Multi-faceted Insight through Interoperable Visual Information Analysis Paradigms

Beth Hetzler, Paul Whitney, Lou Martucci, Jim Thomas
Pacific Northwest National Laboratory*

Abstract

To gain insight and understanding of complex information collections, users must be able to visualize and explore many facets of the information. This paper presents several novel visual methods from an information analyst's perspective. We present a sample scenario, using the various methods to gain a variety of insights from a large information collection. We conclude that no single paradigm or visual method is sufficient for many analytical tasks. Often a suite of integrated methods offers a better analytic environment in today's emerging culture of information overload and rapidly changing issues. We also conclude that the interactions among these visual paradigms are equally as important as, if not more important than, the paradigms themselves.

Keywords: information visualization, user scenario, information analysis, document analysis

1 Introduction

Information analysts face significant challenges dealing with the growing amount of information available and how to gain needed insight from it in a timely manner [12]. Much of this information comes in the form of unstructured text. Analysts simply do not have the time to read or skim all these documents. Information retrieval approaches, however efficient, still require users to have some idea of what they are looking for. What is needed is a selection of visual and interaction tools, using a variety of paradigms, to support the various methods employed by analysts. Using multiple paradigms, users can grasp the context of the whole, explore specifics of subsets or particular documents, compare represented themes to pre-existing ontologies, and examine both topical trends and exceptions to those trends. The key idea is not the efficacy of a particular paradigm but rather the interactions between and among them. We believe that this multi-faceted approach provides a superior set of insights to what can be achieved using any of the visualizations alone.

When faced with a complex analysis problem, analysts will likely combine a number of information exploration methods. Shneiderman's [17] taxonomy of tasks and data types lists seven high-level user tasks, such as overview, zoom, filter, and details on demand. In addition to being able to carry out multiple tasks, we contend that insight often requires the ability to carry out one task using a variety of paradigms. For example, getting an overview of the information can mean many things, including

- getting a sense of the major themes in the information and how strongly they are present
- understanding how themes relate to one another
- seeing how the information relates to another information collection or to a standard ontology
- getting a summary of the attributes of the information, such as source, date, document type, etc.

This paper overviews a variety of strategies employed by analysts and presents a selection of visualization tools that were designed to support those strategies. It then walks through a sample information exploration scenario. The goal is not to represent how all analysts work; different users, different information collections, and different analysis problems may certainly lead to very different approaches. The goal is to provide an example and use it as one method for providing evaluating the prototype tools.

2 Tools to meet a variety of analysis needs

Information analysts will have varied goals, approaches, and interaction needs. Analyst goals may vary from the specific (e.g., "what is this government leader likely to do tomorrow?") to the general (e.g., "what unexpected things can I find in this document collection?"). The first of these questions can be construed as a data comparison question: find how the government leader responded to situations similar to the current situation. The second of these questions can be construed as a summarization question: find a description of the typical situation and determine the unusual parts with respect to the typical.

The tools described in this paper address both types of objectives. Collectively, the tools provide means for sum-

* P.O Box 999
Richland, WA 99352

E-mail: beth.hetzler@pnl.gov
Phone: (509) 375-6690
Fax: (509) 375-6703

The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC06-76RLO 1830.

marizing, zooming, extracting detail, filtering, and subsetting to be accomplished. SPIRE (Spatial Paradigm for Information Retrieval and Exploration) projects information into a document space, and the THEMESCAPE™ view shows the distribution of topics on this space. The Cosmic Tumbleweed provides a "dual" view: documents are mapped into a topic space. A commonality between SPIRE and the Cosmic Tumbleweed is that the base spaces depend on the document collection that is currently being examined. The 32D Hypercube view also maps documents into a concept space; however, the concept space can be the same across many different data sets. The Starstruck visualization provides a detailed view of a smaller number of objects (documents or document clusters). The next sections describe each of these tools in greater detail.

1.1 Reducing data preparation time - THEMESCAPE

Currently, information analysts spend the majority of their time in tasks such as monitoring information sources, preparing data for analysis, and reporting on results. The time remaining for actual information analysis and interpretation is highly limited; it would be much more desirable to switch the task dominance, with much less time spent in preparation and much more time allowed for interacting with the data [12].

