to make the collection of sources of information and resource accessible to a defined community for reference or borrowing. This category includes several user engagement methods, such as posting book covers, showcasing historic archives, sharing new acquisitions, and offering access to digital media [9].

Posting book covers is a method that the librarians share book covers, especially those from new books, special collections, and kid-friendly materials. It can be a great way to attract readers to books they might not have otherwise checked out. Showcasing historic archives is a method that the libraries present the historic literature much more than books in their archives. Sharing new acquisitions keeps patrons in the loop about the library's latest arrivals. Offering up access to digital collections offers a new channel to libraries sharing their digital collections through Twitter.

B. Engaging Topic

"Engaging topic" is a key usage of social media by libraries to engage the users to participate in the library services. This category needs the librarians to create visually appealing topics and contents for users. In the meantime, users can actively create or retweet their favorite contents in turn to improve the topic. This category includes some user engagement methods, such as creating a reading list, encouraging kids and teens to read, collecting learning materials for parents, and running reading programs [9].

Creating reading lists is a method that the librarians create appealing reading lists for the users, and users can post and retweet their favorites to improve the lists in turn. Encouraging kids and teens to read stands for an especially unique purpose that the library collects crafts, books, and materials that appeal to teens. Collecting learning materials for parents helps parents who want to get their kids into reading but just don't know where to look for the right materials to do so. Running reading programs is a method for librarians to find inspiration for summer reading programs and for parents and kids to keep up with the latest reads, fun integrated projects, and much more. Many are already sharing ideas for the year's summer reading programs, which is perfect for parents who want to plan ahead and librarians who are looking for creative ways to keep things fresh.

C. Community Building

"Community building" provides an active way for libraries to directly interact with the users in their community. This category includes some user engagement methods, such as helping patrons start book clubs, building a community of libraries online, and creating collaborative boards with patrons [9].

Helping patrons start book clubs is a great way to not only read more but to also get a chance to discuss a book more in-depth. Some public libraries are supplying links to book clubs kits that provide all the essentials for understanding and analyzing some of the most popular reads of the past few years. Building a community for libraries online builds a new channel for libraries to connect to other libraries and keep track of all the libraries that are showcasing some of their most unique and exciting ideas. Creating collaborative boards with patrons provides a way for libraries to interact with those in the community, sometimes in very fun and creative ways. Libraries create a topic and ask patrons to comment it. Users browse the post and discuss the topic with their idea. With the analysis of feedbacks, libraries can better understand the needs and thoughts of the users, and improve their service.

D. Library Showcasing

"Library showcasing" aims to keep the users staying up-to-date about what their local library is doing, what new programs have been launched by the library, etc. This category includes some user engagement methods, such as promoting library activities, collecting ideas for library displays, and exhibiting pictures of the library [9].

Promoting library activities keeps library users staying up-to-date about what the local library is doing. Many libraries are showcasing fliers about their events right on the image-centric site, showcasing everything from lectures, to job help, to author visits. Collecting ideas for library displays exhibits the library as a magical place for young kids, especially when librarians are creative in designing reading areas, bulletin boards, and other fun places in the library. Showing pictures of the library exhibits the cool stuff of the library and what the library is doing, especially after undergoing a big renovation project. Many of library facility related materials could be showcased, such as photos of their great murals, DVD sections, meeting rooms, children's department, and more. Library users can browse the posted tweets and better understand the library in their community, and give the improvement suggestions for library renovation.

This study examines the tweets collected from public libraries under these four categories. The tweets are clustered and grouped under these four categories, as shown in Table II, formulating a scheme to classify the engagement strategies on libraries' tweets.

V. DATA ANALYSIS

To figure out what user engagement strategies are used by libraries, the collected tweets need to be classified into different "topics". As a type of statistical model for discovering the abstract "topics" that occur in a collection of documents, topic modeling provides a powerful tool for classifying short text messages [12]. In this section, we start with tweets classification with topic modeling on the collected data set. With tweets classification, we check four user engagement strategies in every library. The tweets distribution shows engagement strategies using trends and details of every library. The effectiveness of every engagement strategy has been checked by users' responses analysis. Finally, we analyze 10 tweets with top responses. Such top 10 tweets give us more insights on the engagement strategies. The collected tweets data set includes posts tweeted by libraries or users. The number of user responses to every collected tweet includes two features: favorites and retweets. Twitter favorites is a feature on Twitter that allows you to mark a tweet as a favorite to easily see it later. A retweet means a tweet that had been forwarded or "resent" on Twitter by someone, but was originally written and sent by someone else. Such two features are good sources and indicators for users response study – a tweet with a good engagement strategy will most likely generate high response rate on favorites and retweets. The numbers of favorites and retweets are examined in the study to verify the effects of four engagement strategies.

