Visual Text Analytics for Social Media Conversations

by

Enamul Hoque Prince

B.Sc., Chittagong University of Engineering & Technology, 2007

M.Sc., Memorial University of Newfoundland, 2012

Doctoral thesis proposal

DEPARTMENT OF COMPUTER SCIENCE

The University Of British Columbia

(Vancouver)

November 2014

Abstract

In the last decade, there has been an exponential growth of asynchronous online conversations thanks to the rise of social media. Analyzing and gaining insights from such conversations can be quite challenging for a user, especially when the discussions become very long. Most of the works focused on either introducing advanced text analysis methods or developing new information visualization techniques, however only a little efforts have been made to tightly integrate them together. During my doctoral research, I aim to investigate how to integrate Information Visualization with Natural Language Processing techniques to better support the user's task of exploring and analyzing conversations. For this purpose, I consider the following approaches: applying design study methodology in InfoVis to uncover data and task abstractions; applying NLP methods for extracting the identified data to support those tasks; and incorporating human feedback in the text analysis process when the extracted data is noisy and/or may not match the user's mental model, and current tasks. Through a set of design studies, I aim to evaluate the effectiveness of my approaches of designing visual text analytic systems for social media conversations.

Table of Contents

Abstract . . . . **ii**

Table of Contents . . . . **iii**

List of Tables . . . . **v**

List of Figures . . . . **vi**

Acknowledgments . . . . **viii**

1 Introduction **1**
1.1 The Problem 1
1.2 Scope: Domains and Users 4
1.3 Overview 5

2 Completed Research Works **7**
2.1 ConVis: Exploring asynchronous conversation 7
2.1.1 Problem 7
2.1.2 Related work 8
2.1.3 Methodology 10
2.1.4 Results 13
2.1.5 Contributions 15
2.2 ConVisIT: Interactive topic modeling 15
2.2.1 Problem 15
2.2.2 Related work 17

2.2.3 Methodology 18
2.2.4 Results 24
2.2.5 Contributions 26

3 Ongoing and Future Work 27
3.1 Exploring Social media collections 27
3.1.1 Problem 27
3.1.2 Methodology 29
3.1.3 Expected contributions 37
3.2 Interactive mining and summarization of conversations 37
3.2.1 Introduction 37
3.2.2 Methodology 38
3.2.3 Expected Contributions 39
3.3 Design framework for visual text analytics 40
3.3.1 Introduction 40
3.3.2 Methodology 40
3.3.3 Expected contributions 43

4 Milestones 44

Bibliography 47

A Supporting Materials 55

List of Tables

Table 1.1 User categorization for asynchronous conversation. 5

Table 2.1 A set of tasks (phrased as questions) that a user may likely have to perform/answer while exploring a blog conversation to satisfy her information needs. 10

Table 2.2 Different possible topic revision operations. 19

Table 3.1 MacRumors blog dataset statistics 31

List of Figures

Figure 1.1 My PhD research falls into the cross section between Information Visualization, Natural Language Processing, and Human-in-the-loop computation. 3

Figure 2.1 An overview of the completed work. 8

Figure 2.2 A snapshot of ConVis for exploring a blog conversation: The Thread Overview visually represents the whole conversation encoding the thread structure and how the sentiment is expressed for each comment((middle-left); The Facet Overview presents topics and authors circularly around the Thread Overview; and the Conversation View presents the actual conversation in a scrollable list (right). Here, topics and authors are connected to their related comments via curved links. 13

Figure 2.3 An example showing: (a) The user clicked on a comment (the one with black horizontal outlines) in the Thread Overview. (b) As a result, the system automatic scrolled to the actual comment in the Conversation View. 14

Figure 2.4 Interactive topic modeling framework for exploring asynchronous conversation. 16

Figure 2.5 Three different user operations for topic revision 20

Figure 2.6 An example showing: (a) The user hovers over the topic 'military security' and decides to perform the split operation. (b) As a result, the topic moves to its left while the rest of the topics are pushed along the perimeter of the circular layout to create space for the new children. 23

Figure 2.7 An example showing: (a) The user decides to merge two topics by joining (indicated by orange color). (b) As a result, ConVisIT updates the topic organization where these two topic nodes are merged under the parent topic 'subject sql injection attack'. 24

Figure 2.8 An example showing: (a) The user decides to perform merge by absorption on two topics 'web sites' and 'prototype'. (b) ConVisIT updates the topic organization where the previous link from 'web sites' to 'army server' is removed, and then 'web sites' is absorbed into a more generic parent topic 'web programming' along with 'prototype'. 24

Figure 3.1 An overview of the on-going work. 28

Figure 3.2 An initial sketch of a visual interface for exploring a collection of conversations. **Left:** The overview consists of topics and authors to support faceted exploration and for each conversation, a Timeline graph shows the volume of comments over time with sentiment distribution to provide visual information scent (middle-left) [72]. **Right:** The selected conversation appeared along with an overview of its thread structure. 36

Figure 3.3 Design cycles of ConVis and ConVisIT. 42

Figure 4.1 Timeline for my doctoral research. 45

Figure 4.2 A summary of the completed and on-going work along with publication plan. 46

Acknowledgments

I thank my supervisory committee: Giuseppe Carenini, Tamara Munzner and Raymond Ng for helpful comments and suggestions. I also thank all the current and past members of our Natural Language Processing research group.

Chapter 1

Introduction

1.1 The Problem

Since the phenomenal adoption of Web-based social media, an ever-increasing amount of written conversations are generated everyday [6]. While email remains a fundamental way of communicating for most people, other conversational modalities such as blogs, microblogs and discussion fora have quickly become widely popular. These conversations are primarily asynchronous in nature, where the participants communicate online with each other at different times [41].

An asynchronous conversation such as a blog may start with a news article or an editorial opinion, and later may generate a long and complex thread as comments are added by the participants [6]. When a reader wants to explore such a large conversation, traditional social media sites present the original posts and subsequent replies as a paginated indented list. Thus the reader needs to go through a long list of comments sequentially, until her information needs are fulfilled. Going through such an overwhelming amount of textual data in this way often leads to information overload, i.e., the user finds it very difficult to get insights about the ongoing (or past) discussions [40]. The problem becomes even more compounded when the user is interested in analyzing multiple conversations that are discussing similar issues.

While both the Natural Language Processing (NLP) and the Information Visualization (InfoVis) community individually attempt to address this and similar problems, only little effort have been devoted to integrating NLP and InfoVis techniques in a synergistic way. For instance, several techniques for text mining have been developed recently, which include topic modeling to find the major themes [41], and opinion mining to understand the sentiment expressed within conversations [66]. However, the extracted topics and associated opinions often do not match the user current information needs. Moreover, the output of mining algorithms can be very complex, making it hard to be consumable to the users. On the other hand, existing InfoVis systems either visualize only metadata that do not reveal any content information, or visualize the results of simple, often inaccurate, NLP techniques that were not explicitly designed for conversational data. As a results, the user's ability to explore and analyze the conversation becomes limited.

The primary goal of my research is to develop a comprehensive understanding of how a combination of text analysis and interactive visualization can support users in exploring online conversations. The hypothesis is that *by tightly integrating NLP and information visualization (InfoVis) techniques, we can better support the user's task of exploring and analyzing conversations.* But how NLP and InfoVis techniques can be effectively integrated? More specifically, I pose the following research questions:

1. What tasks users want to perform and what metadata and text analysis results are actually useful to support these tasks?
2. How useful metadata and content can be extracted from the conversation?
3. How the extracted metadata and contents should be visualized to the user?
4. How can we support the user when (she realizes that) the current text analysis results is not helping her anymore?

My research falls into the cross section between three main research areas that are associated with my research questions: information visualization (Q1, Q3), natural language processing (Q2, Q4), and human-in-the-loop-computation (Q4). The overlap between these three areas defines the scope of my doctoral research, i.e., designing a visual text analytic system for asynchronous conversations (see Figure 1.1). The distinct role of each area in my research is as follows:

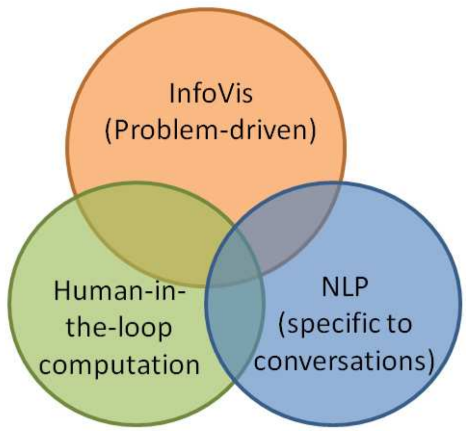

Figure 1.1: My PhD research falls into the cross section between Information Visualization, Natural Language Processing, and Human-in-the-loop computation.

-*Why InfoVis?* To address Q1 and Q3, I focus on applying human-centered design methodologies from the InfoVis literature (e.g., [54, 61]). Starting from an analysis of user behaviours and needs in the target conversational domain, such methods help uncover useful task and data abstractions. On the one hand, task and data abstractions can characterize the type of information that needs to be extracted from the conversation (Q1); on the other hand, they can inform the design of the visual encodings and interaction techniques (Q3). More tellingly, as both the NLP and the InfoVis components of the resulting system were designed by referring to a common set of task and data abstractions, they are more likely to be consistent and synergistic.

- *Why NLP for conversations?* To address Q2, I focus on devising and applying text mining and summarization methods specific to asynchronous conversations. Most of the existing visual text analytic systems use NLP methods that are originally devised for generic documents. These methods generally do not exploit the specific characteristics of asynchronous conversations (e.g., use of quotation, dialog acts), while it has been shown that text analysis results are more accurate when these specific characteristics are taken into account [41]. In order to address this limitation, I aim to apply text mining and summarization approaches that take advantage of the conversational features.

- *Why human-in-the-loop computation?* To address Q4, I focus on considering the human feedback in the text analysis process. The motivation for such an approach is that the results of NLP systems can be noisy and may not match the user's mental model, and current tasks. In such situations, I aim to support the user in providing feedback to the underlying NLP system, so that the results can better match her information needs.

1.2 Scope: Domains and Users

In the initial, exploratory phase of my research, I try to understand and analyze the broad range of domains, users, and data for asynchronous conversations with the aim of better defining the scope for the thesis. Here, first I provide an overview of the domains, users, and data, followed by the design scope of this thesis.

Conversational Domains: The phenomenal adoption of novel Web-based social media has led to the rise of asynchronous conversations in many different domains, ranging from blogs, to discussing forums, to social networking sites. Social news blog sites such as Reddit, Slashdot and Digg contain user-generated stories that are ranked based on popularity. Users can comment on these posts, and these comments may also be ranked. Some of these sites focus on a specific topic, such as MacRumors which is dedicated to the discussion of news relating to the Apple Inc. Finally, for many users, Microblogs such as Twitter and social networking sites such as Facebook have become part of their online life.

Users: As shown in Table 1.1, users in a conversational domain can be categorized into two groups based on their activities: (a) *participants* who have already contributed to the conversations, and (b) *non-participants* who have not contributed to the conversations yet. Depending on different user groups the tasks might vary as well, something that needs to be taken into account in the design process.

For example, imagine a *participant* who has expressed her opinion about a major political issue. After some time, she may become interested to know what comments were made supporting or opposing her opinion, and whether those comments require a reply right away. On the contrary, a *non-participant,* who is interested in joining the ongoing conversation on that particular political issue, may want to decide whether and how she should contribute by quickly skimming through a long

Conversation status / Users	Ongoing conversation	Inactive/archived post
Participant (Have some prior knowledge about the conversation)	• **Author of the initial post** Created the initial post and wants to know what people are saying about it. • **Other participants** Already joined the conversation (wants to get updated and possibly post new comments)	Wants to delve into the past conversations and to explore what was discussed, what she said, what other people replied, etc.
Non-participant (Possibly do not have any prior knowledge about the conversation)	• **Potential participant** wants to join the conversation • **Analyst** (wants to analyze the ongoing conversation, but does not intend to join)	Wants to analyze and gain insight about the past conversation.

Table 1.1: User categorization for asynchronous conversation.

thread of blog comments. Another group of users may include the analysts who does not wish to join the conversation, but may want to analyze and gain insights from conversations. For instance, a journalist may want to summarize the major arguments that were used to support or oppose the political issue. Another example is an analyst who wants to discover important insights from conversations and present to a policy maker for supporting the decision making process.

Scope: In this thesis, I would like to conduct design studies focusing on supporting a set of tasks of exploring and analyzing conversations, that are commonly performed across different domains and user types. However, for the purpose of design and evaluation, I would like to apply our system on the domain of blogs, because blog conversations often have complex thread structure with large number of participants and comments [41], making it a more challenging problem from visualization perspective. As we will discuss in the thesis, many of the tasks that we have identified, can be arguably generalized to other conversational domains.

1.3 Overview

In this thesis, I initiate a set of design studies with an aim to integrate InfoVis with NLP for developing visual text analytics interfaces to social media conversations. In my completed research, I focused on supporting a reader in exploring an archived blog conversation. Due to the large number of comments with complex thread structure [41], mining and visualizing a long blog conversation can become a challenging problem. Through an iterative design process, I developed two interactive interfaces for exploring blog conversations, namely ConVis and ConVisIT. An overview of these design cycles is provided in (Chapter Section 2).

