

Visual Analytics Support for Intelligence Analysis

Carsten Görg*
University of Colorado

Youn-ah Kang†
Google Inc.

Zhicheng Liu‡
Stanford University

John Stasko§
Georgia Institute of Technology

ABSTRACT

Intelligence analysis challenges investigators to examine large collections of data and documents and come to a deeper understanding of the information and events contained within them. Visual analytics technologies hold great promise as potential aids for intelligence analysis professionals. We describe our research to better understand intelligence analysis processes and analysts, learn how visual analytics can help investigators, and design visual analytics systems to serve in this role. To illustrate these ideas, we present a hypothetical intelligence analysis scenario that explores a collection of text documents using the Jigsaw system that we have created. The system combines computational analysis of document text with interactive visualizations of the document contents and analysis results. Evaluating such systems is very challenging and the article concludes by discussing potential evaluation methodologies for these types of systems.

Keywords: Visual analytics, investigative analysis, intelligence analysis, information visualization, knowledge acquisition, data exploration, case study, qualitative user study.

1 INTRODUCTION

Visual analytics is a relatively new research field that integrates the interactive visualization and exploration of data with computational data analyses [8]. Intelligence analysis has been one of the key application domains of visual analytics since the area's inception in 2004, facilitated by the creation of the National Visualization and Analytics Center by the Department of Homeland Security. An initial research roadmap [11] described challenges and goals of the new field and identified tasks, data, and analytical scenarios focused on homeland security and prevention of terrorism.

Enabling insights through the analysis of large amounts of diverse and dynamic data was the underlying grand challenge in the research agenda. As stated in [11], “The analysis of overwhelming amounts of disparate, conflicting, and dynamic information is central to identifying and preventing emerging threats, protecting our borders, and responding in the event of an attack or other disaster. This analysis process requires human judgment to make the best possible evaluation of incomplete, inconsistent, and potentially deceptive information in the face of rapidly changing situations to both detect the expected and discover the unexpected.”

Intelligence analysis requires investigators to gather as much available data as possible in order to better understand a situation and then make judgments about the appropriate next steps to take. Two fundamental types of investigative scenarios exist within the intelligence domain: (1) targeted analysis scenarios, in which analysts are tasked with examining specific people, organizations, or incidents, as well as locations and dates, in order to either investigate past events or uncover an imminent threat, and (2) open ended,

strategic analysis scenarios, in which analysts are tasked with learning as much as possible about a person, organization, country, or situation in order to gain a deeper understanding, conduct an accurate assessment, and possibly make a prediction on the likely chain of events that will occur at a later point in time. Examining and understanding large collections of textual documents plays an important role in both types of these scenarios. Analysts must gather nuggets of information within textual documents from diverse sources, ranging from reports from field agents to open source news articles. Examining textual documents is fundamentally a slow process (due to the sequential nature of reading) and it is challenging for the analysts to keep track of what they discovered and form an internal mental model that represents a coherent picture of the events, people, places, and organizations discussed in the documents. Uncovering and understanding the connections between those entities across a large collection of documents is one of the key challenges they face.

Based on a cognitive task analysis of working analysts, Pirolli and Card [9] identified a number of “pain points” in the intelligence process that are particularly challenging to human analysts. These pain points include the costs of scanning, recognizing (assessing), and selecting items for further attention; the costs of shifting attention and control; the limited span of attention for evidence and hypotheses; and the difficulty of generating alternative hypotheses. All these challenges are exacerbated when the amount of data to examine grows larger and larger. Today’s “big data” technologies often make the acquisition of data easier, but they present increasing challenges to analysts who must review and investigate all that data. In this article we highlight a number of our research projects on intelligence analysis from the last five years, including an observational study to gain a better understanding of the intelligence analysis process and its characteristics, the development of a visual analytics system that integrates computational text analyses with interactive visualization in order to explore collections of documents, and an evaluation of the utility of the system via a controlled laboratory experiment as well as observational case studies of extended use of the system in the field.

2 INTELLIGENCE ANALYSIS PROCESS

Analyzing and understanding end-users needs and tasks is one of the fundamental requirements for creating useful computational tools. To better understand intelligence analysis, it is important to explore the mindset and methodologies of analysts as well as the fundamental processes they conduct. Heuer [4] examined the psychology of intelligence analysis and the types of mental reasoning analysts must engage in. In particular, he identified a number of challenges analysts must confront in the analytical reasoning process. For example, when people encounter a situation of uncertainty, they typically will develop a single hypothesis explaining the situation and will work to gather evidence confirming the hypothesis. Intelligence analysts, however, are trained to develop multiple hypotheses and seek out information that can discredit many of the hypotheses.