A. Classification with Topic Modeling

Topic modeling is defined as a machine learning technique that looks for patterns in the use of words and it is an attempt to inject semantic meaning into vocabulary, in which a “topic” consists of a cluster of words that frequently occur together [1]. Topic models provide a simple way to analyze large volumes of unlabeled text [28]. Using contextual clues, topic models can connect words with similar meanings and distinguish between uses of words with multiple meanings. The tool we used in our topic modeling is a machine learning toolkit for language named “MALLET” [18]. It provides an efficient way to build up topic models based on Latent Dirichlet Allocation (LDA) model [1]. In natural language processing, LDA is a generative probabilistic model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar.

In this study, MALLET takes every tweets and looks for patterns in the use of words; MALLET helps us attempt to inject classified engagement strategies into vocabulary. MALLET does not know anything about the meaning of the words in a text. Instead, MALLET assumes that any tweet is composed by selecting words from possible baskets of words where every basket corresponds to an engagement topic. If that is true, then MALLET becomes possible to mathematically decompose every tweet into the probable baskets from whence the words first came. MALLET goes through this process

repeatedly until it settles on the most likely distribution of words into baskets, which we call engagement topics.

To apply topic models to discovering the underlying thematic structure in the collected tweets, we need to choose an appropriate number of topics (k) to reduce the perturbations in clustering. Choosing too few topics will produce results that are overly broad, while choosing too many will result in the “over-clustering” the collected data into many small, highly-similar topics [1]. In this study, we select a bunch of k ($= 4, 8, 12, 24$) values to manually check the stability of the output with repeated clusterings. With the stability test, we finally set $k = 4$ in the data processing. Such number setting keeps the maximum distance among the topics and generates the consistent classification with repeated clusterings. The other settings ($k = 8, 12, 24$) for the number of topics just over-cluster the collected data into fine-grained similar topics. Due to the space limit, we only list the associated words with $k = 8$ to show such over clustering in Table II.

With the setting $k = 4$, four terms baskets have been calculated by MALLET. Such four baskets are mapped into the classification of engagement strategies described in “Users Engagement Strategies Classification” section IV. We list samples of most related words for each topic in Table II. The very common English adverbs, conjunctions, pronouns and prepositions, have been ignored by MALLET with “remove-stopwords” option. So, the words list doesn’t include any high frequent stopwords, such as the, and, both etc.

TABLE II. LIST OF TOPICS AND TOP ASSOCIATED TERMS

Literature Exhibits		Engaging Topic		Community Building		Library Showcasing	
archives books list read bestselling articles cookbook benefit photo cover		readinglist school excited encourage happy love wisdom enjoyed parents favorite		join celebration libraries opening program party creative story questions today		public librarian visit staff branch exhibit welcome opening event program	
books read list bestselling picture history cookbook author	reads info archives articles news photo read check	readinglist excited readeverywhere learning happy wisdom encourage	encourage school kids parents favorite blog http show	community party art family summer program opening today	join fun celebrate creative club check questions story	public event food opening visit event library Friday	staff exhibit opening exhibit visit friends park branch

Table II shows the topic name and the samples of most related terms. After the topic modeling, we have the word clusters for each topic which has not yet been labeled. To label every word clusters, we re-checked every engagement strategies described in Section IV. Literature exhibits is to share any documentation that can be found in the library online. There are over 80% top N words, such as bestselling, archives, articles, cookbook, cover, books, read, photo etc.