In my ongoing and future research, I focus on supporting a user in exploring a *collection* of archived conversations. Exploring topics of interest, that are potentially discussed over multiple conversations, is arguably an even more challenging problem, as the volume and complexity of the conversational data increases. An overview of this planned work is provided in Section 3.1 and 3.2. Finally, during the last phase of my PhD, I plan to summarize findings and reflections from my studies into a generalizable design framework for visual text analytics as discussed in Section 3.3.

Chapter 2

Completed Research Works

During the summer 2013, I started designing a visual text analytic system, named ConVis, for exploring a blog conversation. ConVis presented topics, sentiment and a set of metadata to support the user in exploring and navigating through the conversation. Later, informed by the preliminary evaluation of ConVis, I extended the system to incorporate interactive topic modeling. Both works are documented in published and submitted research papers [36–38]. A visual summary of the work is presented in Figure 2.1 .

2.1 ConVis: Exploring asynchronous conversation

2.1.1 Problem

Traditional social media sites present the original posts and subsequent replies as a paginated indented list. Thus, the reader needs to go through a long list of comments sequentially, until her information needs are fulfilled. Consider a scenario, where a reader opens a blog conversation about Obama's healthcare policy. The reader wants to know why people are supporting or opposing *ObamaCare*. However, since some related discussion topics like *student loan* and *job recession* are introduced, the reader finds it hard to keep track of the comments about *ObamaCare*, which end up being buried in the long discussion.

To support the readers in dealing with this and similar situations we designed

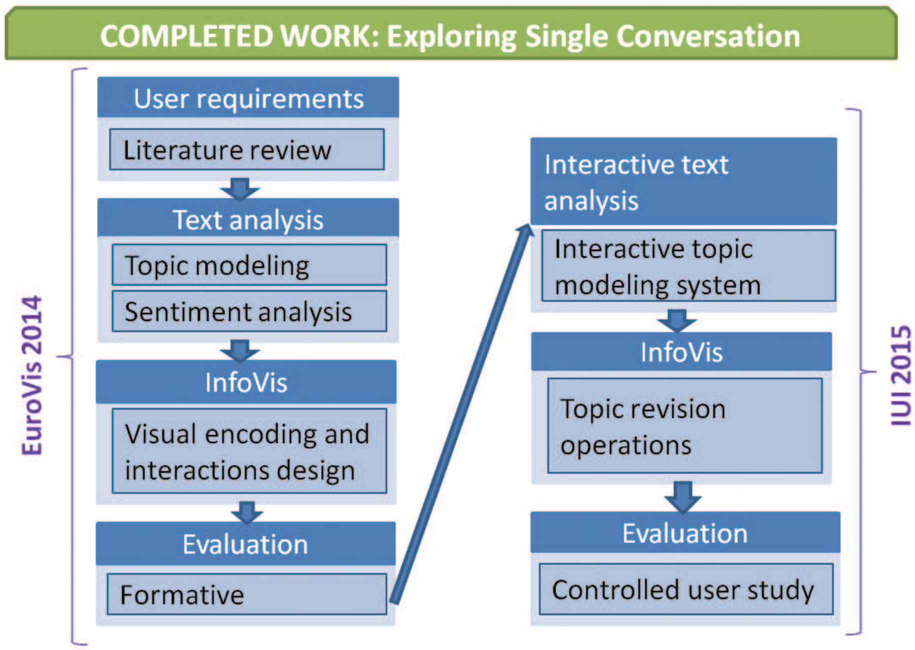

Figure 2.1: An overview of the completed work.

ConVis. ConVis tightly integrates interactive visualization with text mining techniques that are especially devised to deal with conversational data.

2.1.2 Related work

Previous work on visualizing asynchronous conversations can be classified into two categories: metadata-based and content-based visualization; depending on whether the focus of the research was more on visualizing system and user generated metadata (e.g., thread structure), vs. the results of some text analysis (e.g., finding topical clusters).

Metadata-based Visualization: Earlier works for visualizing asynchronous conversations primarily focused on revealing the structural and temporal patterns of a conversation [57, 70]. Typically, the goal was to effectively represent the thread structure of a conversation using tree visualization techniques, such as thumbnail metaphor (a sequence of rectangles) [70] and radial tree layout [57]. Various interaction techniques, such as highlighting user-specified search terms [70] and zooming into an area of the thread overview [57] were proposed to deal with space constraints for larger threads. Other works visualize various system and user generated metadata such as timestamp [21]; comment length and moderation score [55].

Even though metadata-based visualizations help to understand the social interaction patterns or the quality of the comments in a conversation, they may be inadequate to support users in most of the tasks shown in Table 2.1. For example, if the user is reading a political blog to know *"what do people think about Obama's recent healthcare policy?"*, knowing how nested the thread structure is or how many replies are made to a particular post would be insufficient.

Content-based Visualization: Recently, there has been more focus on performing content analysis of the conversations, such as identifying primary themes (or topics) within conversations [18, 60], and visualizing the content evolution over time [26, 68, 71]. Commonly, these approaches use probabilistic topic models such as Latent Dirichlet Allocation (LDA), where topics are defined as distributions of words and documents are represented as mixture of topics. For instance, the TIARA system applies the Themeriver metaphor [32], where each layer in the graph represents a topic and the keywords of each topic are distributed along time. From the height of each topic and its content distributed over time, the user can see the topic evolution. More recently, a hierarchical version of the Themeriver metaphor was also designed to explore the temporal changes of topics [26], by generating a topic tree based on computing the distance between the probability distributions of topics.

While there has been a clear trend of moving beyond using only metadata to an increasing use of text analysis within the interactive visualization process, current systems generally suffer from two fundamental limitations. First, they use generic text analysis techniques. Secondly, current systems only convey one type of mined information (e.g., either topic or opinion), thus limiting the user's ability to perform most of the tasks in Table 2.1. In this work, we have addressed both limitations.

2.1.3 Methodology

In order to tightly integrate interactive visualization with text mining techniques, I followed the nested model for visualization design [54], where I started by characterizing the domain of blogs. Here I provide a brief overview of the approach, the detailed description is provided in [36].

No	Questions (Q)	Topic	Author	Opinion	Thread	Comment
1	What this conversation is about?	X				X
2	Which topics are generating more discussions?	X				
3	What do people say about topic X?	X		X	X	X
4	How controversial was the conversation? Were there substantial differences in opinion?	X	X	X	X	X
5	How other people's viewpoints differ from my current viewpoint on topic X?	X		X	X	X
6	Why are people supporting/opposing an opinion?			X	X	X
7	Who was the most dominant participant in the conversation?		X		X	X
8	Who are the sources of most negative/positive comments on a topic?	X	X	X	X	X
9	Who has similar opinions to mine?		X	X		X
10	What are some interesting/funny comments to read?	X	X	X		X

Table 2.1: A set of tasks (phrased as questions) that a user may likely have to perform/answer while exploring a blog conversation to satisfy her information needs.

User requirements analysis: Blog reading has been extensively studied in the fields of computer mediated communications (CMC) [22, 43, 74], social media [30, 34], human computer interaction (HCI) [3, 18, 53], and information retrieval

[40, 44, 49, 52, 63]. This literature provides a detailed analysis of *why and how people read blogs*. Based on this analysis, we characterize the data and tasks in the domain of blogs and then identify the user requirements, which are finally translated into a set of design principles.

Tasks and Data: From the user requirement analysis, we compile a list of tasks (see Table 2.1) and associated data variables that one would wish to visualize to support these tasks. These tasks can be framed as a set of questions, for instance, ‘what do people say about topic X?’ The data variables (columns in Table Table 2.1) include: *Topic*, *Author*, *Opinion*, *Thread*, and *Comment*. We also compute average and maximum counts for different types of data to better understand what scale the visualization needs to deal with. These values are computed based on a set of Slashdot blogs which comes with human generated topic annotations [41].

Text analysis: According to our task abstractions in Table 2.1, most of the questions require the user to know the topics and sentiment information. Therefore, we apply both topic modeling and sentiment analysis to our conversation dataset. In the topic modeling phase our goal is to group the sentences of an asynchronous conversation into a set of topical clusters/ segments (*segmentation*), and then assign representative key phrases to each of these segments (*labeling*). We adopt a novel topic modeling approach that captures finer level conversation structure in the form of a graph called Fragment Quotation Graph (FQG) [41]. The FQG is exploited in both topic segmentation and labeling, as this conversational structure has shown to improve performance in both cases. To determine the sentiment polarity of each sentence of the conversation we used SoCAL [66], which has been shown to work well on user-generated content.

Visual encoding and interaction design: Once we have characterized our domain, we derive a set of design principles, which then guide the visual encoding and interaction techniques of ConVis. ConVis is primarily an overview + details interface, since this design has been found to be more effective for text comprehension tasks than other approaches such as zooming and focus+context [14]. The overview consists of the whole thread as well as the topics and authors of the conversation. The interactions between these views are performed in a coordinated way. Below, we describe the design of each component along with careful justification of crucial design decisions[1].

The Thread Overview visually represents each comment of the discussion as a stacked bar, where each stacked bar encodes three different metadata (comment length, position of the comment in the thread, and depth of the comment within the thread). A set of five diverging colors was used to visualize the distribution of sentiment orientation of a comment in a perceptually meaningful order, ranging from purple (highly negative) to orange (highly positive). Thus, the distribution of colors in the Thread Overview can help the user to perceive the kind of conversation they are going to deal with. For example, if the Thread Overview is mostly in strong purple color, then the conversation has many negative comments.

The primary facets of the conversations, namely topics and authors are presented in a circular layout around the Thread Overview (Figure 2.2). Both topics and authors are positioned according to their chronological order in the conversation starting from the top, allowing the user to understand how the conversation evolves as the discussion progresses. The font size of facet items helps the user to quickly identify what are the mostly discussed themes and who are the most dominant participants within a conversation. To indicate topic-comment-author relationship, the facet elements are connected to their corresponding comments in the Thread Overview via subtle curved links. These visual links allow the user to perceive the related entities more quickly and with greater subjective satisfaction than plain highlighting [64]. Finally, the Conversation View displays the actual text of the comments in the discussion as a scrollable list.

We designed a set of user interactions that can be easily triggered without causing drastic modifications to the visual encoding, thus allowing the user to comprehend the effect of interactions without much cognitive overload. The user can start exploring the conversation by hovering the mouse on topics, which highlights the connecting curved links and related comments in the Thread Overview. As such, one can quickly understand how topics may be related to different comments and authors. Then, if the reader becomes further interested in a specific topic/author, she can click on it. As a result, a thick vertical outline is drawn next to the corresponding comments in the Thread Overview. Such outlines are also mirrored in the

[1] A live demo of the system is available at: https://www.cs.ubc.ca/cs-research/lci/research-groups/natural-language-processing/ConVis.html

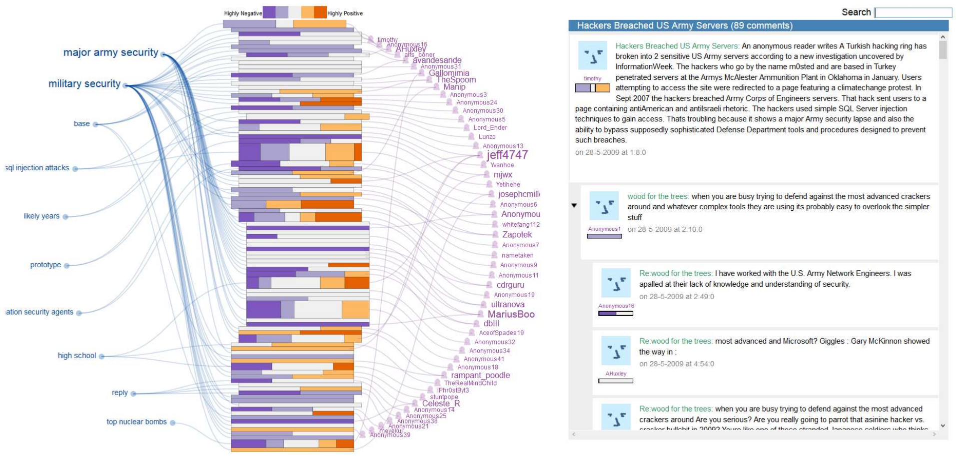

Figure 2.2: A snapshot of ConVis for exploring a blog conversation: The Thread Overview visually represents the whole conversation encoding the thread structure and how the sentiment is expressed for each comment((middle-left); The Facet Overview presents topics and authors circularly around the Thread Overview; and the Conversation View presents the actual conversation in a scrollable list (right). Here, topics and authors are connected to their related comments via curved links.

Conversation View. Besides exploring by the topics/authors, the reader can browse individual comments by hovering and clicking on them in the Thread Overview. In particular, when the user hovers over a comment its topic is highlighted, while when the user clicks on a comment, the actual text for that comment is shown in the Conversation View (by scrolling). In this way, the user can easily locate the comments that belong to a particular topic.

