

Improving the Scalability of Interactive Visualization Systems for Exploring Threaded Conversations

A. McNutt, IEEE student member, & G. Kindlmann

Department of Computer Science, University of Chicago

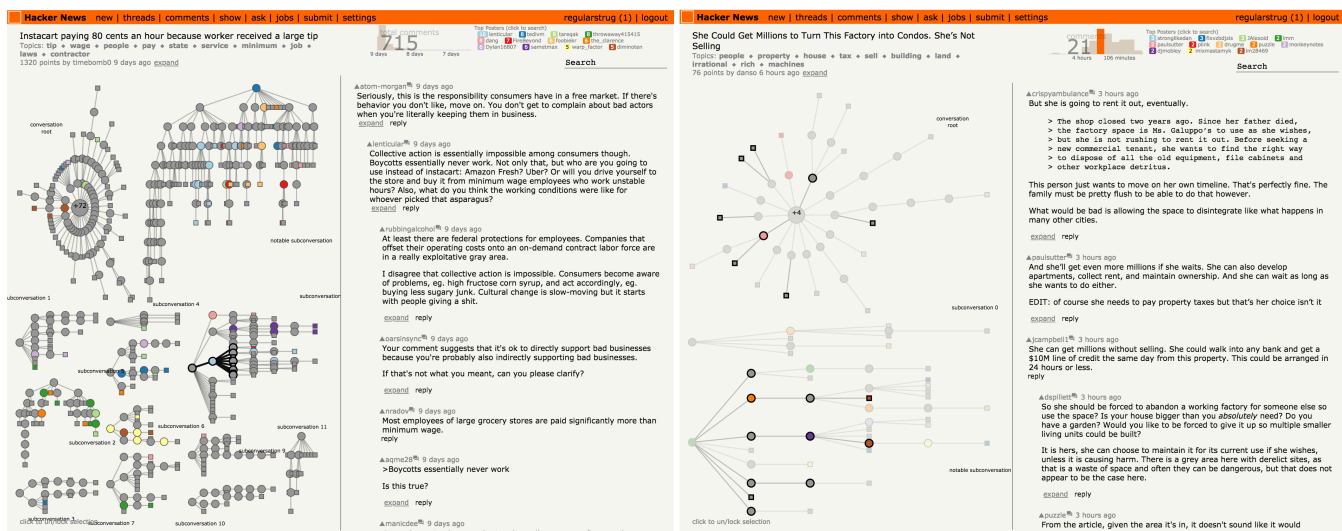


Figure 1: Two very differently sized HackerNews conversations (715 and 85 comments respectively) rendered using our ForumExplorer application. The right image shows a temporal search state.

Abstract

The ABSTRACT is to be in fully-justified italicized text, between two horizontal lines, in one-column format, below the author and affiliation information. Use the word “Abstract” as the title, in 9-point Times, boldface type, left-aligned to the text, initially capitalized. The abstract is to be in 9-point, single-spaced type. The abstract may be up to 3 inches (7.62 cm) long. Leave one blank line after the abstract, then add the subject categories according to the ACM Classification Index (see <http://www.acm.org/about/class/class/2012>)

CCS Concepts

• **Computing methodologies** → Collision detection; • **Hardware** → Sensors and actuators; PCB design and layout;

1. Introduction

Conversation on the internet takes many forms and shapes, including question+answer forums, synchronous messaging, and asynchronous threaded conversation. Of particular interest are conversations that take place in an asynchronous threaded environments, such as reddit or slash dot, as they offer a mechanism for communities to have wide collections of related conversations within a single topic. In many cases the participants in these conversations

are experts on the topic and might provide valuable insights. Unfortunately the design of these digital spaces typically do not allow for users to interact with the conversational corpus as a whole, which can limit or impede understanding of the community opinions and insights about a topic.

Building systems that provide an overview of a complex or non-linear data space is a favorite class of problems in the information visualization community, and threaded conversations have been no

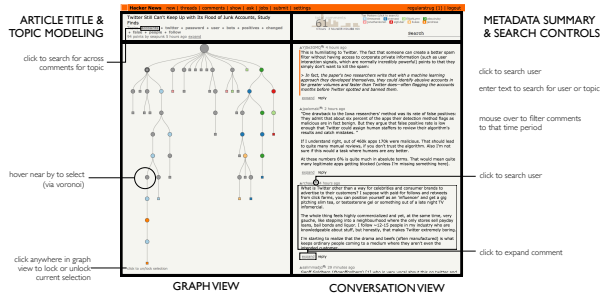


Figure 2: Annotated view explaining the components of the UI.

exception. Previous visualization works have developed a fascinating collection of UI paradigms to address this space. A common trend among these works features a overview of the conversational thread by representing it as graph like structure, which the user then interacts with in a details-on-demand [Shn96] pattern to expose components of the discourse. While these tools are uniformly well received by their evaluation audiences, they have failed to gain widespread usage; possibly because they do not possess an accessible or online implementations, are not aligned to the specific community they are trying to affect, and are designed around a single ideal size of conversation (and thus become cumbersome or difficult to use when conversations of interest fall outside of that target domain). Further, it is possible that this type of tool is in fact not necessary for the types of tasks that most users pursue within these spaces.

In this work we present Forum Explorer, a Chrome Extension that repurposes the conventional layout of yCombinator’s HackerNews [yCo] to facilitate better data exploration through the use of tree visualization techniques. Our system expands upon previous techniques developed and contributes several novel visual encodings to facilitate. Specifically, we introduce a novel Forested Tree View which splits threaded conversations at the root into a collection of smaller and more legible trees, this allows for ample visual space to provide in-situ annotations and textual guides. We construct a solution for reducing visual noise in large conversations by collapsing rooted stumps into the root, while still making those nodes available. Our solution seeks to make this type of tool easily available through the use of easily reproducible components.