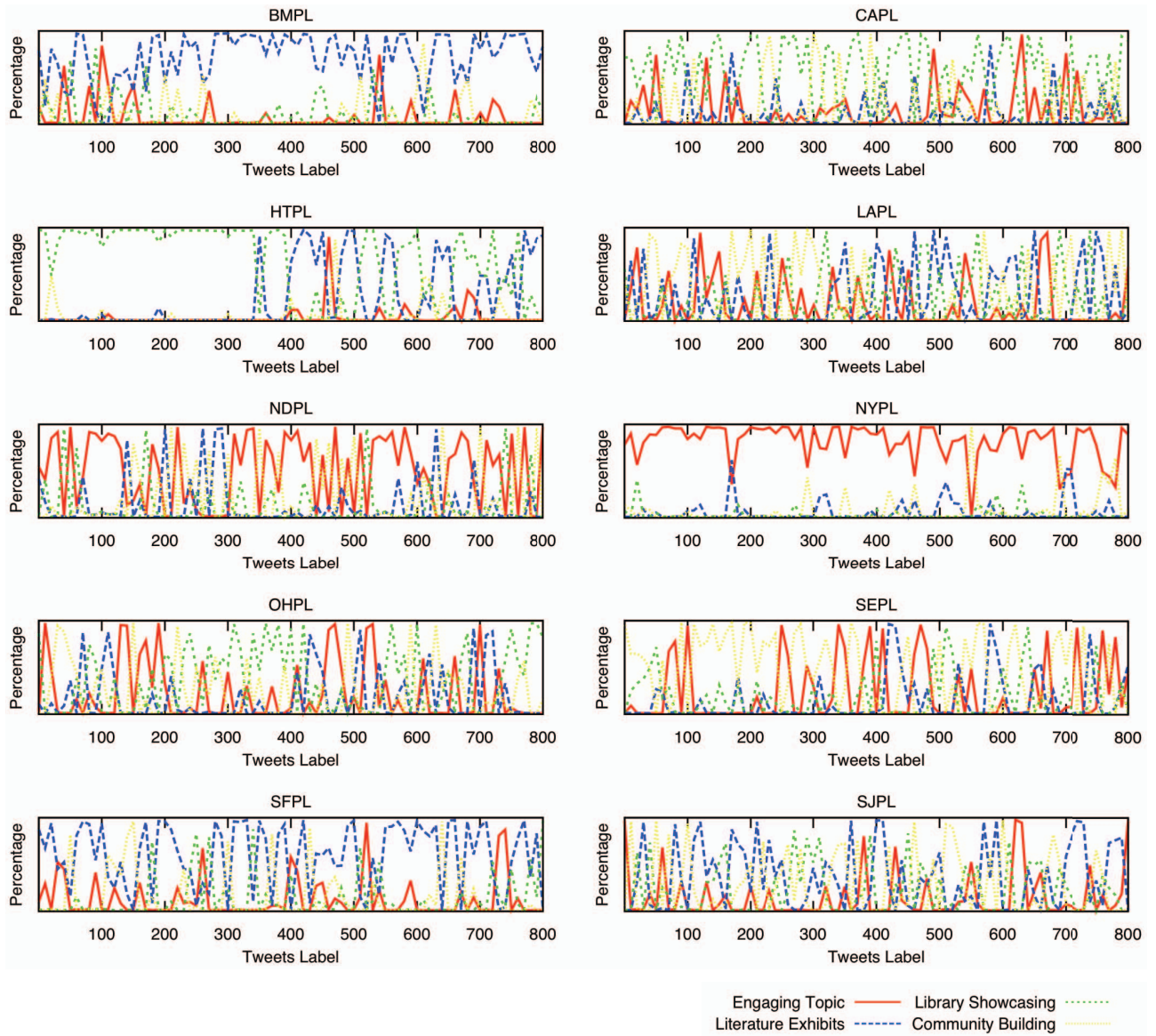
directly serve the purpose of literature exhibits. In addition, we randomly check over 100 tweets in 10 libraries, and there are over 80% tweets perfectly matching the topic of literature exhibits. For example, the tweet “ Hungry? Whet your appetite with our menu archive, and dream of a time where a steak cost 75 cents!” posted by NYPL directly show the steak menu collection for users. OHPL shares an old photo with users in the tweet “ Historic Photo Friday. Snow covered Ohio

Statehouse in 1898”. SFPL retweets a Harry Potter covers tweet “Kick off your shoes and peruse these 35 HARRY POTTER covers from around the world” on its twitter. We also check the other three engagement strategies, engaging topic, community building, and library showcasing with words and example tweets check, two steps check. Based on the checking results, we match the labels with our classification – Literature Exhibits, Engaging Topic, Community Building, and Library Showcasing, which are defined in “Users Engagement Strategies Classification” section. The strategy names have been listed in Table II as topics.