2.1.4 Results

During the design and implementation of ConVis, we conducted formative evaluations to identify potential usability issues and to iteratively refine the prototype. Once the prototype was completed, we ran an informal evaluation [45] with the aim to: 1) understand to what extent the overall visualization and its specific components are perceived to be useful by the potential users; 2) identify differences in how different users performed the tasks; and 3) solicit ideas for improvements and enhancements.

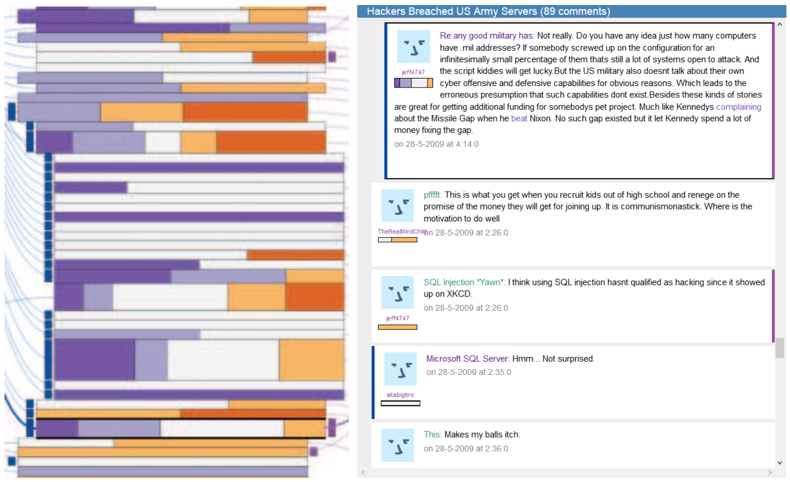

Figure 2.3: An example showing: (a) The user clicked on a comment (the one with black horizontal outlines) in the Thread Overview. (b) As a result, the system automatic scrolled to the actual comment in the Conversation View.

We conducted the study with five participants (age range 18 to 24, 2 female), who are frequent blog readers (four of them reported to read blogs at least several times a day and one reported several times a week). We asked the participant to explore the conversations according to her own interests and write down a summary of the key insights gained while exploring each conversation. During the study, we primarily focused on gathering qualitative data such as observations, user-generated summaries, and semi-structured interviews. In addition, we logged interface actions to better understand the usage patterns of ConVis.

The participants' feedback from our informal evaluation suggests that ConVis can help the user to identify the topics and opinions expressed in the conversation; supporting the user in finding comments of interest, even if they are buried near the end of the thread. We also analyzed the user generated summaries from the evaluation to reflect on the task abstraction. After mapping each sentence of the summaries to one or more possible tasks in Table 2.1, we found that some of the tasks were performed more frequently than others. For instance, topic related questions were used more frequently than other variables (e.g., authors). Finally, users found that sometimes topics and sentiment do not match the current information (they were too general or too specific).

2.1.5 Contributions

The primary contributions of this work are:

1) We performed a user requirements analysis based on extensive literature review in the domain of blogs. The analysis includes data and task abstractions for the problem domain and a set of design principles to support the user requirements.

2) To the best of our knowledge, ConVis is the first visual text analytic system for blog conversations that visualizes both *topic* and *opinion* mining results along with a set of metadata (e.g., authors, position of the comments), which were identified as primary means for browsing and navigation from the user requirement analysis. Previous systems either visualize some metadata or only one type of content information from the conversations (e.g., the topics covered but not opinions), thus limiting the ability of the user to explore and analyze the conversation.

3) We present the design, implementation, and an informal evaluation of ConVis. ConVis visually represents the overview of a blog and then allows the user to explore this conversation based on multiple facets (e.g., topics and authors). This is a major shift from traditional blog reading interfaces which provide a long list of paginated comments, thus only supporting a linear way of navigation.

2.2 ConVisIT: Interactive topic modeling

2.2.1 Problem

The preliminary evaluation of ConVis suggested that users could benefit from a greater control over the topic modeling process. This was particularly evident from the interviews and observational data, where users expressed a pressing need for enhancing their ability to revise the topic model according to their own information needs.

The need for revising the topics may arise due to three different reasons. First, sometimes the current information seeking tasks may require a topic model at a different level of granularity, e.g., if the user needs more specific information about ‘ObamaCare’ she might be interested in exploring its potential sub-topics such

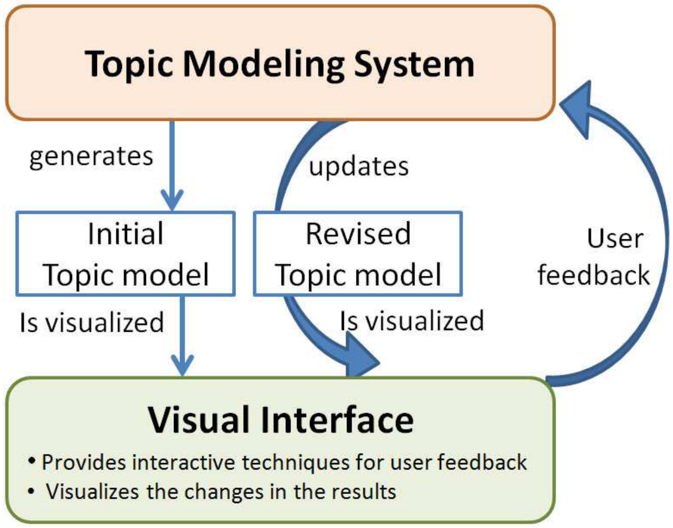

Figure 2.4: Interactive topic modeling framework for exploring asynchronous conversation.

as 'health insurance', 'healthcare cost' and 'drugs'. Second, the interpretation of topics may vary among users according to their expertise and mental model. In fact, in a topic annotation study humans sometimes disagreed on the number of topics and on the assignment of sentences to topic clusters [41]. For instance, for one of the conversations from their corpora, one annotator produced 22 topics, while another annotator reported only 8 topics. Finally, in some cases the results of topic modeling can be simply incorrect, in the sense that the generated topics would not make sense to any user [12, 41]. For example, two semantically different topics 'Obama health policy' and 'job recession' might be wrongly grouped together with the misleading topic 'Obama recession'.

To address the aforementioned problems, we propose a novel topic modeling system for asynchronous conversations that revises the model on the fly based on user's feedback. Figure 2.4 illustrates our proposed topic modeling framework. Given an asynchronous conversation (e.g., blog), the system generates the initial set of topics, which are presented in the visual interface along with other conversational data. The interface then supports the user in exploring the conversation. However, whenever the user realizes that the current topic model is not helping her, she can provide topic revision feedback to the system through interactions. Subsequently, the system updates the topic model and the new results are presented in the interface.

2.2.2 Related work

Since system-generated topic models can be noisy, some recent work investigate how user supervision can be introduced to improve the results. These works mainly focused on answering two research questions: 1) How to incorporate user feedback in the topic model? 2) How an interface can support the user in expressing such feedback? To answer the first question, in the dominant LDA topic modeling framework the original unsupervised LDA method was modified to allow the introduction of human supervision [2, 39, 59]. For instance, Andrzejewski et al. incorporate user's domain knowledge in LDA by adding constraints in the form of must-link (enforces that sets of words must appear together in the same topic) and cannot-link (enforces that sets of words must be in different topics) using Dirichlet forest prior [2]. However, this method requires to rerun Gibbs sampling from scratch after a set of constraints is added, leading to high latency. Since, such latency is undesirable for real-time interactions, [39] proposes an efficient inference mechanism that aims to minimize user's waiting time. Unfortunately, all these approaches were designed for generic documents. In contrast, we devise a new interactive topic modeling framework that is designed to take advantage of conversational features.

The question of how a visual interface can support the user in expressing her feedback has been addressed in [9, 13, 47]. Chuang et al. extend Termite [10], which visualizes the term-topic distributions produced by LDA topic models, and allows the user to revise the model by clicking on words to promote or demote a words usage in a topic [13]. Similarly, [47] visualizes topic modeling results from LDA, and allows the user to interactively manipulate the topical keyword weights and to merge/split topic clusters. More recently, user feedback was incorporated through a scatter plot visualization, that steers a semi-supervised non-negative matrix factorization (NMF) method for topic modeling [9]. The authors show that the NMF-based approach has faster empirical convergence and offers more consistency in the results over traditional LDA-based approach. They visually present each topic cluster and then allows the user to directly manipulate the documents and keywords within each cluster to specify topic revisions. A fundamental limitation of all these works is that the visual interfaces for interactive topic model was not evaluated with real users. Therefore, a set of critical research questions remained unanswered. For instance, would users be really interested in performing all the operations provided with such a complex interactive visualizations? What operations are actually useful to the users for performing exploratory tasks in a specific domain? To answer these questions, we applied a systematic approach, where we first devised a set of topic revision operations which are most useful according to our tasks analysis, and then performed a user study to measure the utility of these operations.

2.2.3 Methodology

Since it may take some effort from the users to express different topic revision operations, it is important to devise the minimal set of operations that would be both intuitive and sufficient to support user's tasks [13]. For this purpose, we first identified eleven different possible topic revision operations (see Table 2.1) based on reviewing existing work on interactive topic model [2, 9, 39, 47]. Next, we prioritized the operations based on the following criteria ordered by their importance: 1) *Task relevancy:* To what extent this operation is relevant to the tasks involved in exploring conversation as identified in [36]? 2) *Topic model relevancy:* Is this operation applicable to our topic model approach? 3) *Redundancy:* Is this operation already covered by other operations, which are stronger on the previous two criteria?

The three operations at the bottom of Table 2.1 (9-11) are eliminated based on both task and topic model relevancy criteria. Not only these operations are designed to fix the term-topic distribution which is not applicable to our topic modeling approach; but more importantly they are arguably not very useful to support the high level exploratory reading tasks as identified in [36] and therefore the users may not be motivated to perform such operations. On the contrary, we selected the top three operations in Table 2.2 (i.e., 'split a topic', 'merge topics by join', and 'merge topics by absorption'), because we identified them as the most rele-

No	Operation	Why?	Criteria			Reference
			Task relevancy	**Topic Model relevancy**	**Redundancy**	
1	Split a topic	This topic is too generic	high	yes	no	[9, 47]
2	Merge by joining	These topics are talking about similar things	high	yes	no	[9, 47]
3	Merge by absorption	A group of sentences are wrongly clustered into a different topic	high	yes	no	[9]
4	Split by keyword	This keyword should be separated into a new topic	medium	yes	yes	[9]
5	Change the overall granularity level of topics	Too few topics/ too many specific topics are generated	medium	yes	yes	-
6	Remove the topic from the display	This topic does not make any sense (i.e., off-topic)	low	yes	yes	[47]
7	Assign a label for this topic	The current label of this topic does not represent the actual topic	low	yes	yes	[28]
8	Increase the weight of this keyphrase	This keyphrase should be included in the topic label list	low	yes	yes	[28]
9	Apply must-link constraint	Those words **must be** in the same topic	low	no	no	[2, 39]
10	Apply cannot-link constraint	Those words **must not be** in the same topic	low	no	no	[2, 39]
11	Change keyword weights	This keyword is more related to the topic	low	no	yes	[9, 47]

Table 2.2: Different possible topic revision operations.

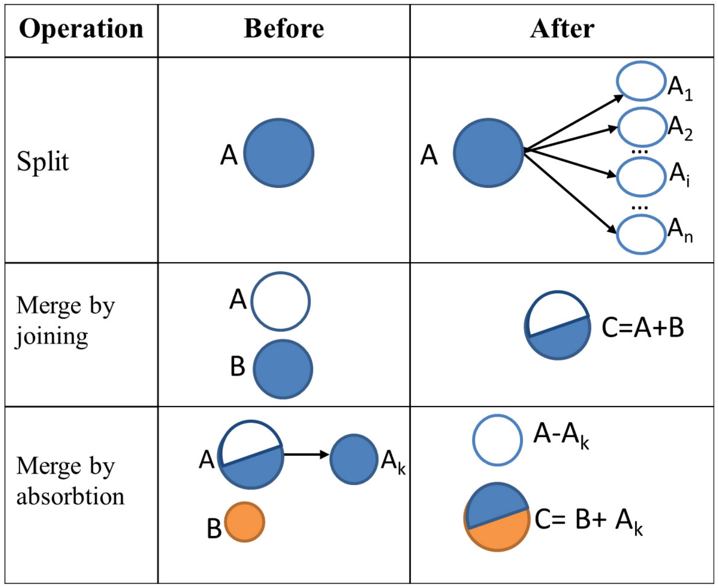

Figure 2.5: Three different user operations for topic revision

vant to our exploratory reading tasks that require the user to dynamically change the granularity level of different topics. Also, by selecting them some other candidate operations with lower task relevancy become redundant and therefore they are eliminated, such as 'change the overall granularity level of topics' (covered by topic splitting and merging) and 'split by keyword' (covered by topic splitting). In the reminder of the section, we describe how each of these operations support user's tasks, and how the underlying topic model is revised according to these operations.

Split a topic: Topic splitting allows the user to explore more specific sub-topics of a given topic, thus changing the topic granularity to a finer level. Consider an example, where initially the system creates a topic named 'military security'. As the user starts exploring this topic, she finds it to be too generic with respect to her information needs, and therefore she wants to split it into more specific sub-topics.