The SPIRE system [2, 18] was explicitly designed to help with this task switch. SPIRE automatically produces a suitable knowledge base of topics (key words) that can be used to distinguish groups within the document collection under analysis. The system then creates n-dimensional vectors characterizing each document with respect to those topics. The document vectors are clustered and projected from n-space into 2-space; the 2D projection is used to create the visual representations. In the THEMESCAPE visualization, themes within the document spaces appear as a relief map of natural terrain, where taller peaks show dominant themes. It is particularly good for helping an analyst jump-start his or her understanding of the collection – it conveys the main themes in a collection and an overall sense of how they are related.

1.2 Giving more explicit meaning to the information space – Cosmic Tumbleweed

One technique an analyst will want to employ is to view the entire information collection from a variety of perspectives. Dimensionality reduction techniques such as multidimensional scaling [16] help determine the 2D layout for the THEMESCAPE view. However, as noted in Brodbeck [1], such projection techniques create a space with no inherent meaning other than the overall information relationships reflected in the collection. Many clients and users requested a visualization where the meaning of a particular location in a projection is more explicit – even when there is nothing in that location.

A prototype visualization, the Cosmic Tumbleweed [11], was developed to show complex relationships among a set of interrelated topics or themes. In this 3-D visualization, the collection's main topics are projected into 3-space and form the vertices of a tumbleweed-like framework (see Figure 1).

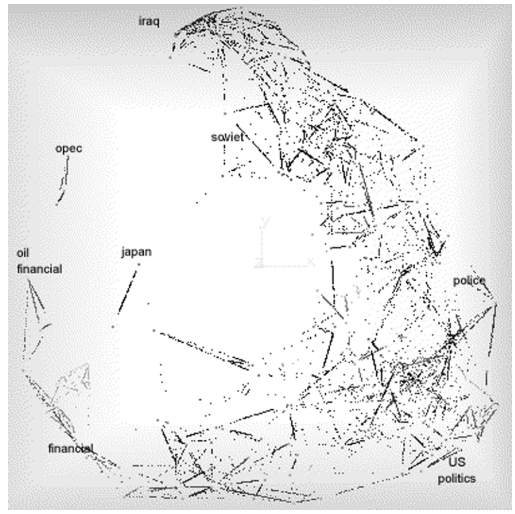

Figure 1: Cosmic Tumbleweed

To simplify this image, most of the topic names have been hidden. In the interactive prototype, the ability to rotate and zoom the 3-D structure substantially increases the appeal and understandability of this visualization.

The first steps in creating the visualization are to identify the topics and to create context vectors for both the topics and the documents. Context vectors for documents can be derived using a multitude of techniques (for example, see Deerwester [3]); context vectors for topics can often be obtained as a natural adjunct of the document context vector. For the Tumbleweed visualization, the technique used must enable calculating the contribution of each topic to the document context vector.

Once the context vectors for the topics are created, principal components analysis is used to transform the topic vectors from n-dimensional coordinates to normalized 3-dimensional coordinates. Finally, each document vector is projected near the topics that are most closely related to the document. The three most dominant topics are identified for each document. For example, dominance can be measured by Euclidean distance between topic vectors and the document vector, with closer topics being considered more dominant.

The document's base position is computed as a linear mixture of the positions of the two most dominant topics. Optionally, the third most dominant topic can be given a

small weight relative to the other weights, so that it perturbs the base position (see Figure 2).

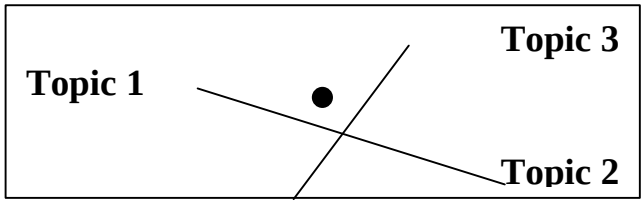

Figure 2: Calculating document location in the Cosmic Tumbleweed

Because the vertices of the Cosmic Tumbleweed represent actual topics, the visualization provides insight into the relationships between those topics, as well as a more explicit mapping for the spatial layout of the visualization. Also, placing documents on or near a line between their primary and secondary topics allows the user to tell at a glance what the topical thrust of a document is without any reading required. For example, near the upper left of Figure 1 is a small isolated triangle of documents with one vertex labeled "opec." The other vertices represent "oil" and "barrels"; documents on a line between opec and oil have those two as their dominant topics.