In the clustering, every tweet can be composed by any selecting words from four possible topics. The MALLET

calculates the conditional distribution (also called the posterior distribution) of every tweet under four different topic, rather than immediately decides a strategy for every tweet. Therefore, every tweet has four proportions to show the possibility of each topic that the tweet could belong to. In our study, the topic with the primary proportion is selected as the principal topic of the tweet. For example, if tweet 1 has four proportions: 0.011, 0.788, 0.014, 0.185 for topic 0, topic 1, topic 2, and topic 3 respectively, topic 1 (with 0.788 proportion) is the principal topic of the tweet 1. With the principal topic analysis, Fig. 1 shows the engagement strategies changes in different libraries. Each sub-figure shows the changing trend of one library, which includes four changing curves for the four topics. To present the figures clearly enough to observe the underlying trends, we

Fig. 1. Strategy Changes in Different Library.



only illustrate the conditional distribution of 800 tweets for every library. The observed trend remains in the rest tweets.

In Fig. 1, most libraries are mainly using literature exhibits and engaging topics to show their information resource and offer services. Among the 10 libraries, NYPL shows its very distinct trend and dedication to creating engaging topics to engage its users. HTPL primarily uses library showcasing to show its resource and service in the first part of our analysis, and this trend has shifted towards more literature exhibits tweets being posted for user engagement. Library showcasing and literature exhibits become two dominant strategies, which are mainly used by HTPL. CAPL takes library showcasing and community building as the majority of strategies for engaging their users. LAPL, NDPL, OHPL, and SJPL don't have a preference in a certain engagement strategy. Four engagement strategies have been equally employed to introduce their resources and services. BMPL and SFPL mainly use literature exhibits and engaging topic to show their resources and services. However, literature exhibits dominates most tweets posted by BMPL. SEPL evenly uses three topics – engaging topic, library showcasing, and literature exhibits to engage its users after a short community building period. In summary, most of the libraries are consistent with the engagement strategy adopted. HTPL and SEPL are the only two libraries that changed their strategies in different time periods.

Fig. 3. Libraries Tweets Distribution.

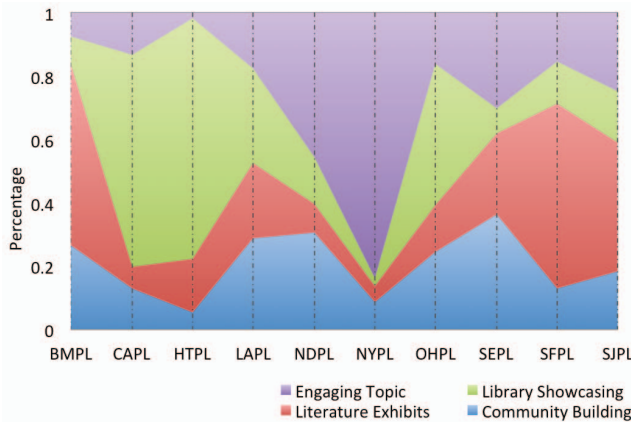


Fig. 1 shows the topics change trend of every library. Fig. 2 further shows how every library uses four strategies in a stacked area chart, in which the percentage distribution of topics in every library are compared against each other. The proportion of each topic is presented with different colors in the stacked area chart. Both NYPL and NDPL post many engaging topic-related tweets to engage the users. CAPL and HTPL use many tweets to show the topic related with the library. The rest of libraries don't show specific preference for any engagement strategy.

B. Users Response Checking

After discussion on the topic changes and distribution, we further examine the user feedback (favorites and retweets)

collected from every library to evaluate the engaging consequence of every library in this sub-section. Fig 3 shows a bar chart with overall number statistics on favorites and retweets. NYPL and SEPL have over 15,000 favorites and retweets from the users. SJPL, OHPL, NDPL, HTPL, and CAPL all have less than 2,000 favorites and retweets. SFPL, LAPL, and BMPL have a range between 2,000 and 4,000 favorites and retweets.