Method: Assume that the user wants to split a topic A into multiple sub-topics (see Figure 2.5). Upon user's request, the underlying topic model creates a subgraph $G_A(V_A, E_A) \subset G(V,E)$ from the original graph $G(V,E)$ generated in the initial topic segmentation, where V_A represents the vertices (sentences) of topic A, and and each edge $w(x,y)$ in E_A represents the edge weights of topic A.

Next, the system splits the chosen topical cluster A into further n sub-clusters $A_1, A_2, ..., A_n$, by applying the same graph partitioning algorithm used in the initial topic segmentation phase, i.e., approximate solution to n-Cut [62] on $G_A(V_A, E_A)$. Here, n is the optimal number of sub-topics, which is automatically determined by finding the value of n for which an objective function Q is maximized according to the formula proposed by Newman and Girvan [56],

$$Q_n = \sum_{c=1}^{n} \frac{\sum_{x\in V_c, y\in V_c} w(x,y)}{\sum_{x\in V_A, y\in V_A} w(x,y)} - \Big(\frac{\sum_{x\in V_c, y\in V_A} w(x,y)}{\sum_{x\in V_A, y\in V_A} w(x,y)}\Big)^2 \tag{2.1}$$

Here, Q_n measures the quality of a clustering of nodes in the graph $G_A(V_A, E_A)$ into n groups, where $\sum_{x\in V_c, y\in V_c} w(x,y)$ measures the within-cluster sum of weights, $\sum_{x\in V_A, y\in V_A} w(x,y)$ measures the sum of all edge weights in the graph, and $\sum_{x\in V_c, y\in V_A} w(x,y)$ measures the sum of weights over all edges attached to nodes in cluster c. In essence, according to (2.1), the nodes in high quality clusters should have much stronger connections among themselves than with other nodes in the graph.

We apply equation (2.1) for the value of $n = 2, 3, 4, 5$ and select the value of n, for which Q_n is maximum. The highest possible value is capped to 5 because of time constraint imposed by the interactive nature of the operation. Notice however, that this limitation is not too penalizing. Our analysis of the Slashdot corpus shows that in 86% cases of splitting a topic, the best value of Q_n is with $n \leq 5$.

Once the parent topic is segmented into n different sub-clusters, representative keyphrases are generated for each sub-topic. This is done by running our topic labeling method only on the sub-conversation covered by A.

Merge by joining: This operation allows the user to aggregate multiple similar topics into a single one. As opposed to topic splitting, the result is a topic with coarser granularity. Consider an example, where the initial topic model produces two different topics namely 'secure code' and 'simple sql server injection'. The user may find that both topics are too specific, therefore joining them into a more generic topic may help her to better perform the subsequent tasks.

Method: Assume that the user decides to merge by joining two topics A and B (see Figure 2.5). To perform this operation, the topic modeling system creates another topic C and assigns its vertices as $V_C = V_A \bigcup V_B$ and edges as $E_C = E_A \bigcup E_B$. After that, a label for C is generated. This is done by running our topic labeling method only on the sub conversation covered by C.

Merge by absorption: If a sub-topic is more related to a different topic than its current parent topic, merge by absorption allows the user to separate this sub-topic from its current parent and merge it with the one to which it is more related. Unlike the previous merge operation (which joins two independent topics), this operation allows a sub-topic that is already placed under a topic to be absorbed by a different parent topic. Consider an example, where the sentences related to two different topics, namely 'Obama health policy' and 'job recession' are wrongly grouped together under the topic 'Obama recession'. The user may realize that the sub-topic 'job recession' should be separated from its parent topic and merged with the 'unemployment' topic to which it is more related.

Method: Upon receiving a merge by absorption feedback from the user on A_k and B, the topic modeling system removes the sub-topic A_k from its current parent A and merge it with the topic B (see Figure 2.5). The system then creates a new parent topic C and then assigns vertices such that,

$$V_C = V_{A_k} \cup V_B, V_A = V_A \setminus V_{A_k} \tag{2.2}$$

and edges such that,

$$E_C = E_{A_k} \cup E_B, E_A = E_A \setminus E_{A_k} \tag{2.3}$$

After that, the topic labeling method takes the portion of the conversation that consists of the sentences in V_C, thus generating a label for C that potentially represents descriptive keyphrases from both topics A_k and B.

Interactive visualization for topic revisions: We have designed a set of interactive topic revision operations within the interface through some intuitive direct manipulation methods. As the user performs these operations, the system updates the topic model and changes the visual encoding of the topic list from the initial flat list of topics into a multi-rooted tree organization. Such updates to the topic organization becomes visible to the user through perceptually meaningful animations, following the design guidelines of effective animation presented in [35]. In particular, we have designed staged animation for each operation, i.e., we break up the corresponding transition into a set of simple sub-transitions, allowing multiple changes to be easily observed.

For instance, when the user splits a topic by double clicking on it, the following

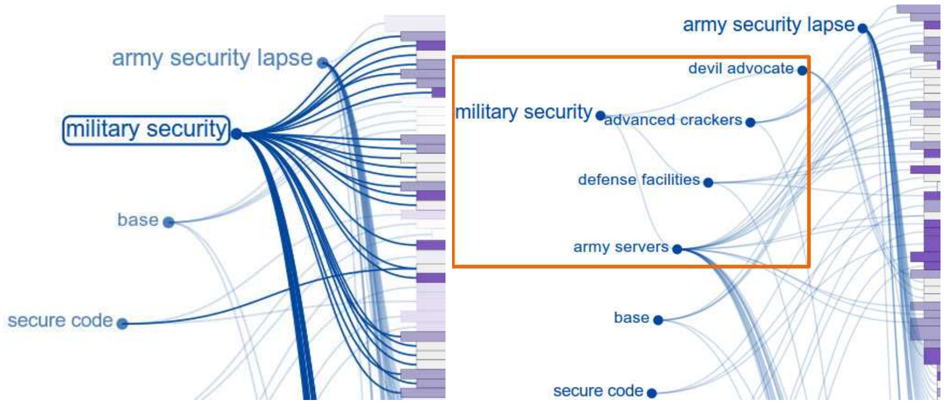

Figure 2.6: An example showing: (a) The user hovers over the topic 'military security' and decides to perform the split operation. (b) As a result, the topic moves to its left while the rest of the topics are pushed along the perimeter of the circular layout to create space for the new children.

sub-transitions occur. First, the clicked topic A moves to the left along with its parent node(s) (if any), while existing nodes at the deepest level are pushed towards their new positions (up/down) around the circular layout to create angular spaces for the new sub-topics. Second, the new sub-topics $A_1, A_2, ...A_n$ appear and move from their parent's position (A) to their new positions. Third, labels appear for these sub-topics (see Figure 2.6). Double clicking on A again causes it to collapse by following the exact reverse order of animation, i.e., the labels of the children move from their current positions to their parent and fade away, and then the parent moves to its previous position while other nodes move closer to the parent node to fill the gaps left by the removed children nodes.

Merging of two topics can be performed by dragging a topic A over another topic B, which causes the system to update the topic model. As a result, a new parent topic C appears to the left and curved links are drawn from C to A and B to indicate parent-child relationship (see Figure 2.7). The user can subsequently double click on C to collapse it, which hides its sub-topics. Finally, if a child topic A_k is discovered to be wrongly placed under a topic A instead of under a more appropriate topic B, the user can drag A_k over B. As a result, the link of A_k with its parent A is removed and then a new parent node C appears that connects both A_k and B (see Figure 2.8).

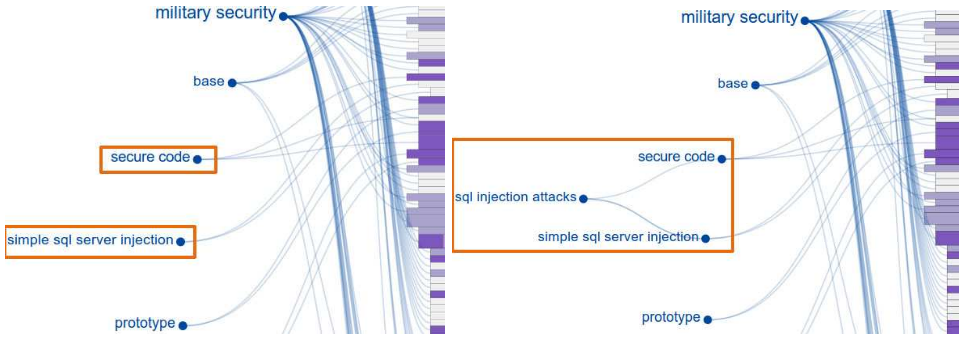

Figure 2.7: An example showing: (a) The user decides to merge two topics by joining (indicated by orange color). (b) As a result, ConVisIT updates the topic organization where these two topic nodes are merged under the parent topic 'subject sql injection attack'.

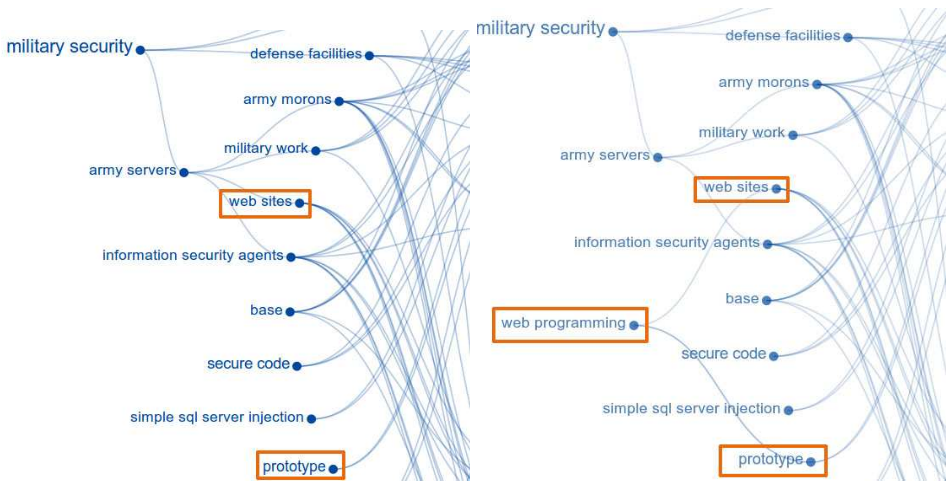

Figure 2.8: An example showing: (a) The user decides to perform merge by absorption on two topics 'web sites' and 'prototype'. (b) ConVisIT updates the topic organization where the previous link from 'web sites' to 'army server' is removed, and then 'web sites' is absorbed into a more generic parent topic 'web programming' along with 'prototype'.

2.2.4 Results

We run a user study in controlled settings to understand how the introduction of visual interfaces for exploring conversations may influence the user performance and subjective opinions compared to more traditional interfaces. 20 subjects (aged 19-43, 8 females) with considerable prior experience of reading blogs participated in the study. The study was designed with three interfaces as conditions: a) *Slashdot*, b) *ConVis*, and c) *ConVisIT*. The Slashdot interface follows a typical blog reading interface design and it serves as a suitable baseline for our experiments. It provides an indented list based representation of the whole conversation as well as common functionalities of blog interfaces such as scrolling up and down, collapsing a sub-thread, and searching for terms. The primary reason for comparing between ConVis and ConVisIT was to verify whether any potential influence in performance and user behaviour are due to the visualization features common between them, or due to the interactive topic revision feature (which is only present in ConVisIT).

In the study, our first research question was whether there is any difference in user performance and subjective reactions due to the interface condition. We found that overall ConVisIT was the most preferred interface, and was rated higher over its counterparts on the *findInsightfulComments* measure. On the contrary, Slashdot was the least preferred interface, and it received significantly lower rating on three different measures. As for ConVis, it seems to provide a middle ground between the other two interfaces and its topic organization, although static, was found to be visually less cluttered than the one of ConVisIT. In general, this shows that while interactive topic model can be beneficial to the user, such feature may introduce visual clutter and interaction costs at least for some users. Finally, there were no significant differences among the interfaces in terms of *easeOfuse*, in spite of the higher complexity of ConVis and ConVisIT.

The second key research question was what specific visualization features/ components of the interfaces are perceived as more/less beneficial by the potential users (e.g., interactive topic revision, Thread Overview, and relations between facets)? We found that in general, the visualization features of ConVis and ConVisIT received higher rating than the ones of Slashdot. Interestingly, we found that subjective reactions about different features of the interfaces such as split, merge, and clicking on topic directly correlates with their frequency of use. More importantly, we found that not all interactive topic revision operations were equally received. For example, the split operation was used more frequently than its counterparts. This issue needs to be further investigated.

2.2.5 Contributions

The primary contributions of this work are three-fold:

1) A novel **interactive topic modeling system** specifically designed for asynchronous conversation. Existing systems (e.g., [9, 39, 47]) are mainly devised for generic documents without considering the unique features of conversations. On the contrary, we analyze the information seeking tasks in our target domain to select a minimal set of topic revision operations that are critical to the user. Then, we devise computational methods for each of these operations to be performed by the system.