Many other researchers have studied the intelligence analysis process in order to construct abstract models of it. While a number of process models exist, most involve some form of iterative cycle of exploration, including steps such as data collection, pro-

*e-mail: Carsten.Goerg@ucdenver.edu

†e-mail: ykang@google.com

‡e-mail: zcliu@cs.stanford.edu

§e-mail: stasko@cc.gatech.edu

cessing, analysis and production, dissemination, and planning and direction [1].

Pirolli and Card's notional model of the sensemaking loop for intelligence analysis [9] has been widely cited and adopted by researchers within the visual analytics community. It consists of a linear set of states characterizing both data and process flow in an investigation. Analysts iterate through this process over the course of an investigation. At a high level, the model contains two primary loops: a foraging loop in which analysts collect data and evidence, and a sensemaking loop in which analysts reflect on the data in order to generate schema and hypotheses about the situation and ultimately construct a presentation of the findings. Each loop contains three stages that further refine the process and both loops are connected through an overarching reality/policy loop.

This model broadly characterizes the workflow used for analysis activities and it has guided the development of a number of computational tools. However, it abstracts substantially from analysts' work in the real world and does not provide an adequate level of detail necessary to develop tools that analysts can integrate seamlessly with their existing workflow. Furthermore, not all analysts agree that the linear structure of the model captures the way they work. Dr. Kristan Wheaton, Professor at the Department of Intelligence Studies at Mercyhurst College, proposed an alternative model in which modeling, collection, analysis, and production stages take place in parallel, just with different emphases over the course of an investigation. At the beginning, emphasis focuses on modeling and then throughout the investigation it shifts to collection, analysis, and finally production.

To better understand the analytical process and its requirements in the intelligence domain, we conducted our own qualitative user study [6]. Professor Wheaton provided us with the opportunity to observe three teams of intelligence analysts in training within the intelligence program at Mercyhurst College. The student teams each conducted an intelligence analysis project throughout an entire academic term (ten weeks). One team consisted of four undergraduate students and performed analysis on a project for which we served as a "client"; the other two teams consisted of graduate students and conducted a structured analysis on projects provided by external clients.

We found that four processes dominated the overall workflow: construction of a conceptual model, collection, analysis, and production. The study helped us to better understand some misconceptions that visual analytics researchers may harbor about intelligence analysis. For instance, analysis is typically not about finding an answer to a specific problem and it does not evolve in a sequential process. Instead, analysis is often about determining how to answer a question, what to research, what to collect, and what criteria to use. The process is often organic and parallel. Another misconception is that intelligence analysts typically operate as lone investigators, researching some problem. We, conversely, found that collaboration is commonplace and crucial, frequently being asynchronous. Also, the student analysts we observed did not seek grand, monolithic computational analysis tools. Instead, the teams used a variety of computational tools with many being small applications used for one specific purpose. They sought ways to integrate existing tools and easy-to-use new tools that leveraged existing analysis methods.

Finally, our study surfaced a number of recommendations for visual analytics technology developers:

- Externalize the thinking process - Help analysts continuously build a conceptual model
- Support source management - Enable managing both pushed and pulled information and organizing sources meaningfully
- Support analysis with constantly changing information - Integrate collection and analysis in a single system and help analysts use structured methods during collection

- Help analysts create convincing production - Support insight provenance and sanity checks of analytical products
- Support asynchronous collaboration rather than synchronous collaboration for exploratory analysis

3 EXAMPLE SCENARIO EMPLOYING VISUAL ANALYTICS

In order to better demonstrate how visual analytics can aid intelligence analysis, we present an example scenario. "The 9/11 Commission Report" is a publicly available report about the 9/11 terrorist attacks on the World Trade Center in New York. One version of the report is stored as a pdf document with 585 pages. In order to better simulate a larger collection of short intelligence reports, we split this document into 585 pages and consider each as a separate document. We use the page breaks as separators since the report does not have a natural structure that would lend itself to being split into short documents of a few paragraphs.

To illustrate this scenario, we employ the Jigsaw visual analytics system [10]. Jigsaw combines automated text analyses with interactive visualizations for exploring and analyzing collections of unstructured and semi-structured text documents. It automatically identifies entities of interest in the documents, such as people, places, and organizations, and then shows connections between those entities across the entire collection, as well as connections between documents and entities. Connections are defined by co-occurrence: if two entities co-occur in the same document, they are connected to each other as well as to that document. If entities co-occur in many documents they have a stronger connection. Even though this untyped connection model based on co-occurrence is very simple, it has turned out to be a powerful tool for investigative analyses. It works best if the documents are not too large, as it is often the case for news articles or case reports that usually span a few paragraphs.

We present the scenario from the point of view of a hypothetical intelligence analyst who is examining the document collection. To begin, the analyst imports the 585 single-page documents and runs an automatic entity identification. She uses the integrated OpenCalais webservice to identify people, locations, and organizations; the integrated GATE package to identify money entities, and built-in regular expression matching algorithms for identifying date entities. The analyst removes all entities that occur in only one document (they would not contribute to any connections) and performs a basic entity clean-up process, including removing wrongly identified entities and aliasing entities with multiple representations such as "George Bush" and "George W. Bush". The entire process results in a document collection with 369 people, 200 location, 252 organization, 12 money, and 464 date entities across the 585 documents.