Fig. 2. Libraries Retweets and Favorites Comparison

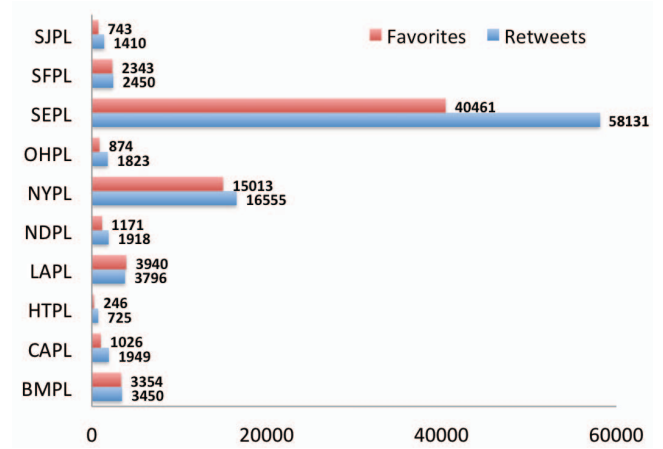


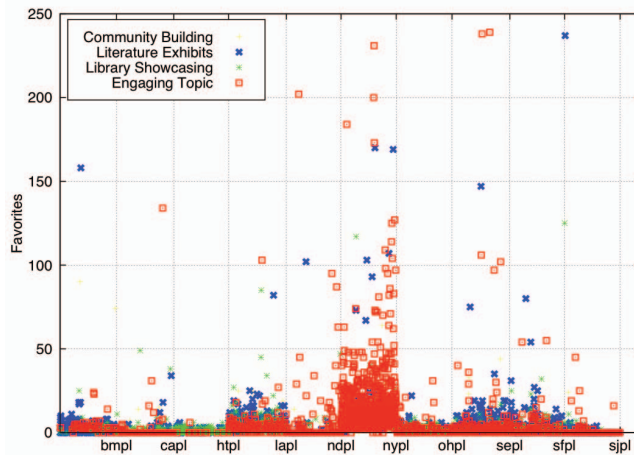
Table I shows the total tweets amount of every library. NYPL, LAPL, BMPL, and SEPL all have over 11K tweets, which means they have over 7 posted tweets per day on average. Such four libraries received much more users feedback than the other libraries. This observation resonates with the observation that active participatory services do help libraries better engage their users. In Fig. 3, NYPL uses engaging topic as the main strategy to attract its users, which in turn generates positive user feedback for its posts. Although engaging topic presents great potential in user engagement for NYPL, this strategy wasn't widely adopted by all libraries. For example, SEPL equally weights engaging topic, literature exhibits, and community building as its engagement strategies. However, it still received over 40,000 favorites and 58,000 retweets. SEPL shows a very special case in all libraries and prompts our next analysis, which focuses on figuring out which strategies really contribute to the majority of users feedback.

C. Response Distribution with Four Strategies

Our findings so far suggest that the four engagement strategies do not work equally well in terms of user feedback in every library. This subsection further investigates which strategies really work on the users feedback. Fig 4. and Fig. 5 show the favorites and retweets distributions for every library in detail. To help us better understand the distributions, we remove the popular tweets with over 250 favorites and retweets from the figures. The removed tweets have been specially analyzed in the next subsection as a special case.

Fig. 4 and Fig. 5 show a clear pattern that engaging topic contributes the most to user feedback. Literature exhibits also helps libraries better engage their users. Library showcasing

Fig. 4. Libraries Favorites Analysis.

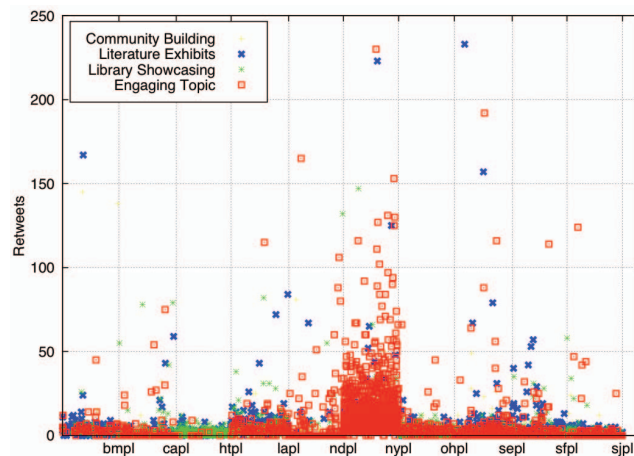


and community building received relatively less feedback from the users. NYPL is the most popular library after removing the top popular tweets. Small very popular tweets contribute most of the users feedback to SEPL in contract to NYPL. Such popular tweets will be further analyzed in the next subsection.