2) **ConVisIT, a visual interface** which provides a set of interactive features that allow the user to revise the current topic model. In response, the interface updates and re-organizes the modified topics by means of intuitive animations, so that the user can better fulfill her information needs.

3) **A user study** to understand how the visual interfaces for exploring conversations (ConVis and ConVisIT) may influence user performance and subjective opinions when compared to more traditional interfaces. This evaluation also provides insights into the potential advantages and relative trade-offs of interactive topic visualization approaches (i.e., ConVis vs. ConVisIT).

Chapter 3

Ongoing and Future Work

Recently, I have started working on a project where the goal is to support readers in exploring a collection of social media conversations related to a given query. An overview of this planned work is provided in Section 3.1 and 3.2. Moreover, I discuss my plan to summarize the reflections and findings from the completed and ongoing studies into a generalizable design framework for visual text analytics in Section 3.3. A summary of the on-going work is provided in Figure 3.1.

3.1 Exploring Social media collections

3.1.1 Problem

Now-a-days, social media conversations provide abundance of variety of opinions about a user-provided query, such as ‘ObamaCare’, ‘US immigration reform’, ‘iWatch release’. Often readers are interested in the set of conversations about their query of interest to perform some information and opinion seeking tasks [34, 43]. However, traditional blog sites provide only limited functionalities to get an overview of a set of conversations. Generally, they present latest posts and allow the user to search using keywords terms. If the user provide a query, they present a set of blogs as a paginated list, ordered by their recency. Therefore, users are required to rely on their own to go through this list and read the individual conversations in details to understand what people are saying about the topic of their

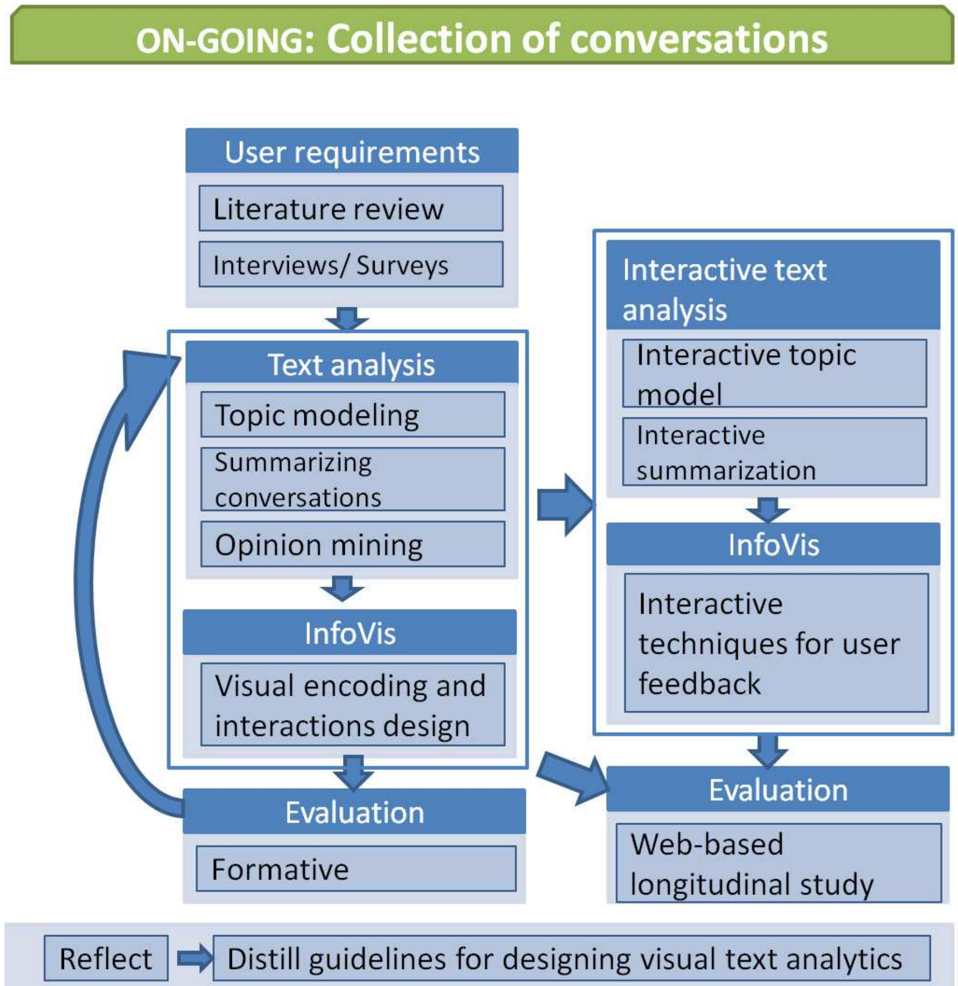

Figure 3.1: An overview of the on-going work.

interest. This navigational support is often inadequate to explore a set of blogs that sparked lively discussion and are of interest to the new reader [67].

To illustrate the scenario, consider a casual blog reader John, who is interested to buy a new smartphone. He heard that Apple recently launched the new iPhone6. However, some people claim that this new phone can easily bend in the pocket while sitting on it. John is really eager to know what are people's reaction about the problem, as he wants to make up his mind about buying an iPhone6. In particular, he wants to know: what people had to say about this bending issue? To what extent people think that the bending issue is serious and why? John is looking for interesting conversations that discussed this issue in Macrumors [1], a blog site about Apple related news. By searching in this site with a query ‘iPhone bend’, John finds a dozen of conversations covering related issues such as ‘what users reported about the bending issue’, and ‘what Apple says to defend its new product’. John needs to get an overview of the major points and opinions, but going through all of these conversations, that overall consist of more than a few thousand comments, in a reasonable amount of time is impractical. Therefore, after going through some of the initial comments, John decides to stop reading the conversations without fulfilling his needs.

3.1.2 Methodology

From the above problem scenario we can summarize the following main limitations of current interfaces that hinder the user’s ability to perform their information and opinion seeking tasks: the lack of high-level abstraction of the conversations, i.e., the user needs to read the detailed conversations to get an overview; the lack of presentation beyond a long paginated list; and the lack of navigation cues for exploring interesting contents. Arguably, in order to support this task we need to combine NLP and InfoVis techniques in a synergistic way as we have done for a single conversation. While NLP techniques could provide a high-level abstraction of the content of the conversations, InfoVis techniques could improve the presentation and interaction techniques to support the tasks more effectively.

To tightly couple NLP with InfoVis techniques, I plan to apply a similar design approach taken in our completed work in Chapter 2, where I followed the nested model for visualization design [54]. The rest of this section describes how we will approach the problem in details.

1) User requirement analysis: In order to understand why and how people explore a collection of conversations, we are reviewing relevant literature from the fields of computer mediated communications (CMC) [22, 43, 74], social media [30, 34], human computer interaction (HCI) [53], and information retrieval [40,

[1] http://macrumors.com/

44, 49, 52, 63].

Since we are still focusing on conversational data as we did in our completed work in Chapter 2, many of the user requirements described in Section 2.1 are also relevant here. For example, we previously mentioned that information and opinion seeking are the major two reasons for people to explore blogs [18, 43, 44, 52]. Similarly, some reading behaviours are also common (and arguably more frequent) in exploring a blog collection, such as skimming behaviour i.e., the user frequently switches between exploratory and focus state [55, 74].

However, there are also significant differences in why and how the readers explore and analyze a collection of conversations compared to a single conversation. It has been argued that the temporal evolution of discussion themes is a key characteristic of a collection of social media conversations [19, 34], as opposed to a single conversation where the positions of comments or topics are more important than their timestamps. For instance, to explore a popular topic X from a collection of blogs, the user often wants to perform the following task: " Find out what are people thinking or feeling about X over time." [34].

While reviewing the literature is a reasonable starting point for my study, it alone may not be sufficient to fully uncover the user requirements. Therefore, to complement the literature review, I will consider conducting interviews and online surveys with a set of target users as part of analyzing user requirements. While interviews can provide more qualitative information gathered from a small number of participants, online questionnaires would allow us to reach a wide range of participants [58]. The goal of the study would be to collect the following information about the participants: 1) their background, 2) how they currently explore and read a set of conversations about a popular issue and what problems they usually face, 3) what features they would expect from a visualization designed to support their tasks. The analysis of the study results may help us to understand how we can synthesize and abstract the raw conversational data, so that the resulting data abstraction can better support the users.

Dataset: In order to guide the data abstraction, I have been collecting different sets of conversational data and performing exploratory analysis on these datasets. The goal here is to understand the types and the scales of the data that the visualization may need to deal with. Since blog readers are generally more interested in exploring recent issues, existing blog corpora (such as the corpora at the TREC conference [49]) are not suitable for our experiments, because they are fairly old. Therefore, we have decided to create new datasets for our design and experiments. So far, I have focused on Macrumors, which is a large online forum dedicated to the discussion of recent news and opinion relating to the Apple Inc.

Dataset	#Conversations	#Comments	#Authors	Average Duration per conversation
iPhone bend	13	4461	1736	11.63 days
iWatch battery	29	9400	2720	7.85 days
IPad air release	29	3283	1443	5.45 days

Table 3.1: MacRumors blog dataset statistics

In order to create new dataset from MacRumors, I developed a crawler to extract conversations for a set of queries. A summary of the resulting dataset is provided in Table 3.1.

2) Related Work: Some earlier works have focused on how to support the exploration of a blog archive using only metadata, for example, by visualizing tags and comments arranged along a time-axis [67], or by providing faceted visualization widgets for visual query formulation according to time, place, and tags [23]. While these works may assist users to find the blogs they are looking for, they are not designed to support users in understanding the actual content of these blogs. On the contrary, our goal is to support the tasks that require the user to get overviews of the actual contents of a collection of conversations, such as " Find out what are people feeling about *X* over time."

In contrast to the metadata of the conversations, some visualizations attempt to represent the actual topics being discussed within a collection of conversations [24, 68, 71]. For example, Themail visualizes how topics in personal email conversations develop over time by arranging keyword lists along a horizontal time axis [68]. TIARA represents the temporal evolution of topics from an email collection using stacked graph representation [71]. Using the same metaphor, Visual Backchannel [24] presents keyword-based topics that are extracted from the ongoing conversations in twitter, as well as a set of metadata i.e., participants and photos in a multiple coordinated view interface. Inspired by these works, we also plan to consider topical, temporal and social aspects of the conversations into our visualization, as we have identified them as crucial facets in our preliminary user requirements analysis.

More recently, topics from a news corpus were organized into a hierarchy based on computing the distance between the probability distributions of topics [26]. Using the same algorithm, topic hierarchies were built from multiple corpora (i.e., news, blogs, and microblogs), followed by matching these hierarchies using a graph-matching technique, so that the common and distinctive topics from different corpora can be visualized [48]. Organizing topics into a hierarchy can be useful to our work as well, because the number of distinctive topics in a collection of conversations may be quite high, compared to a single conversation.

There is a growing interest in visualizing the subjective aspects of online conversations, mostly focusing on microblogs [20, 50, 73]. Diakopoulos et al. presented Vox Civitas [20] that displayed sentiment and tweets volume over time for events discussed in microblogs to support the tasks of journalistic inquiry. TwitInfo [50] was also designed for visualizing microblogs with a focus on providing more accurate aggregation of sentiment information over a collection of tweets. Unlike these works, OpinionFlow focused more on visualizing the diffusion of opinions about a particular topic (e.g., 'US government shutdown') among participants with a combination of a density map and a Sankey diagram [73]. Contrary to these works, which only focused on how to show sentiment information, I aim to show a higher level opinion information, e.g., a summary of the major arguments about topic X. This may allow us to support a broader range of opinion seeking tasks, such as "Why people are supporting or opposing an opinion?".

3) Text analysis: Informed by the user requirements analysis, we would like to devise a set of NLP techniques to extract useful information from collections of conversations. Based on our preliminary user requirement analysis, we are considering the following possible text analysis methods to be applied to our datasets.

- **Topic modeling:** The primary purpose of applying topic modeling technique is to create a high-level overview of the whole collection of conversations. In ConVis [36], we used a topic modeling method that takes advantage of conversational structure (e.g., the use of quotation) to improve accuracy over traditional topic model such as LDA. While in general this method has been found to perform well for a single conversation, we may need to extend it to create a topic model over a large set of conversations. Currently, I am considering two possible ways of extending the current method: 1) Bottom-up approach: We can apply topic segmentation and labeling method to each conversation (local topic model) and then aggregate these local topic models into a global hierarchy of topics based on similarly measures between topic segments, such as mutual information [65]; 2) Top-down approach: We can create a topic tree over the whole set of conversations by iteratively splitting topics into their sub-topics, for instance, by using the hierarchical phrase mining method recently proposed in [69].

- **Opinion mining:** According to our preliminary user requirements analysis, the user often wants to perform opinion seeking tasks, such as 'why people were supporting or opposing an opinion?'. Unfortunately, most of the existing visualizations (including ConVis [36]) present sentiment information based on simple lexicon-based sentiment classification techniques that may not effectively support the opinion seeking tasks.