1.3 Relative vs. absolute space

Both the THEMESCAPE and Cosmic Tumbleweed visualizations are based on relative placement of themes and documents. Different information collections can produce entirely different topical layouts. For certain tasks, analysts may be interested in establishing a standard space that could be used, for instance, to compare a particular collection to a known layout (perhaps determined by a familiar information collection) or to more easily compare two different collections in terms of a reference set.

The 32D Hypercube visualization shows how a collection of documents relates to a selected ordered set of concepts, for example, the MeSH (Medical Subject Headings), the Dewey Decimal System, or an ordered version of the key concepts present in the collection itself [11]. The latter option is used in the example presented here. For each document, we determine which concept from the ordered set of concepts is most strongly represented in that document, which is second most strongly represented, and so on, up to an arbitrary limit of 32 total concepts.

A 3-dimensional visualization is created using the ordered set of concepts for each axis. The x-value for a document is the index of the concept that is most strongly represented in that document; the y- and z-values are the indices of the second and third most strongly represented concepts. Similar visualizations can be shown for any trio of concept strengths desired. Our system provides rotation to help the user examine the visualization. Concentrations of dots indicate where closely related concepts show up strongly in multiple documents (see Figure 3).

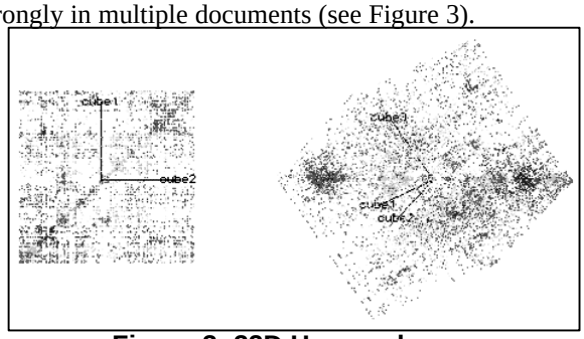

Figure 3: 32D Hypercube

The 32D Hypercube also has appeal when dealing with massive and incremental data sets. Because the vector summary is created with respect to a reference ontology, context vectors for documents can be calculated in parallel and without the need to know the full corpus.

1.4 Viewing explicit theme combinations—Starstruck

An analyst may often interact with combinations of alternative overviews and focus on a particular subset of documents to examine in greater detail. The Starstruck visualization was intended to meet two needs:

- represent differing levels of granularity in information spaces and give access at multiple levels of detail
- give meaning to the shape of clusters in the information space.

In the Starstruck visualization, each document is represented as a planar starburst of topic rays. The ray length indicates the strength of the corresponding topic or theme in that document. The order of the theme rays from the 12 o'clock position clockwise is the same for each document and is determined by a mathematical correlation of the themes within the collection. This means that themes that tend to occur together in the document collection will be near each other in the planar starburst, making it much easier to identify areas of theme strength and theme weakness. Figure 4 shows several individual documents. The similarity between the first and third stars represents the similar topics in the corresponding documents.

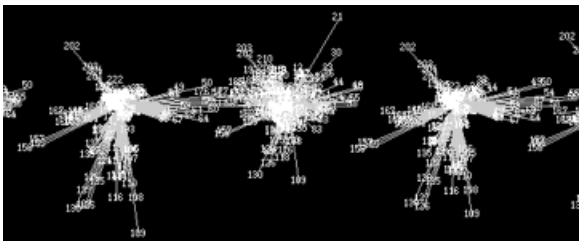

Figure 4: Individual document stars in Starstruck

The Starstruck visualization is intended to show thematic patterns over a group or cluster of documents. The individual planar starbursts for the group are aligned along an invisible rod in space. When the view is changed to look along that rod, all the starbursts are shown superimposed (see Figure 5) and the shape of the group representation gives an impression of the thematic distribution within that group. Simply by varying the view rotation, the user can switch from seeing the thematic strengths of the group as a whole to panning over individual documents.