The analyst begins the investigation seeking an overview of the entities. She uses the List View to display lists for Location, Person, Organization, and Money and change the list ordering from alphabetic to frequency-of-occurrence to see the most frequent entities in the document collection (Figure 1). The small bar to the left of each entity indicates the number of documents in which it occurs. Entities with aliases are shown in italic font and the aliases are displayed as tool tips, as shown for *Usama Bin Ladin*. The most frequent locations are *United States* (364 occurrences), *Afghanistan* (184), *Pakistan* (98), *New York* (77), and *Saudi Arabia* (71); the most frequent persons are *Bill Clinton* (65), *George W. Bush* (59), *Usama Bin Ladin* (59), *Richard Clarke* (50), and *George Tenet* (36); and the most frequent organizations are *al-Qeda* (233), *Central Intelligence Agency* (214), *Federal Bureau of Investigation* (181), *White House* (93), and *Federal Aviation Administration* (91).

The analyst next selects *Usama Bin Ladin* in the person list and reorders the other lists by strength of connection to the selection in order to see the entities most common with him (Figure 2, left).

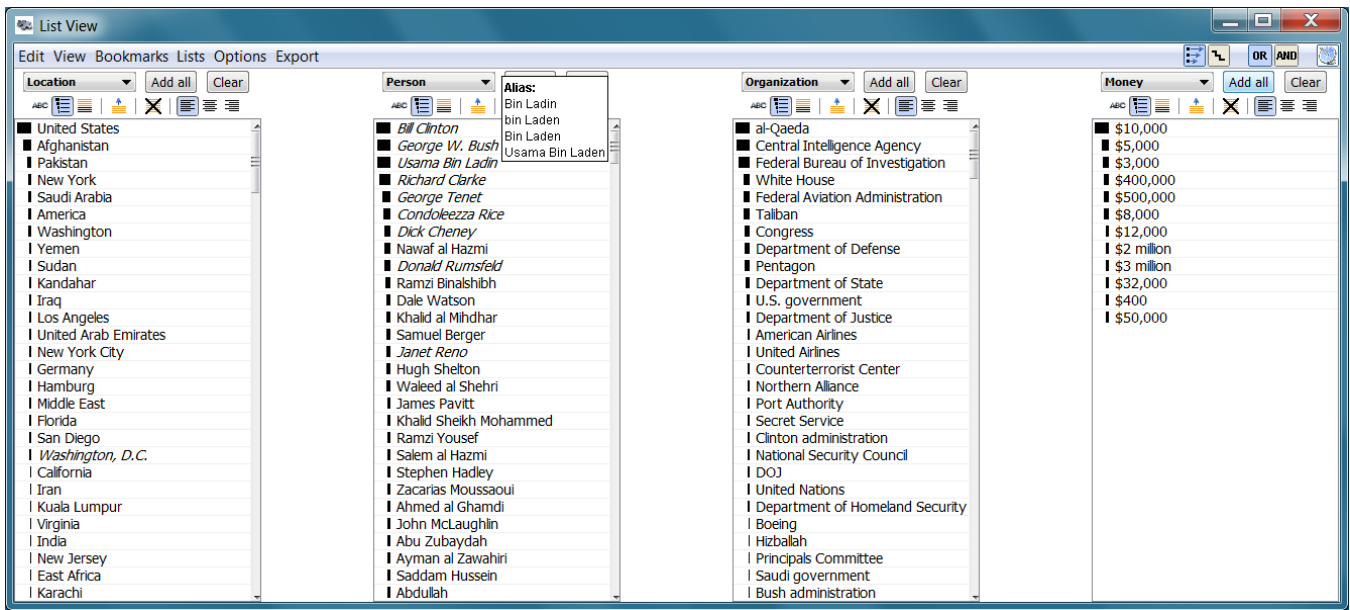


Figure 1: List View showing an overview of the 9/11 Commission Report, focusing on the Location, Person, Organization, and Money entities. All entities are sorted by frequency.