  We realize that opinion seeking tasks could be better supported by developing more advanced text analysis techniques to detect disagreement in conversations. A recent work from our group shows that incorporating higher level discourse relations obtained from the conversations improves the accuracy in detecting disagreements [1]. Therefore, I plan to apply and extend this work to detect disagreements on topics, and possibly extract the utterances from conversations that expressed major disagreements within a collection of conversations. While such new techniques could benefit the opinion seeking tasks for a single conversation, they could be even more useful for understanding opinions expressed within a collection of conversations, because the number of comments in a collection might drastically increase.

- **Text summarization:** Generally, each topic in a topic model is represented by a set of words or a keyphrase, which can be regarded as a short summary of that topic. However, if the reader is interested to know what people are actually saying about a particular topic, she may need to read through the related comments. For example, if the reader is interested about a topic 'military security lapses', only after reading the detailed comments she may understand that the participants were commenting on what were the major security lapses behind the hacking of US Army Server. While it may be possible for the reader to quickly scan through some comments of a single conversation, when she needs to deal with a topic that was discussed over a set of conversations it becomes much more difficult and time consuming. This could be due to two reasons: 1) the number of related comments could be much higher, 2) the discussions about a topic may scattered around multiple conversations, making it difficult to drill down to different conversations for understanding the discussions about the topic.

A promising solution to this problem could be to provide a multi-sentential summary for each topic, that captures the most salient information. First, it may allow the user to quickly understand what discussion was primarily covered by a topic. Second, from a summary sentence the user may able to drill down to the conversation where that sentence appears to know more details on the topic of interest.

To summarize topics, I aim to devise and extend graph based approaches [29], where I plan to consider the conversational features such as rhetorical structure of the posts and fragment quotation graph along with sentiment analysis in an unified way. Our expectation is that by taking these features into account the quality of the summaries may potentially improve and better support the users.

In order to devise and implement various text analysis methods, I plan to collaborate with Shafiq Joty, a former member of our NLP group who is currently a Scientist at Qatar Computing Research Institute (QCRI). I plan to pursue a research internship at QCRI, where in collaboration with the Language Technology and Social Computing groups, we will develop new NLP methods for social media conversations.

4) Design and implementation: Once we complete the task and data abstractions, we will map them to a set of appropriate visual encoding and interaction techniques. Based on our preliminary analysis, I am considering a broad range of visualization design choices. For example, it has been found that providing *navigation cues* through various metadata can be useful for exploring conversations [67]. Therefore, it may be useful to design a set of scented widgets, as they have been shown to improve social data analysis by providing navigational cues [72]. We have also mentioned the exploratory information-seeking behaviour of blog reading i.e., tendency to switch between different sub-topics, which may become even more frequent in exploring multiple conversations. I consider supporting this through novel focus+context techniques such as the monadic exploration [25]. Using monadic exploration technique the user can navigate through the current topic of interest with the closely related topics being attracted towards the center, while less relevant topics are still shown on the periphery of a circular layout, so that the user can switch to a different topic at any time. Finally, Hearst et al. suggested that the designers should support the exploration of a social media collection through various facets (e.g., time, topic, authors etc.) [34]. Therefore, we also aim to explore various design choices of faceted exploration [4, 27, 46].

During an exploratory session, a user may need to deal with three different scales of conversations: 1) *All*: Initially the user encounters the whole collection of conversations, some of which may not be relevant to her information needs; 2) *Some*: The user gradually filters the list to a smaller set of conversations that are potentially relevant to her information needs; and 3) *One*: The user decides to read a specific conversation. An important question is which scale among these should form the primary view and how to encode connections between these three different scales? Another question is when the user moves from one scale to the next, how we should map different facets such as topics and authors for each of these three views? We need to find the most suitable techniques for such mapping among various alternative visual encoding/ interaction techniques (e.g., linked highlighting, showing explicit links, appear in other view based on selection in one view).

While I am drawing some early sketches based on our current understanding of task and data abstractions, they will be iteratively modified through several phases of sketching, prototyping, and getting feedback from real users and InfoVis experts. One example of such early sketches is shown in Figure 3.2.

5) Evaluation: During the design, we plan to conduct formative evaluations to

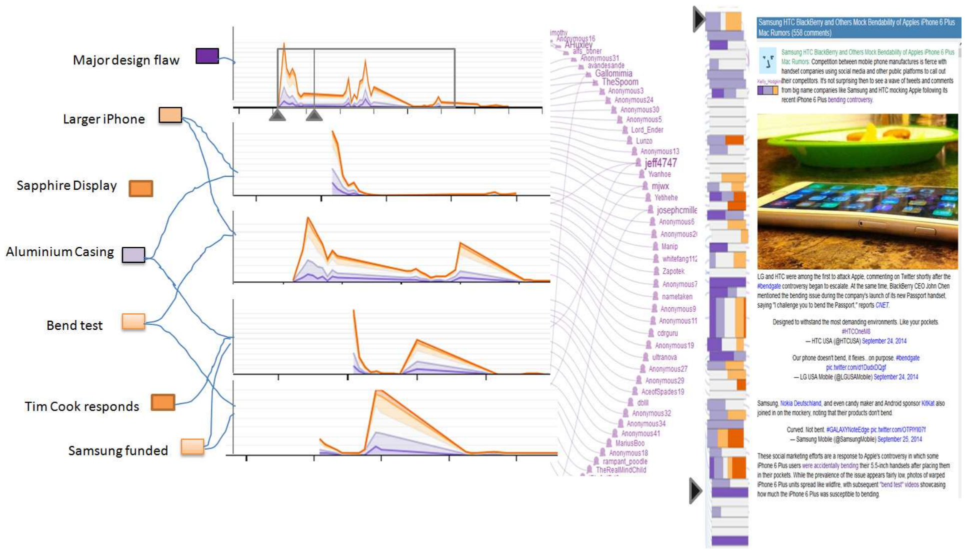

Figure 3.2: An initial sketch of a visual interface for exploring a collection of conversations. **Left:** The overview consists of topics and authors to support faceted exploration and for each conversation, a Timeline graph shows the volume of comments over time with sentiment distribution to provide visual information scent (middle-left) [72]. **Right:** The selected conversation appeared along with an overview of its thread structure.

identify potential usability issues and to iteratively refine the prototype. Once the high-fidelity prototype is developed, we aim to conduct informal evaluations and case studies [45] with a different set of target users to understand to what extent the overall visualization and its specific components are perceived to be useful by the potential users and to solicit ideas for improvements and enhancements. We are considering three possible groups of users for evaluation: 1) frequent blog readers, 2) social scientists 3) journalists.

6) Reflection: At the end of evaluating the high fidelity prototype, I aim to critically reflect on our design study. In particular, I would like to revisit the task abstractions and analyze to what extent these abstractions meet real user's goal. I also hope to investigate the possibility of expanding our task abstractions to include the tasks that may be relevant only for someone participating to an ongoing conversation. For instance, what is the nature of these tasks? Subsequently, to what extent the visualization needs changes for participants compared to non-participants?

3.1.3 Expected contributions

We anticipate that through the user-centered design process, the resultant visual text analytic system will provide a novel way of searching and exploring a social media collection. While previous work has mainly focused on visualizing metadata [24, 67] and results of simple text analysis techniques, our goal is to combine a set of metadata and the results of advanced text analysis techniques in a synergistic way. We expect that the resulting system will enhance the user's ability to fulfill information needs that are hard, if not impossible, to satisfy with traditional interfaces.

Moreover, the text analysis methods that we plan to develop to support the user tasks may lead to additional contributions, such as novel way for detecting the disagreements for controversial topics, and for summarizing conversations by combining a set of conversational features.

3.2 Interactive mining and summarization of conversations

3.2.1 Introduction

After designing and evaluating the first version of the system for exploring a collection of conversations, I plan to perform a similar analysis that was conducted for ConVis. In particular, I will examine to what extent the data extracted from the conversations was useful the real user. This analysis may reveal that: first, some of the data abstractions are simply wrong (i.e., the user does not care about those); second, the data extracted is appropriate, but they are not sufficiently accurate; finally, the data extracted do not match the expertise, mental model and tasks of the user. We anticipate that in the second and third scenarios, a promising direction would be to explore opportunities for the human-in-the-loop computation, as we have already found this idea to be useful for exploring a single conversation.

3.2.2 Methodology

Identify the potential scope for human-in-the-loop computation: We anticipate that from our preliminary evaluation, we will be able to identify where introducing human feedback in the text analysis systems may enhance the performance of the user in a significant way. Based on the text analysis methods that we discussed in Section 3.1.3, we anticipate the following possible scopes:

1) Interactive topic modeling: In Section 2.2, we have already showed that the user can benefit from the ability to interactively change the granularity of topics. The potential utility of interactive topic modeling might be even greater, when the topic model is extended over a large set of conversations. This could be due to the fact that as the volume of conversational data increases, so does the chances that the granularity of the topics will not match the expertise, mental model, and current task of the user. Therefore, if the results of the preliminary evaluations will indicate this to be the case, we will examine whether the introduction of interactive topic modeling has potential value, and what could be useful topic revision operations in this context.

In Section 3.1, I have discussed the plan for extending the topic modeling method that is currently designed for single conversation to run over a collection of conversations. We may modify this new topic modeling method by devising a set of interactive topic modeling operations, so that the user can provide interactive topic revision feedback.

2) Interactive summarization: While presenting a textual summary, a significant design challenge may arise due to the tradeoff between showing a long, informative summary and minimizing the screen space by showing a short snippet [33]. Depending on the current tasks, sometimes the user may need to read a long informative summary, in other cases a short snippet may be just sufficient for her. This is evident from the finding that adding more information to the summary significantly improved performance for information seeking tasks, but degraded performance for navigational tasks [17].

However, traditional summarization methods provide only a fixed length summary snippet and do not provide the user with control over the summary generation process. In this context, it might be useful to allow the user to modify the summary snippets dynamically based on different possible user feedback. For instance, the user may be allowed to dynamically change the overall summary length, as users have been found to benefit from customizing the summary length [42]. Another variation of this operation could be to expand summary with respect to only a portion of the current summary, e.g., 'give me more detailed summary related to only this summary sentence'.

Design: Here, our goal is to incorporate user's feedback within the visual text analytic system through a set of interactive techniques. Once the system revises the text analysis results based on user feedback, the interface will update and reorganize the modified results by means of intuitive animations, so that the user can perceive the changes in the results. After deciding on the type of user feedback and modified results, we will explore various design choices for the interface.

An important question related to the design process is: how can we use low-fidelity prototyping techniques to simulate the interactive feedback loop, prior to designing high fidelity prototypes? Since implementing the interaction techniques and the generation of system's responses require significant efforts, how can we utilize a prototyping method that closely resembles the real interactions.

Evaluation: In order to enhance the ecological validity of the evaluation [8], we would like to perform a longitudinal study by running the new system on a Web server and then invite participants to carry our their tasks online using our systems. The aim of the study would be: 1) to understand the potential utility of visual interfaces for exploring a collection of conversations; 2) to study the behavioral aspects of human-in-the loop topic, for instance to what extent real participants care about providing their feedback to the system and how that potentially enhance their performance.

3.2.3 Expected Contributions

We anticipate two primary contributions from this work: 1) Devising a human-in the-loop computation, by identifying the potential limitations of the current text analysis methods through a human-centered design approach; 2) An ecologically valid evaluation of the system among large set of users. Most of the existing systems that provide human-in-the-loop computation [9, 13, 47] have not been systematically tested to verify whether real users actually care about all the interactive operations that aim to improve the text mining and analysis results.

3.3 Design framework for visual text analytics

3.3.1 Introduction

After completing the design studies, I will take a step back to reflect on the wider context of designing visual text analytics. In particular, I aim to critically analyze the role of the designer in the process of integration between NLP and InfoVis techniques. During the final phase of my PhD, I plan to summarize these reflections and findings into a generalizable design framework for visual text analytics.

3.3.2 Methodology

Background: Within the visualization community, there has been significant advancement in the field of design study methodologies [51, 54, 61], which provides guidelines on how to perform design activities and how to validate different stages of design. However, when designing a visual text analytic system, it is also required to devise a set of *text analysis methods* and validate the interpretability, accuracy, and usefulness of the output generated by these methods. Arguably, devising suitable text analysis methods is just as critical as visualization design in determining the effectiveness of a visual text analytic system. Therefore, within a user-centered design approach, a designer must consider what should be the most suitable text analysis methods, and how to iteratively modify these methods when the output of these methods are not sufficiently interpretable, accurate, and useful.

Unfortunately, when designing visual text analytic systems, many researchers treat text analysis models as black boxes without considering whether they are the most suitable models for a particular problem domain. For example, many text visualizations select the terms to be displayed based on their frequency [24] or their TF-IDF scores [68], even though more sophisticated techniques are available [31] that could select better descriptive keyphrases. By not considering the most suitable text analysis methods, often the system fails to effectively support the real world tasks.