D. Case Study – Top 10 Tweets Analysis

In the collected data set, there are total 18 tweets, which have over 250 favorites and retweets, being removed from the analysis in Fig 4. and Fig 5. Such removed tweets should be further analyzed to help us understanding library user engagement strategies and their effectiveness in depth. Due to the space limit, we only list first 10 tweets with top favorites and retweets in Table III for analysis. In the table, every tweet contains tweet contents, library name, favorites, retweets, strategy, and date, six features for analysis. In order to facilitate discussion, we also label every tweet from 1 to 10 in the table. In Table III, most of the popular tweets come from SEPL. This

Fig. 5. Libraries Retweets Analysis.



explains why SEPL received over 40,000 favorites and 58,000 retweets. However, top 10 tweets contributes almost 28,300 favorites and 47,800 retweets for the users feedback of SEPL. We check top ten tweets and only the tweet with label 3 (label 3 tweet) is created by NYPL library itself. All of the other tweets are retweeted from the other popular twitter sites. Label 3 tweet shows 28 famous libraries and their quotes around the world. The libraries include New York Public Library, Trinity College Library of University of Dublin, Admont Abbey Library, Beinecke Rare Book & Manuscript Library of Yale University, Queen's College Library of Oxford University etc. This tweet is retweeted over 1,000 times by library lovers.

If the librarian cannot create an engaging tweet by himself, some very popular tweets can be retweeted to the library engage the users. Label 4 tweet announces a revival news about the TV show "Twin Peaks". Twin Peaks became one of the top-rated shows of 1990 and was a critical success both nationally and internationally. It captured a devoted cult fan base and became a part of popular culture that has been referenced in television shows, commercials, comic books, video games, films and song lyrics. On October 6, 2014, it was confirmed that the series will return for a nine-episode limited series to air in early 2016 on showtime. Under this big background, SEPL retweets the original producer of "Twin Peaks" - David Lynch's first hand announcement on the same day. The fans of the TV show responses for this news immediately. Label 8 tweet created by George R. R. Martin, who is an American novelist and short story writer in the fantasy, horror, and science fiction genres, and a screenwriter and television producer. Most of his fans response his joke on his new book. So, finding out the popular tweets and retweeting them also help engage the users if the library cannot originally create the tweet with engaging topic.

In the top 10 tweets, there are five tweets (label 1, 4, 6, 7 and 8) related with top-rated tv shows or movies. "True Detective" is an American television crime drama series. The first season premiered on Jan. 12, 2014, and consisted of eight episodes, concluding on March 9, 2014. The series has received widespread critical acclaim. The label 1 tweet releases the news on the new season renews. Label 7 tweet is related with the latest supernatural horror film – "Annabelle". This film premiered in Hollywood, Los Angeles, on Sep. 29, 2014. Label 6 tweet shows a historic picture of film "Psycho". "Psycho" is a 1960 American horror-thriller film. The film is praised as a work of cinematic art by international film critics and film scholars. Therefore, any vogue related topic could be easy way to create an attractive tweet and help library engaging users. Oct. 31 is Halloween. Label 5, 6 and 7 tweets show the related topic – horrible with this big holiday in U.S.

The quantitative analysis of top 10 popular tweets demonstrate that library users prefer participating in active services with popular trends created or retweeted by the library, and the library needs to spend much more time and efforts on contributing to a creative theme or design with popular trends on some tweets, rather than simply posting some messages on twitter.

TABLE III. TOP 10 TWEET IN THE COLLECTED DATA SET

Label	Contents	Library	Favorites	Retweets	Strategy	Date
1	"True Detective renews for another season..."	BMPL	1,453	1,214	Engaging Topic	27 May 2014
2	"RT to spread the word → On Sunday, entrance fees will be waived for all national wildlife refuges"	CAPL	422	694	Engaging Topic	10 Oct. 2014
3	"28 Beautiful Quotes About Libraries"	NYPL	860	1,132	Engaging Topic	2 Sep. 2014
4	"Dear Twitter Friends... it is happening again."	SEPL	17,737	29,655	Engaging Topic	6 Oct. 2014
5	"I'll just be wearing normal clothes for Halloween. Aren't our own identities the biggest masquerade of all?"	SEPL	767	705	Engaging Topic	19 Oct. 2014
6	"Anthony Perkins on the set of Psycho, 1960"	SEPL	1,945	964	Literature Exhibits	19 Oct. 2014
7	"The True Story Behind 'Annabelle' http://allday.com/g/anna "	SEPL	994	758	Engaging Topic	18 Oct. 2014
8	"I'm in a rush to finish the next book, so the entire second half will be written from Hodor's perspective."	SEPL	10,671	18,249	Engaging Topic	17 Apr. 2014
9	"This is what we see looking down while being inside an #aurora."	SEPL	4096	3604	Engaging Topic	2 Sep. 2014
10	"Need a recipe? We have 16,000 of them. The entire NYT recipe archive is now yours to cook"	SFPL	520	415	Literature Exhibits	19 Sep. 2014