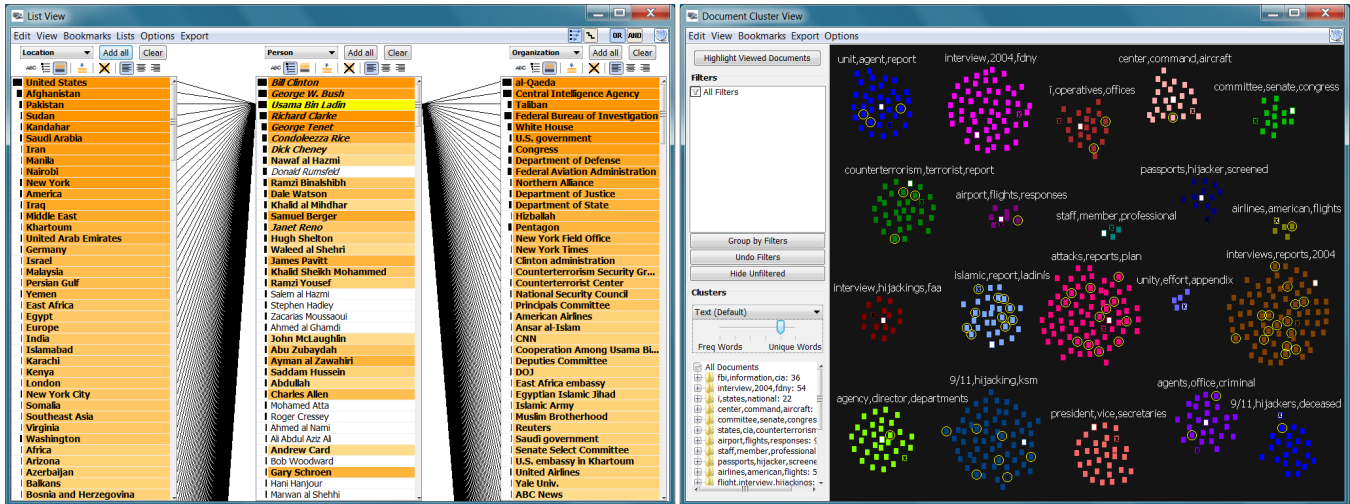


Figure 2: List View (left) showing locations, persons, and organizations connected to *Usama Bin Laden*. Document Cluster View (right) showing different clusters of related documents (small rectangles in different colors). Documents connected to *Usama Bin Laden* are selected (surrounded by a yellow circle).

The List View highlights entities connected to a selected entity (yellow) via an orange background. Darker shades of orange indicate or stronger (more frequent) connection. Entities that are not directly connected have a white background. *United States*, *Afghanistan*, *Pakistan*, *Sudan*, and *Kandahar* are the most connected locations to *Usama Bin Laden*; *al-Qaeda*, *Central Intelligence Agency*, *Taliban*, *Federal Bureau of Investigation*, and *White House* are the most connected organizations. He is also strongly connected to the people *Bill Clinton*, *George W. Bush*, *Richard Clarke*, and *George Tenet*.

To better understand the themes and topics in the report, and in particular those in which *Usama Bin Laden* is mentioned, the analyst opens the Document Cluster View and displays the documents clustered by text similarity (Figure 2, right). Each document is displayed as a small rectangle and each cluster is labeled with three keywords. Text analysis algorithms integrated in Jigsaw automatically compute the clusters and summaries. The cluster summaries represent important topics in the report, including counterterrorism, hijackings, attacks, interviews, and president.

Cross-view selection and filtering are important capabilities in visual analytics systems. Since *Usama Bin Laden* is still selected in the List View (Figure 2, left), the documents he appears in are also selected in the Document Cluster View (Figure 2, right), indicated by a yellow circle. He is connected to more than ten documents in the “islamic, report, ladin’s”, “attacks, reports, plan”, and “interviews, reports, 2004” clusters and to seven documents in the “9/11, hijacking, ksm” cluster. The analyst also could use cross-view selection in the opposite direction for a different kind of exploration: when she selects the “president, vice, secretaries” cluster, she observes in the List View that *George W. Bush*, *Dick Cheney*, *Condoleezza Rice*, and *Donald Rumsfeld* are the most connected people to that cluster. Interestingly, *Donald Rumsfeld* is not connected to *Usama Bin Laden* (Figure 2, left).

To learn more about Rumsfeld, the analyst opens the Graph View and explores the people and organizations connected to him using a “circular layout” approach (Figure 3, left). This approach positions the documents that mention *Donald Rumsfeld* on a circle (white