Contrary to this trend, Chuang et al. opened the black-box and focused on how to devise an interpretable and trustworthy text analysis model by aligning the model with the tasks and user expertise in a particular problem domain [11]. From their experience of designing a dissertation browser, they distilled the following process-oriented guidelines: *align* the model with the tasks, user expertise and visual encoding; *verify* the model to assess how well they fit an analyst's goals; iteratively *modify* the model when the model's output is incorrect or incomplete; and *progressively disclose* data at multiple levels of data abstractions, so that analysts can switch between different levels of abstractions to interpret and verify model's output. Within their guidelines, they discussed *what* are the possible approaches to improve a candidate model (e.g., modify model parameters, modify the model structure, add more training data, and leverage interactive machine learning techniques). However, no adequate guidelines were provided on *when* and *how* to choose a particular model modification approach. For instance, even though they acknowledge that modifying the model by introducing interactive machine learning techniques is a challenging problem, no further guidelines were provided on when and how the designer should devise such techniques, so that the resulting system improves significantly.

Our design approach: Echoing the call for aligning the model with the tasks and visual encoding [11], in our completed work we focused on the problem domain characterization and task abstractions first and then devised the suitable models. Based on our data and task abstractions, we also made modeling choices, i.e., choosing the suitable topic modeling and sentiment analysis methods.

During the preliminary evaluation of the initial prototype, we *verified* the performance of the topic models with end users. Through this evaluation, we identified that the current model sometimes does not match the user's mental model and current tasks. Therefore, we pondered *how* the model could be modified so that it can support users in performing their tasks more effectively. Since our analysis revealed that the perceived usefulness of the topic model depends on user's mental model and current tasks, introducing some user feedback to the model was deemed to be more promising than other alternative approaches (e.g., modifying the model parameters, and the model structure).

Once we had decided to introduce interactive topic modeling, we faced another

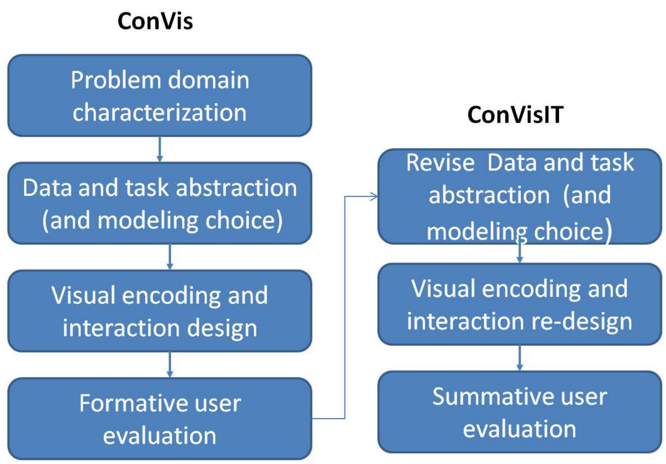

Figure 3.3: Design cycles of ConVis and ConVisIT.

important challenge of *how* to devise a minimal set of interactive topic revision operations, that the real user would care about. To approach this issue, we first identified a set of candidate operations and then we prioritized the operations based on three criteria, i.e., 1) task relevancy, 2) topic model relevancy, and 3) redundancy. These criteria helped us to tie the model modification process with the task abstractions and the current topic model. As a result, we were able design of a new interactive topic modeling method that better matches the user's goal. An overview of our design process is illustrated in Figure 3.3.

I would like to continue to analyze and understand *how* text analysis models can be chosen and modified within the design process. During the ongoing project, through a series of collaborations with NLP researchers, I plan to iteratively refine not only the visualization choices but also the text analysis models. In particular, I aim to understand when and how the designer should introduce human-in-the-loop computation, instead of improving the model without human supervision. If the designer decides to choose human-in-the-loop computation, how should she select the set of interactive operations for modifying the model? Through the experience of on-going design study I would like to critically analyze these questions. Finally, I aim to use the completed and ongoing design studies as walk-through examples to distill guidelines for different stages of design.

3.3.3 Expected contributions

Since, guidelines for designing visual text analytics are limited, our expected contribution is a novel set of lessons and design guidelines that could significantly benefit the visual text analytics community.

Chapter 4

Milestones

I began my PhD study in September 2012, and completed the comprehensive course requirements within the first two terms. During this time period, I also worked with the Intelligent User Interfaces group as a Research Assistant to build up my background knowledge. I co-authored two conference papers ([7] and [16]) and two workshop papers ([5]) and [15] in the process. After that, I positioned a journal paper in Computer Graphics Forum (Proc. EuroVis 2014) [36], presented a workshop paper at the ACL conference [38], and will present another paper at the ACM Intelligent User Interfaces conference (IUI 2015) [37], all of them focusing on my PhD dissertation.

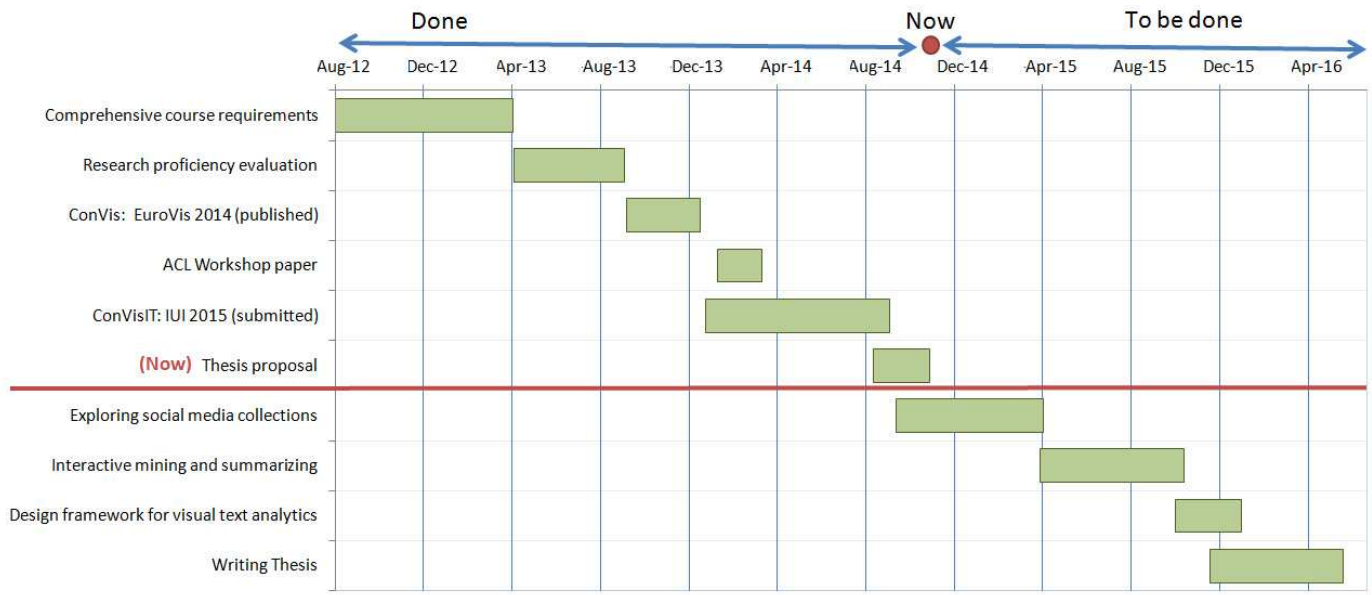

Figure 4.1: Timeline for my doctoral research.

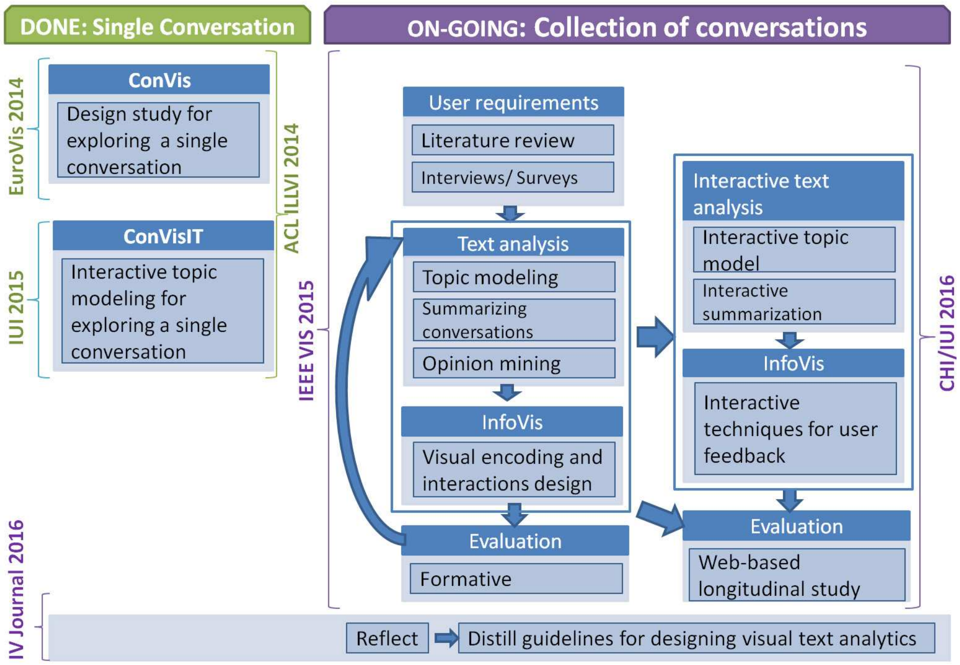

Figure 4.2: A summary of the completed and on-going work along with publication plan.

The remaining milestones in my thesis are: completing the proposed projects described in Chapter 3, as well as writing and defending my PhD dissertation. I aim to achieve these milestones with a target completion date in *Summer 2016*. Figure 4.1 shows an estimated timeline for my PhD research.

Publication plan: In the near future, I plan to continue working on exploring social media collection (Section 3.1) and submit a paper to *IEEE VAST/InfoVis 2015*. Subsequently, I will aim for positioning another paper on human-in-the-loop computation (Section 3.2) to *IUI 2016/CHI 2016*. Finally, one possibility to submit a paper on design framework for visual text analytics (Section 3.3) would be the *Information Visualization* Journal. A summary of the proposed work along with publication plan is presented in Figure 4.2.

Bibliography

[1] K. Allen, G. Carenini, and R. T. Ng. Detecting disagreement in conversations using pseudo-monologic rhetorical structure. In *Proceedings of the EMNLP*, 2014. → pages 33

[2] D. Andrzejewski, X. Zhu, and M. Craven. Incorporating domain knowledge into topic modeling via dirichlet forest priors. In *Proc. International Conf. on Machine Learning*, pages 25–32, 2009. → pages 17, 18, 19

[3] E. Baumer, M. Sueyoshi, and B. Tomlinson. Exploring the role of the reader in the activity of blogging. In *Proc. Conf. CHI*, pages 1111–1120, 2008. → pages 10

[4] N. Cao, D. Gotz, J. Sun, Y.-R. Lin, and H. Qu. Solarmap: Multifaceted visual analytics for topic exploration. In *IEEE Conf. on Data Mining (ICDM)*, pages 101–110, 2011. → pages 35

[5] G. Carenini, E. Hoque, and B. Steichen. User task adaptation in multimedia presentations. In *Proc. International Workshop on User-Adaptive Information Visualization (WUAV 2013), in conjunction with conf. on User Modeling, Adaptation and Personalization (UMAP 2013)*. → pages 44

[6] G. Carenini, G. Murray, and R. Ng. *Methods for Mining and Summarizing Text Conversations*. Morgan Claypool, 2011. → pages 1

[7] G. Carenini, C. Conati, E. Hoque, B. Steichen, D. Toker, and J. Enns. Highlighting interventions and user differences: informing adaptive information visualization support. In *Proc. CHI*, pages 1835–1844. ACM, 2014. → pages 44

[8] S. Carpendale. Evaluating information visualizations. In *Information Visualization*, pages 19–45. 2008. → pages 39

[9] J. Choo, C. Lee, C. K. Reddy, and H. Park. Utopian: User-driven topic modeling based on interactive nonnegative matrix factorization. *IEEE Trans. Visualization & Comp. Graphics*, 19(12):1992–2001, 2013. → pages 17, 18, 19, 26, 39

[10] J. Chuang, C. D. Manning, and J. Heer. Termite: Visualization techniques for assessing textual topic models. In *Proc. International Working Conf. on Advanced Visual Interfaces*, pages 74–77, 2012. → pages 17

[11] J. Chuang, D. Ramage, C. Manning, and J. Heer. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proc. Conf. CHI*, pages 443–452. ACM, 2012. → pages 41

[12] J. Chuang, S. Gupta, C. Manning, and J. Heer. Topic model diagnostics: Assessing domain relevance via topical alignment. In *Proc. Conf. on Machine Learning*, pages 612–620, 2013. → pages 16

[13] J. Chuang, Y. Hu, A. Jin, J. D. Wilkerson, D. A. McFarland, C. D. Manning, and J. Heer. Document exploration with topic modeling: Designing interactive visualizations to support effective analysis workflows. In *NIPS Workshop on Topic Models: Computation, Application and Evaluation*, 2013. → pages 17, 18, 39

[14] A. Cockburn, A. Karlson, and B. B. Bederson. A review of overview+ detail, zooming, and focus+ context interfaces. *ACM Computing Surveys (CSUR)*, 41(1):2, 2008. → pages 11