^b. The data has been collected on Oct. 21, 201

VI. CONCLUSIONS

The social media is changing the traditional functionality of libraries. The library has been converted from a book reference place to information sharing and communication center. Twitter provides a good place to post, collect and classify the information, which can greatly engage library users with active participatory. However, Twitter doesn't create a fully bi-way communication channel between libraries and users. Users can post their comments or retweet the posts to show their interests. This direction is not enough to make librarians to fully understand the needs of the users in a timely fashion. In this paper, we classify the library user engagement methods into four categories – literature exhibits, engaging topic, community building, and library showcasing. The tweets of 10 selected U.S. public libraries have been collected and analyzed. Based on the topic modeling analysis, we focus our research specifically on the two categories: literature exhibits and engaging topic. The conclusions and suggestions have been derived from our quantitative analysis on the case study with four popular board themes.

Twitter provides a new platform for libraries to engage their users. However, such engagement mode needs to be carefully studied and prepared. In this research, we find that tweets with vogue-related topics receive more attention from users. Even in the Librarian's List the self-wrote annotation can greatly help users to pick the books. Libraries should not and could not use Twitter sites only as a new legacy of the traditional mission of the library, in which the library needs to make the collection of sources of information and resource accessible to a defined community for reference or borrowing. The library should creatively launch some popular trends related topics and share newly created themes for users. Favorites, retweets, and comments can help the library to better

understand the needs of their users and better improve their service to engage the users.

REFERENCES

- [1] D. Greene, D. O'Callaghan, P. Cunningham, How Many Topics? Stability Analysis for Topic Models, Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science Vol 8724, 2014, pp 498-513
- [2] K. Cahill, User-generated Content and Its Impact on Web-based Library Services, Oxford: Chandos Publishing, 2009.
- [3] D. M. Blei, Surveying a Suite of Algorithms that Offer a Solution to Managing Large Document Archives, Communications of The ACM, DOI:10.1145, 2012.
- [4] K. Cahill, User-generated Content and Its Impact on Web-based Library Services, Oxford: Chandos Publishing, 2009.
- [5] M. Casey, L. Savastinuk, Library 2.0: Service for the Next-generation Library, Library Journal, 131(14), 40-44, 2006.
- [6] L. Charnigo, P. Barnett-Ellis, Checking Out Facebook.com: The Impact of a Digital Trend on Academic Libraries. Information Technology and Libraries, 26(1), 23-34, 2007.
- [7] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia. Who is tweeting on Twitter: Human, Bot, or Cyborg? Proceedings of the 26th Annual Computer Security Applications Conference. Association for Computing Machinery, New York, 21-30, 2010.
- [8] C. De Rosa, J. Cantrell, A. Havens, J. Hawk, L. Jenkins, B. Gauder, D. Cellentani, Sharing Privacy and Trust in Our Networked World: A Report to the OCLC Membership. Dublin, OH: OCLC Online Computer Library Center.
- [9] J. Dunn, 20 Ways Libraries Are Using Pinterest Right Now, URL: <http://www.edudemic.com/20-ways-libraries-are-using-pinterest-right-now/> Online, Published by Edudemic, 2012.
- [10] D. J. Factiva, Quick Study: Direct Correction Established Between Social Media Engagement and Strong Financial Performance, PR News, 2009.
- [11] J. Fernandez, A SWOT analysis for Social Media in Libraries, URL: <http://pqasb.pqarchiver.com/infotoday/doc/199913917.html> Online 35-37, 2009.