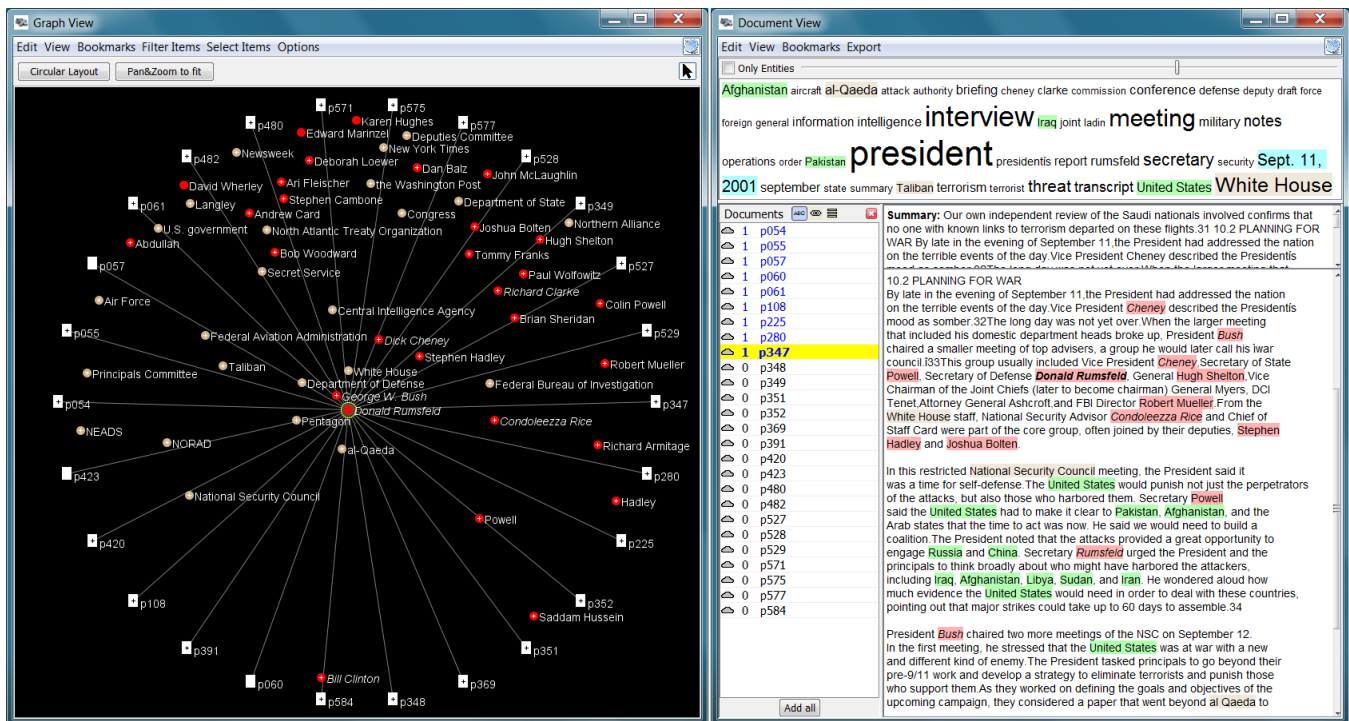


Figure 3: Graph View (left) showing a Circular Layout with documents (white rectangles) that mention *Donald Rumsfeld*. Persons (red circles) and organizations (tan circles) are positioned inside the circle of documents, the more connected they are the closer they are positioned to the center of the circle. Document View (right) showing all documents mentioning *Donald Rumsfeld*. Above the selected document's text (right) is a one sentence summary and below are the affiliated entities. The word cloud (top) summarizes all documents loaded in the view.

rectangles) and the related entities within that circle (red circles for people, tan circles for organizations). More highly connected entities are placed closer to the center of the circle. The layout shows that *Donald Rumsfeld* is strongly connected to *George W. Bush*, *Dick Cheney*, and *Stephen Hadley*, as well as to the organizations *White House*, *Department of Defense*, and *Pentagon*. The Graph View also supports interactive exploration of the connection network via expand and collapse operations. A double click on a document or an entity expands that item and brings in all other items that are connected to it. Items having additional connections that are currently not shown are indicated with a plus sign. A double click on an item that already shows all its connected items (e.g. *Donald Rumsfeld*) collapses that item and hides all its connected items.

The analyst next wants to read the documents about Donald Rumsfeld, so she opens them in the Document View (Figure 3, right). The document list (left) shows the 26 documents (pages) that mention him; documents shown in blue have already been examined, and the number in front of the document indicates how often it was displayed. The word cloud (top) summarizes the currently loaded document set using the most frequent words in those documents. The selected (yellow) document in the list is presented on the right. Above the documented text is a one-sentence summary of the document computed by a text summary analysis. To support quick scanning of documents, entities in the word cloud and in the document itself are highlighted: people in red, locations in green, dates in blue, and organizations in tan. This document (p347) talks about a restricted National Security Meeting on the night of the attacks in which “Rumsfeld urged the President and the principals to think broadly about who might have harbored the attackers, including Iraq, Afghanistan, Libya, Sudan, and Iran.” (The numbers in some sentences in the document are footnote references.)

To investigate if similar documents exist in the collection, the analyst opens the Document Grid View and sorts and colors the

documents by their similarity to document p347 (Figure 4, top). The shading of blue indicates document similarity: dark blue indicates similar documents, light blue indicates documents that do not have much in common with the reference document. A tooltip provides the one sentence summary of a document. A similar document p215 mentions “Bonk told Bush that Americans would die from terrorism during the next four years. During the long contest after election day, the CIA set up an office in Crawford to pass intelligence to Bush and some of his key advisors.” It seems that there might have been some miscommunication in the post election transition. To understand the role of the former president *Bill Clinton*, the analyst displays his name in the Word Tree View (Figure 4, bottom). A Word Tree [12] shows all occurrences of a word or phrase across all documents in the context of the words that follow it. Each word can be explored further by a click. The Word Tree View for *Bill Clinton* shows that his “administration effectively relied on the CIA to take the lead in preparing long-term offensive plans” and that “One of the great regrets of my presidency is that I didn’t get him [Bin Ladin] for you”.