[15] C. Conati, E. Hoque, D. Toker, and B. Steichen. When to adapt: Detecting user's confusion during visualization processing. In *Proc. International Workshop on User-Adaptive Information Visualization (WUAV 2013), in conjunction with conf. on User Modeling, Adaptation and Personalization (UMAP 2013)*, 2013. → pages 44

[16] C. Conati, G. Carenini, E. Hoque, B. Steichen, and D. Toker. Evaluating the impact of user characteristics and different layouts on an interactive visualization for decision making. *Computer Graphics Forum (Proc. EuroVis)*, 33(3):371–380, 2014. ISSN 1467-8659. → pages 44

[17] E. Cutrell and Z. Guan. What are you looking for?: an eye-tracking study of information usage in web search. In *Proc. CHI*, pages 407–416. ACM, 2007. → pages 38

[18] K. Dave, M. Wattenberg, and M. Muller. Flash forums and forumreader: navigating a new kind of large-scale online discussion. In *Proc. ACM Conf. on CSCW*, pages 232–241, 2004. → pages 9, 10, 30

[19] M. De Choudhury and H. Sundaram. Why do we converse on social media?: an analysis of intrinsic and extrinsic network factors. In *Proceedings of the 3rd ACM SIGMM international workshop on Social media*, pages 53–58. ACM, 2011. → pages 30

[20] N. Diakopoulos, M. Naaman, and F. Kivran-Swaine. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 115–122. IEEE, 2010. → pages 32

[21] J. Donath. A semantic approach to visualizing online conversations. *Communications of the ACM*, 45(4):45–49, 2002. → pages 9

[22] J. Donath, K. Karahalios, and F. Vigas. Visualizing conversation. *Journal of Computer-Mediated Communication*, 4(4):0–0, 1999. ISSN 1083-6101. → pages 10, 29

[23] M. Dork, S. Carpendale, C. Collins, and C. Williamson. Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Trans. Visualization & Comp. Graphics*, 14(6):1205–1212, 2008. → pages 31

[24] M. Dork, D. Gruen, C. Williamson, and S. Carpendale. A visual backchannel for large-scale events. *IEEE Trans. Visualization and Computer Graphics*, 16(6):1129–1138, 2010. → pages 31, 37, 40

[25] M. Dörk, R. Comber, and M. Dade-Robertson. Monadic exploration: seeing the whole through its parts. In *Proc. CHI*, pages 1535–1544, 2014. → pages 35

[26] W. Dou, L. Yu, X. Wang, Z. Ma, and W. Ribarsky. Hierarchicaltopics: Visually exploring large text collections using topic hierarchies. *IEEE Trans. Visualization & Comp. Graphics*, 19(12):2002–2011, 2013. → pages 9, 32

[27] M. Drk, N. H. Riche, G. Ramos, and S. Dumais. Pivotpaths: Strolling through faceted information spaces. *IEEE Trans. Visualization & Comp. Graphics*, 18(12):2709–2718, Dec. 2012. → pages 35

[28] A. Endert, P. Fiaux, and C. North. Semantic interaction for visual text analytics. In *Proc. Conf. CHI*, pages 473–482. ACM, 2012. → pages 19

[29] G. Erkan and D. R. Radev. Lexrank: Graph-based lexical centrality as salience in text summarization. *J. Artif. Intell. Res.(JAIR)*, 22(1):457–479, 2004. → pages 34

[30] T. Furukawa, Y. Matsuo, I. Ohmukai, K. Uchiyama, and M. Ishizuka. Social networks and reading bahavior in blogosphere. In *International AAAI Conf. on Weblogs And Social Media(ICWSM)*, 2007. → pages 10, 29

[31] K. S. Hasan and V. Ng. Automatic keyphrase extraction: A survey of the state of the art. In *Proc. ACL*, 2014. → pages 40

[32] S. Havre, E. Hetzler, P. Whitney, and L. Nowell. Themeriver: visualizing thematic changes in large document collections. *IEEE Trans. Visualization & Comp. Graphics*, 8(1):9–20, 2002. ISSN 1077-2626. → pages 9

[33] M. Hearst. *Search user interfaces*. Cambridge University Press, 2009. → pages 38

[34] M. A. Hearst, M. Hurst, and S. T. Dumais. What should blog search look like? In *Proc. ACM Workshop on Search in Social Media*, pages 95–98, 2008. → pages 10, 27, 29, 30, 35

[35] J. Heer and G. G. Robertson. Animated transitions in statistical data graphics. *IEEE Trans. Visualization & Comp. Graphics*, 13(6):1240–1247, 2007. → pages 22

[36] E. Hoque and G. Carenini. Convis: A visual text analytic system for exploring blog conversations. *Computer Graphics Forum (Proc. EuroVis)*, 33(3):221–230, 2014. ISSN 1467-8659. → pages 7, 10, 18, 32, 33, 44

[37] E. Hoque and G. Carenini. ConVisIT: Interactive topic modeling for exploring asynchronous online conversations. In *Proc. ACM conf. on Intelligent User Interfaces*, 2015. → pages 44

[38] E. Hoque, G. Carenini, and S. R. Joty. Interactive exploration of asynchronous conversations: Applying a user-centered approach to design a visual text analytic system. In *Proc. Workshop on Interactive Language Learning, Visualization, and Interfaces (ILLVI 2014), in conjunction with the ACL-2014*. → pages 7, 44

[39] Y. Hu, J. Boyd-Graber, B. Satinoff, and A. Smith. Interactive topic modeling. *Machine Learning*, 95(3):423–469, 2014. → pages 17, 18, 19, 26

[40] Q. Jones, G. Ravid, and S. Rafaeli. Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research*, 15(2):194–210, 2004. → pages 1, 11, 29

[41] S. Joty, G. Carenini, and R. T. Ng. Topic segmentation and labeling in asynchronous conversations. *Journal of Artificial Intelligence Research*, 47: 521–573, 2013. → pages 1, 2, 3, 5, 6, 11, 16

[42] M. Kaisser, M. A. Hearst, and J. B. Lowe. Improving search results quality by customizing summary lengths. In *ACL*, pages 701–709. Citeseer, 2008. → pages 39

[43] B. K. Kaye. Web side story: An exploratory study of why weblog users say they use weblogs. AEJMC Annual Conf., 2005. → pages 10, 27, 29, 30

[44] D. Kim and T. J. Johnson. Political blog readers: Predictors of motivations for accessing political blogs. *Telematics and Informatics*, 29(1):99–109, Feb. 2012. → pages 11, 30

[45] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Trans. Visualization & Comp. Graphics*, 18(9):1520–1536, 2012. → pages 13, 36

[46] B. Lee, G. Smith, G. G. Robertson, M. Czerwinski, and D. S. Tan. Facetlens: exposing trends and relationships to support sensemaking within faceted datasets. In *Proc. CHI*, pages 1293–1302, 2009. → pages 35

[47] H. Lee, J. Kihm, J. Choo, J. Stasko, and H. Park. ivisclustering: An interactive visual document clustering via topic modeling. In *Computer Graphics Forum*, volume 31, pages 1155–1164. Wiley Online Library, 2012. → pages 17, 18, 19, 26, 39

[48] S. Liu, X. Wang, J. Chen, J. Zhu, and B. Guo. Topicpanorama: a full picture of relevant topics. → pages 32

[49] C. Macdonald, R. L. Santos, I. Ounis, and I. Soboroff. Blog track research at TREC. *SIGIR Forum*, 44(1):58–75, 2010. ISSN 0163-5840. → pages 11, 30, 31

[50] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proc. CHI*, pages 227–236. ACM, 2011. → pages 32

[51] S. McKenna, D. Mazur, J. Agutter, and M. Meyer. Design activity framework for visualization design. *IEEE Trans. Visualization and Computer Graphics*. → pages 40

[52] G. Mishne. Information access challenges in the blogspace. In *Workshop on Intelligent Information Access (IIIA)*, 2006. → pages 11, 30

[53] S. A. Munson and P. Resnick. Presenting diverse political opinions: how and how much. In *Proc. Conf. CHI*, pages 1457–1466, 2010. → pages 10, 29

[54] T. Munzner. A nested model for visualization design and validation. *IEEE Trans. Visualization & Comp. Graphics*, 15(6):921–928, 2009. → pages 3, 10, 29, 40

[55] S. Narayan and C. Cheshire. Not too long to read: The tldr interface for exploring and navigating large-scale discussion spaces. In *Hawaii Conf. on System Sciences (HICSS)*, pages 1–10, 2010. → pages 9, 30

[56] M. E. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004. → pages 21

[57] V. Pascual-Cid and A. Kaltenbrunner. Exploring asynchronous online discussions through hierarchical visualisation. In *IEEE Conf. on Information Visualization*, pages 191–196, 2009. → pages 8, 9

[58] J. Preece, Y. Rogers, and H. Sharp. *Interaction Design: Beyond Human-computer Interaction*. Interaction Design: Beyond Human-computer Interaction. J. Wiley & Sons, 2002. → pages 30

[59] D. Ramage, D. Hall, R. Nallapati, and C. D. Manning. Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. In *Proc. Conf. on EMNLP*, pages 248–256, 2009. → pages 17

[60] W. Sack. Conversation map: an interface for very-large-scale conversations. *Journal of Management Information Systems*, 17(3):73–92, 2000. → pages 9

[61] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: reflections from the trenches and the stacks. *IEEE Trans. Visualization & Comp. Graphics*, 18(12):2431–2440, 2012. → pages 3, 40

[62] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000. → pages 21

[63] P. V. Singh, N. Sahoo, and T. Mukhopadhyay. Seeking variety: A dynamic model of employee blog reading behavior. *Available at SSRN 1617405*, 2010. → pages 11, 30

[64] M. Steinberger, M. Waldner, M. Streit, A. Lex, and D. Schmalstieg. Context-preserving visual links. *IEEE Trans. Visualization & Comp. Graphics*, 17(12):2249–2258, 2011. → pages 12

[65] B. Sun, P. Mitra, C. L. Giles, J. Yen, and H. Zha. Topic segmentation with shared topic detection and alignment of multiple documents. In *Proc. annual international conference on Research and development in information retrieval*, pages 199–206. ACM, 2007. → pages 33

[66] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011. → pages 2, 11

[67] J. Vassileva, C. Gutwin, et al. Exploring blog archives with interactive visualization. In *Proc. working Conf. on Advanced visual interfaces*, pages 39–46. ACM, 2008. → pages 28, 31, 35, 37

[68] F. B. Viégas, S. Golder, and J. Donath. Visualizing email content: portraying relationships from conversational histories. In *Proc. Conf. CHI*, pages 979–988, 2006. → pages 9, 31, 40

[69] C. Wang, M. Danilevsky, N. Desai, Y. Zhang, P. Nguyen, T. Taula, and J. Han. A phrase mining framework for recursive construction of a topical hierarchy. In *Proc. international conference on Knowledge discovery and data mining*, pages 437–445. ACM, 2013. → pages 33

[70] M. Wattenberg and D. Millen. Conversation thumbnails for large-scale discussions. In *extended abstracts on CHI*, pages 742–743, 2003. → pages 8, 9

[71] F. Wei, S. Liu, Y. Song, S. Pan, M. X. Zhou, W. Qian, L. Shi, L. Tan, and Q. Zhang. Tiara: a visual exploratory text analytic system. In *Proc. ACM Conf. on Knowledge Discovery and Data Mining*, pages 153–162, 2010. → pages 9, 31

[72] W. Willett, J. Heer, and M. Agrawala. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Trans. Visualization and Comp. Graphics*, 13(6):1129–1136, 2007. → pages vii, 35, 36

[73] Y. Wu, S. Liu, K. Yan, M. Liu, and F. Wu. Opinionflow: Visual analysis of opinion diffusion on social media. *IEEE Trans. Visualization and Computer Graphics*, 20(12):1763–1772, Dec 2014. → pages 32

[74] A. R. Zinman. *Me, myself, and my hyperego: understanding people through the aggregation of their digital footprints*. PhD thesis, MIT, 2011. → pages 10, 29, 30

Appendix A

Supporting Materials

The following research papers have been published/ submitted as part of the completed/ ongoing work:

1. E. Hoque. Visual Text Analytics for Social Media Conversations, *Doctoral Consortium, ACM Conf. on Intelligent User Interfaces* (submitted), 2015.

2. E. Hoque and G. Carenini, ConVisIT: Interactive Topic Modeling for Exploring Asynchronous Online Conversations, *ACM Conf. on Intelligent User Interfaces* (to appear), 2015.

3. E. Hoque and G. Carenini, ConVis: A Visual Text Analytic System for Exploring Blog Conversations, *Journal of Computer Graphics Forum (Proc. EuroVis), 33*, 3 (2014), 221-230, 2014.

4. E. Hoque and G. Carenini, and S. Joty. ”Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System.”, In Proc. Workshop on *Interactive Language Learning, Visualization, and Interfaces (ILLVI 2014)*, in conjunction with the ACL-2014, Baltimore, USA.

5. E. Hoque. A Visual Text Analytic System for Exploring Asynchronous Online Conversations, *Research Proficiency Evaluation (RPE) report*, Dept. Computer Science, University of British Columbia, November 2013.