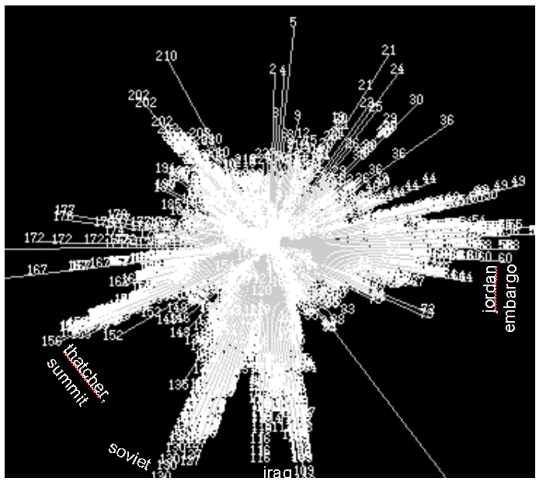

Figure 5: Head-on view of multiple document stars in Starstruck

At first glance, the Starstruck view may seem too complex to be useful. This is where the importance of the correlation ordering of topics and the variable detail approach come into play. The intent is that the overall shape (whether for individual stars or for a superimposed group) will convey a general impression of which theme sets are strongly represented. A strong fan shape of themes between the 1 o'clock and 3 o'clock positions should be distinguishable from a narrower spike near the 8 o'clock position. From a distance, the theme numbers serve only to augment the shape by adding almost a texture, suggesting the variance in theme strength. For more detail, a user can zoom in to identify the actual numbers.

The current Starstruck prototype allows several options for showing theme strength, which can be combined or used separately:

- Both ray length and brightness can be fixed or can vary according to theme strength.
- The endpoints around a starburst can be connected, so that each document forms a spider graph.
- Theme ray labels can be hidden or shown only as vertices.
- Color values can be specified for the rays and the spider graph lines.

3 A sample analysis scenario

The Cosmic Tumbleweed, 32D Hypercube, and Starstruck visualization paradigms are in early development. An obvious next step is to evaluate the merit of these presentation approaches. As one test, we decided to use these three, along with the visualizations available in SPIRE, in a test analysis. One goal was to give us a measure of the possible usefulness of each tool and of the set as a whole. A second goal was to obtain ideas for improvements, based on usage for such a scenario.

Participants included both visualization and analysis experts. The data analyzed are from the 1990 Associated Press data on the TREC5 distribution disks. These newswire data are typical of the information that many analysts encounter daily. The analysis objectives were to ascertain and examine broad trends and anomalies in these data. Early on, a detected anomaly in the data moved the focus of the analysis to documents produced during the time period surrounding the early stages of the Iraq-Kuwait conflict that led to Desert Storm. This section describes some of our findings in the form of a scenario.

The AP collection from 1990 covers the entire year and includes over 105,000 documents. The THEMESCAPE visualization is particularly good for quickly understanding the major theme groupings and their relative dominance within the information collection. Therefore, the analyst in our scenario begins by loading the 1990 data into the SPIRE tool suite.

Faced with universal challenges of time and computation tradeoffs, one of the first strategies an analyst may use is to quickly identify a likely subset of fertile information. The analyst uses SPIRE's Time-Slicer capability to step through the months of the year, watching the THEMESCAPE change to reflect the documents written in each time period. This strategy reveals a major change in the newswire data between July and August. The analyst selects the subset of documents covering July, August, and September (a total of 25,386) for further analysis; a similar change is still seen. Figure 6 shows a region of major change in these months. In July, there are significant peaks labeled "soviet" and "billion." In August and September, these two are not apparent, and a strong "iraq" peak appears.

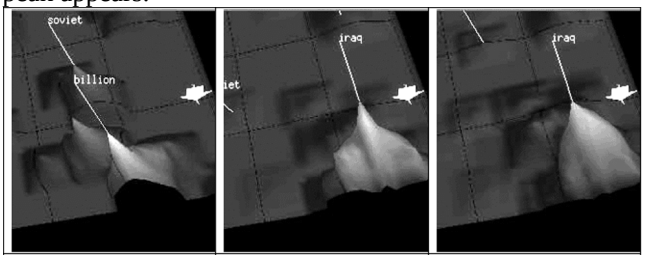

Figure 6: THEMESCAPE detail changes over time

The arrows in the figure show a common location in the information space, as the data are compared from month to month. The differences reflect the changes in content of news articles related to Iraq surrounding the time of the Iraqi invasion of Kuwait (Thursday, August 2, 1990). Examining the key terms and individual documents in the region, the analyst determines that the topic concentration in August is from documents describing Iraqi complaints regarding Kuwaiti oil production. The emerging peaks correspond with topics that describe Iraq's invasion of Kuwait. The proximity of the "iraq" peak in August to the "soviet" peak in July raises questions in the analyst's mind, particularly the following:

- Are there other, related changes in the news during that time?
- What are the secondary topics related to those two topics?
- Are there connections between the "soviet" and "iraq" themes?