SIDEBAR: VISUAL ANALYTICS TOOLS FOR INTELLIGENCE ANALYSIS

A few commercial tools for intelligence analysis employ visual analytics techniques including Analyst’s Notebook from IBM i2 (<http://www.ibm.com/software/industry/i2software>), nSpace from Oculus (<http://www.oculusinfo.com/nspace>), and Palantir’s suite of systems (<http://www.palantir.com>). An extensive discussion of academic research projects employing visual analytics for understanding text and document collections can be found in [2].

The scenario exploring the 9/11 Commission Report and the images used in this article were produced using the Jigsaw visual analytics system. Jigsaw was designed to help investigators ex-

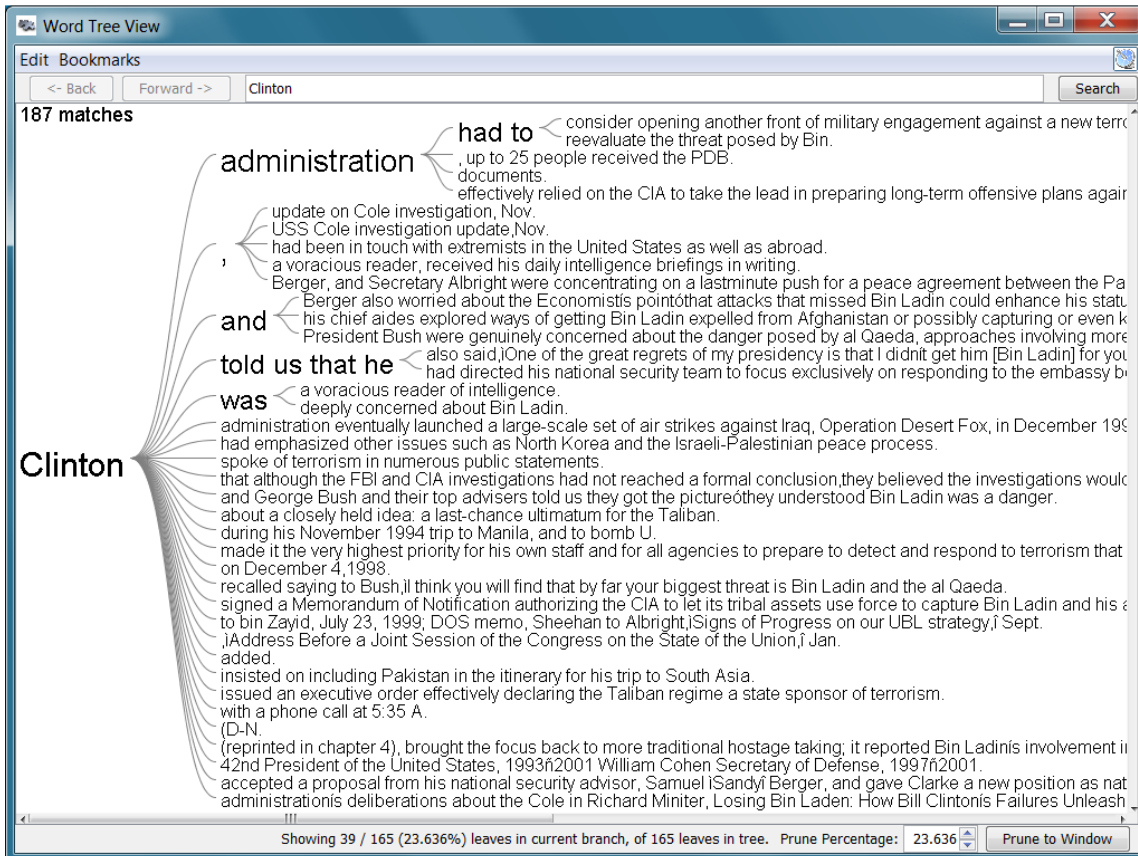
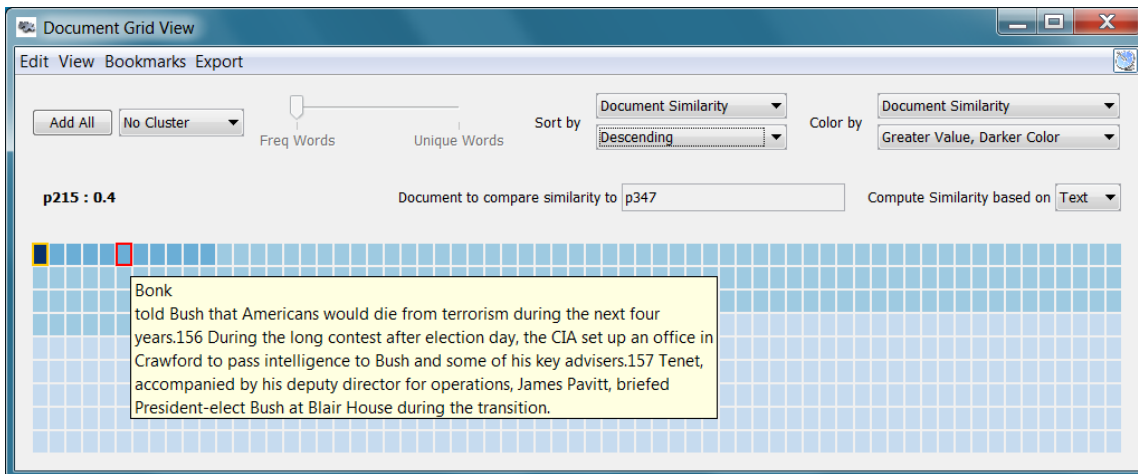