- [12] L. Hong, B. D. Davison, Empirical Study of Topic Modeling in Twitter, In Proc. Of 1 st Workshop on Social Media Analyticsw (SOMA '10), Jul. 2010.
- [13] R. D. Lankes, The Atlas of New Librarianship. Cambridge, MA: The MIT Press.
- [14] R. D. Lankes, J. Silverstein, S. Nicholson, T. Marshall, Participatory Networks: the Library as Conversation, Information Technology and Libraries, 12(4), 17-33, 2007.
- [15] B. R. Lindsay. Social Media and Disasters: Current Uses, Future Options, and Policy Considerations, Congressional Research Service Report for Congress, 2011.
- [16] L. Loudon, and H. Hall. From Triviality to Business Tool: The Case of Twitter in Library and Informaiton Services Delivery. Business Information Review, 27(4), 236-241, 2010.
- [17] D. Mack, A. Behler, and E. Rimland. Reaching Students with Facebook: Data and Best Practices. Electronic Journal of Academic and Special Librarianship, 8(2), 2007.
- [18] A. K. McCallum, Mallet: A Machine Learning for Language Toolkit, 2002. URL: <http://mallet.cs.umass.edu>.
- [19] M. Natalya, Taxonomy: Content Strategy's New Best Friend, URL: <http://johnnyholland.org/2012/04/taxonomy-content-strategys-new-best-friend/> Online, 2012.
- [20] L. C. Nguyen, H. Partridge, S. L. Edwards, Towards an Understanding of the Participatory Library, Library Hi Tech, 30(2) 335-346, 2012.
- [21] S. O'Dell, Opportunities and Obligations for Libraries in a Social Networking Age: A Survey of Web 2.0 and Networking Sites. Journal of Library Administration, 50(3), 237-251.
- [22] N. K. Phillips. Academic Library Use of Facebook: Building Relationships with Students. The Journal of Academic Librarianship, 37(6), 512-522, 2011.
- [23] M. Porter, D. King, Inviting Participation, Public Libraries, 46(6), 34-36, 2007.
- [24] L. L. Rutherford. Building Participative Library Services: the Impact of Social Software Use in Public Libraries. Library Hi Tech, 26(3), 411-423, 2008.
- [25] K. Smeaton, K. Davis, Using Social Media to Create a Participatory Library Service: an Australian Study. Library and Information Research, 54-76, Vol 38, No 117, 2014.
- [26] B. Tiffen, and A. England. Engaging with Clients and Personalising Services at UTS Library: Measuring the Value for Libraries and Their Clients. Australian Library Journal, 60(3), 237-247, 2011.
- [27] E. Walter. The Ultimate Guide to Pinterest for Every Brand. URL: <http://www.fastcompany.com/3015203/work-smart/the-ultimate-guide-to-pinterest-for-every-brand> Online, Posted by Fast Company, 2013.
- [28] D. Wang, D. Irani, and C. Pu. Is Email Business Dying?: A Study on Evaluation of Email Spam over Fifteen Years, EAI-Endorsed Transactions on Collaborative Computing. Vol 1, Issue1, 2014.
- [29] D. Wang, D. Irani, and C. Pu., SPADE: A Social-spam Analytics and Detection Framework, Social Network Analysis and Mining, Vol 4, Issue 1, 2014.
- [30] M. Wanucha, L. Hofschire, U.S. Public Libraries and the User of Web Technologies, 2012, Nov. 2013.ã
- [31] D. Wang, Analysis and Detection of Low Quality Information in Social Networks, 30 IEEE Internatinal Conference on Data Engineering (ICDE 2014), Chicago, 2014.
- [32] D. Wang, D. Irani, and C. Pu. A Study on Evolution of Email Spam over Fifteen Years, In Proc. Of 9 th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2013). Austin, 2013.
- [33] Z. D. Xia. Marketing Library Services Through Facebook groups. Library Management, 30(6/7), 469-478, 2009.
- [34] S. Yardi, D. Romero, G. Schoenebeck, and D. Boyd. Detecting Spam in a Twitter Network. First Monday 15(1): 1-4, 2009.
- [35] Y. Yu, H. Zou, W. Tang, L. Liu, and F. Teng, FlexTuner: A Flexible Container-based Tuning System for Cloud Applications, IEEE International Conference on Cloud Engineer (IC2E 2015), 2015.
- [36] H. Zou, Y. Yu, W. Tang, and M. Chen, FlexAnalytics: A Flexible Data Analytics Framework for Big Data Application with I/O Performance Improvement, Elsevier Journal on Big Data Research, Vol. 1, 4-13, 2014.
- [37] H. Zou, Y. Yu, W. Tang, and M. Chen, Improving I/O Performance with Adaptive Data Compression for Big Data Applications, IEEE International Parallel I/O Systems, Parallel and Distributed Processing Symposium (IPDPSW 2014), 2014
- [38] H. Zou, M. Chen, and S. Dey, Quantitative Analytics for Library User Engagement Strategies Through Social Media: A Perspective from Mining with Pinterest, In Submission.
- [39] H. Zou, M. Slawinska etc. FlexQuery: An Online Query System for Interactive Remote Visual Data Exploration at Large Scale, IEEE International Conference on Cluster Computing (Cluster'13), 2013.
- [40] H. Zou, XH Sun etc. A Source-aware Interrupt Scheduling for Modern IEEE International Parallel I/O Systems, Parallel and Distributed Processing Symposium (IPDPS 2013), 2013.