2. Related Work

Previous work on this topic has developed a number of different strategies for displaying and facilitating novel user interactions under a variety of different design goals. The first work of this type appears to be Donath et al’s Loom, which introduced the graphical exploration of comments as a collection of nodes in space connected together through polylines [DKV99]. This was quickly followed by Sack’s Conversation Map which added a more tree-like structure to the display [Sac00]. Wattenberg et al present a pair of papers which appear to be the first instance of a split pane view, with one providing the graphical overview (which was a graphical abstraction of the multiply indented form threaded conversation are usually depicted in) and the other displaying comments

[WM03, DWM04]. Pascual-Cid et al introduce a space filling radial tree layout. Narayan construct tldr which focuses on reddit and encodes the tree as an icicle diagram [NC10]. Hoque et al’s family of tools breaks from purely metadata visualization by adding NLP-based-summaries like faceted topic modeling and sentiment analysis on top of a split pane interaction with a abstracted indentation graphical summary [HC14, HC16]. These tools provide valuable insight, but come at the cost of forcing their users to use an unfamiliar environment, as well as being not being actually deployed. Most closely related to our technical approach is Treeverse, a chrome extension which allows users to visualize the conversation graph associated with a particular tweet. [Bul] When the user is engaged with a particular tweet they can opt into the Treeverse view which renders the graph as a click and draggable tree with associated secondary panel. Our work improves over this design in that it tightens the response interaction response cycle and enables the user to extract useful information from a single view.

Each of these works features a ringingly positive user evaluation. It is unclear whether or not this result is due to the novelty of the system, and therein experiment bias, or the true utility of the tool. The best evidence that this type of system has substantive utility is that it used by Rao et al [RS18] to study the way that twitter is used SOMETHING domain specific knowledge sharing discussions. This suggests this type of data exploration tool has utility beyond academia.

3. Forum Explorer

TODO NEED TO SCRAPE HACKERNEWS TO FIGURE OUT MAX COMMENT THREAD

We implement a split pane view in which a graphical overview is presented on the left and a detail view on the right. The graphical overview consists of a one of a collection of tree visualizations with the origin comment represents as the root of the tree.

We provide additional detail about the UI in Figure 2. In addition to our Forested View we provide several additional view options to enable users a sense of play and to facilitate exploration within the tool, these include an unsegmented radial layout, a reifold-something tidy tree CITATION, a gridded tree CITATION, and an orbit layout CITATION.

Our primary contribution in this work is addition of a novel visual encoding in the overview representation. We observe that many presentation of trees such as THE RADIAL TREE ONE OR CONVIS scale poorly when presented a large collection of nodes. To address this issue we observe that their tends to be zipfs law distribution of weights for branches incident to the root, so we prune the heaviest branches and layout them out as adjacent conversations

4. Conclusions & Future Work

We have presented Forum Explorer, a tool for exploring threaded conversations on HackerNews. Our primary contribution is a novel Forested Tree layout that facilitates better scalability for systems of this type.

The specific utility of this type of conversation overview tool

remains unclear, although based contemporary developments, such as Treeverse's 2412 active installs as well as its usage in

Previous studies of threaded conversation have typically involved a small scale user study, SPECIFICS. While these studies usefully demonstrate the usability of the structural design elements in general, their context as laboratory based analyses precludes them from providing longitudinal information about the way users might interact with this type of tool when they are able to incorporate it into their day to day workflow. Forum Explorer is well positioned

Our current design process has, in the lens of data feminism, been unfeminist. Our design process happened without consultation of those who might find the application (other than perhaps ourselves). While we are able to rest on the many of the observations from ConVis's prior user studies in the topic, individual forum users tend to have individual expectations about their tools.

The specific utility of this type of conversation overview tool remains unclear, the lessons learned from ours and other systems could readily be applied to graph analysis problems of a similar size and scale. Analysis of locally relevant articles in the scholarship graph are one such example of this type of system.

References

- [Bul] BULTER P.: Treeverse. <https://treeverse.app/>. 2
- [DKV99] DONATH J., KARAHALIOS K., VIEGAS F.: Visualizing conversation. *Journal of computer-mediated communication* 4, 4 (1999), JCMC442. 2
- [DWM04] DAVE K., WATTENBERG M., MULLER M.: Flash forums and forumreader: navigating a new kind of large-scale online discussion. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (2004), ACM, pp. 232–241. 2
- [HC14] HOQUE E., CARENINI G.: Convis: A visual text analytic system for exploring blog conversations. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 221–230. 2
- [HC16] HOQUE E., CARENINI G.: Interactive topic modeling for exploring asynchronous online conversations: Design and evaluation of convisit. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 6, 1 (2016), 7. 2
- [NC10] NARAYAN S., CHESHIRE C.: Not too long to read: The tldr interface for exploring and navigating large-scale discussion spaces. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (2010), IEEE, pp. 1–10. 2
- [RS18] RAO S., STEIN R.: To journey in twitter canoes: Methods to understand the mechanisms and meaning in twitter conversations. In *MW18: Museums and the Web 2018* (2018). 2
- [Sac00] SACK W.: Conversation map: An interface for very large-scale conversations. *Journal of Management Information Systems* 17, 3 (2000), 73–92. 2
- [Shn96] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on* (1996), IEEE, pp. 336–343. 2
- [WM03] WATTENBERG M., MILLEN D.: Conversation thumbnails for large-scale discussions. In *CHI'03 extended abstracts on Human factors in computing systems* (2003), ACM, pp. 742–743. 2
- [yCo] YCOMBINATOR: HackerNews. <https://news.ycombinator.com/>. 2