Figure 4: Document Grid View (top) with the document (small rectangle) order and shading set to correspond to the document's similarity to the selected document p347. Word Tree View (bottom) showing occurrences of the person *Clinton* and the most common phrases that follow him in sentences within the 9/11 Commission Report.

explore and understand collections of text documents, and in particular, to follow trails of ideas embedded across the documents. The system's name comes from the notion of "putting the pieces together." Early versions of Jigsaw emphasized a suite of interactive visualizations portraying the documents' contents and connections between entities in the documents [10]. More recently, we have integrated computational text analysis capabilities [2]. Beyond intelligence analysis, Jigsaw has been used to explore consumer review, academic research, fraud, investigative reporting, law enforcement, business intelligence, and email document collections. The Jigsaw system itself, as well as example datasets (including the one used in this article), tutorials, videos, and re-

lated articles are available for download on the webpage (<http://www.cc.gatech.edu/gvu/ii/jigsaw>).

4 EVALUATING SYSTEMS FOR INTELLIGENCE ANALYSIS

The evaluation of visual analytics systems is a challenging research area in itself: there is no consensus among researchers on how to effectively and objectively measure the contributions of a system to an analyst's generation of insights. This is especially true when intelligence analysis is the domain being studied. In our research, we have used multiple techniques to evaluate the effectiveness of the systems we have built.

In one evaluation study [5], we employed a small synthetic dataset with embedded ground truth, consisting of 50 imaginary short reports with a hidden threat. We then recruited sixteen students, divided them into four groups, and asked them to conduct an analysis of the documents and identify the hidden threat. Participants in the first group only worked with pencil and paper. They received a printout of all the reports and some blank sheets for note taking. Participants in the second group received an electronic copy of the reports and could use basic text editing software for reading and searching the documents. Participants in the third group used only the Document View of the Jigsaw system to read and analyze the document collection. This setup was similar to the previous one, providing functionality for reading and searching; however, the Document View also highlighted identified entities within the documents. Participants in the fourth group used the entire suite of visualizations in Jigsaw to conduct the analysis.

The study participants worked with the documents for 90 minutes and then wrote debriefing statements, which we compared to the ground truth and then graded for accuracy. We also conducted follow-up interviews and collected their notes. Additionally, we videotaped all the sessions and used screen capture software in the settings where participants worked on a computer. We started our analysis with an inductive approach to examine the qualitative data in order to unveil potential concepts and themes and to understand the influence of the tools in the different settings. At a later stage of our analysis we combined inductive and deductive approaches and supplemented them with observations from the video logs, screen captures, and other quantifiable data.

We found that overall the participants using the full Jigsaw system outperformed all other groups on average. Because of the small subject population this result was not statistically significant, however. We did observe four particular strategies that participants employed in their investigations. These strategies ranged from first reading all the documents very carefully to finding an initial clue and following a trail from it. The participants using Jigsaw applied three out of the four strategies and performed well using any of those strategies.

We also used Jigsaw for our own investigations and participated in a number of IEEE VAST Conference Challenges and Contests over the past few years. These contests provide synthetic document collections with embedded ground truth. Participating teams are tasked with finding a hidden threat in the documents. Working with Jigsaw on the contest datasets helped us to gain practical experience in these types of intelligence investigations, improve the system, and develop additional functionality. We further describe the influence of our participation in the VAST contests on the design and development of the Jigsaw system in another article [3].

Because we have made Jigsaw available to anyone to use in their own work, examining real world use by other people is another type of evaluation that we have employed. In order to better understand how professional analysts have been using the system and to determine its benefits and limitations in practice, we interviewed six investigators who had been using Jigsaw for an extended period of time [7]. These individuals included an aerospace engineering researcher, a business analyst investigating fraud, a doctoral candidate in Industrial and Systems Engineering studying enterprise transformation, and intelligence analysts at a national lab, the Air Force, and a police department. The goal of this study was to evaluate whether Jigsaw is helping analysts with their tasks, to understand its application to different types of documents and domains, and to identify useful features and capabilities of the system as well as missing or problematic features.

We identified a number of applications of the system across more than one participant. Many used Jigsaw to find connections and relationships between entities, one of the core goals of the system. Some used it as a search and comparison tool to more conveniently

work with text documents, and many used it to gain a broader understanding or overview of their documents. Surprisingly to us, some of the participants also used Jigsaw as a communication aid to share their understanding with others. We originally created the system as an analysis tool and that application is always how we have thought of it. It was interesting to note that some of the study participants also were using it to present findings and tell a story to their colleagues.

The investigators in the study identified a number of limitations and issues with the system as well. Some of the participants wanted better ways to work with only subsets of their document collections. They wanted to be able to dynamically filter out documents in an investigation, but also maintain the ability to reintroduce filtered documents as desired. Document import was another particular challenge and often required manipulating and translating their original documents into a form that Jigsaw could better analyze. Furthermore, problems that arose in document import or in any other use of the system raised questions in the investigators' and their colleagues' minds about the accuracy of the system. They commented how any kind of issue or usability problem eroded their trust.

Our study identified a number of future objectives for Jigsaw and other visual analytics systems for document analysis. The investigators all believed that entity identification is crucial and they wanted easier and more reliable mechanisms to perform it and correct/modify it. They also sought to have more flexible mechanisms for document management activities such as import, storing, filtering, and maintaining. A number of the users wanted more quantitative and statistical analysis capabilities. For instance, they expressed a desire for more network analysis and modeling metrics. In terms of the user interface, some of the investigators wanted to be able to annotate the system views, highlight particular items, and add notes and comments on top of the visual representations.

5 CONCLUSION

Intelligence analysis requires people and organizations to review and assess large collections of information in order to better understand current situations and take the appropriate next steps. The sheer scale, diversity, and complexity of the information to be explored often makes such analysis cognitively demanding. Furthermore, the information is often recorded as narrative text, not quantitative data, and thus it is not as amenable to automated analysis techniques.

Visual analytics technologies that combine computational text analysis with interactive visualization provide a powerful new paradigm for helping intelligence analysts in their work. While current visual analytics systems have illustrated the potential of the field, many challenges remain. Visual analytics systems must scale to increasingly larger collections of data in order to keep up with our growing ability to log and record information. Additionally, visual analytics systems should assist investigators in the complex processes of analytical reasoning, hypothesis formulation, and decision making.

6 ACKNOWLEDGEMENTS

This research is based upon work supported in part by the National Science Foundation via Awards IIS-0414667, IIS-0915788, and CCF-0808863, by the National Visualization and Analytics Center (NVACTM), a U.S. Department of Homeland Security Program, and by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001.

REFERENCES

- [1] Central Intelligence Agency. *A Consumer's Guide to Intelligence*. Diane Pub Co, 1999.

- [2] C. Görg, Z. Liu, J. Kihm, J. Choo, P. Haesun, and J. Stasko. Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw. *IEEE Transactions on Visualization and Computer Graphics*, 2013. to appear.
- [3] C. Görg, Z. Liu, and J. Stasko. Reflections on the Evolution of the Jigsaw Visual Analytics System. *Information Visualization*, 2013. to appear.
- [4] R. Heuer. *Psychology of Intelligence Analysis*. Center for the Study of Intelligence, Central Intelligence Agency, 1999.
- [5] Y.-a. Kang, C. Görg, and J. Stasko. How Can Visual Analytics Assist Investigative Analysis? Design Implications from an Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 17(5):570–583, May 2011.
- [6] Y.-a. Kang and J. Stasko. Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In *IEEE VAST*, pages 21–30, Oct. 2011.
- [7] Y.-a. Kang and J. Stasko. Examining the use of a visual analytics system for sensemaking tasks: Case studies with domain experts. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2869–2878, 2012.
- [8] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. *Information Visualization: Human-Centered Issues and Perspectives*, pages 154–175, 2008.
- [9] P. Pirolli and S. Card. Sensemaking Processes of Intelligence Analysts and Possible Leverage Points as Identified Through Cognitive Task Analysis. In *International Conference on Intelligence Analysis*, May 2005.
- [10] J. Stasko, C. Görg, and Z. Liu. Jigsaw: Supporting Investigative Analysis through Interactive Visualization. *Information Visualization*, 7(2):118–132, 2008.
- [11] J. J. Thomas and K. A. Cook. *Illuminating the Path*. IEEE Computer Society, 2005.
- [12] M. Wattenberg and F. B. Viégas. The Word Tree, an Interactive Visual Concordance. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1221–1228, 2008.