The full 3-month document collection was also examined using the Cosmic Tumbleweed. The Tumbleweed provides a view of the relationships of topics in the data. The Cosmic Tumbleweed broadly divided the news collection into financial and political topics. Figure 7 shows the portion of the tumbleweed relating to Iraq. Within the political topics, "iraq" and associated topics are close to but not directly connected with the "soviet" related topics. This tells the analyst that the collection does not contain many documents for which "soviet" and "iraq" were the two strongest topics. If there are connections, they are more subtle.

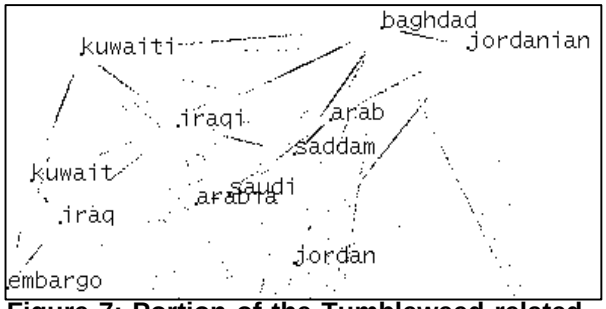

Figure 7: Portion of the Tumbleweed related to the topic "iraq"

Next, more detailed views of the data are constructed by further reducing the document collection to those documents containing the "iraq" or "soviet" related dominant themes. This smaller collection of documents is examined via the 32D Hypercube and the Starstruck visualizations.

The Hypercube representation explicitly orders the topics for each document. Then the analyst can view combinations of particular strength, such as viewing the distribution of documents' first, second, and third most dominant topics (shown in Figure 3). One of the most useful aspects of this is a clear representation of what is not covered, often a particularly useful piece of information for analysts. To emphasize this aspect, the analyst in our scenario examined the dominant topics in the 3-month subset when viewed against time (see Figure 8).

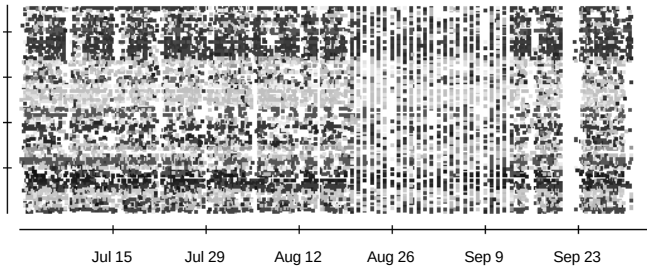

Figure 8: Dominant Topic vs. Time

This view provides a different perspective on the data than the THEMESCAPE sequence shown above. It is apparent in this view that in general, all of the topics are covered across the time period; there are no large horizontal holes. However, there are some gaps in the time coverage that the analyst examines in greater detail. The linear pattern around early September and the vertical gap near September 23 seem to be related to the collection in general—perhaps caused by changes in the date/time representation or missing data. The collection also shows an interesting periodic pattern. Each weekend appears to be light on news; the topical coverage is less over the weekends, as shown by the sparser vertical sections by July 8, July 15, July 22, etc. There are also some smaller gaps in specific topic areas, which might be worth investigating further.

With a representation like the Hypercube's, both identifying topic combinations of interest (here "soviet" and "iraq") and selecting documents related to those topics are straightforward. The analyst creates a smaller subset containing documents for which either "iraq" or "soviet" were relatively strong topics, and makes a similar view plotting the dominant topic versus time for the selected documents (see Figure 9 on the next page). This view shows that, within this subset, articles concerning superpower interactions die out after early August. The Russia/Gorbachev thread dies out along the same time-scale. The difference between the two Iraq-related threads is that the bottom tends to emphasize the U.S. more. An unexplained gap is found in the Iraq/embargo threads near July 29, which was not reflected in the overall data set. Such gaps may be highly interesting to analysts.

To get a broad sense of the other themes within this subset, the analyst examines the iraq/soviet documents in the Starstruck visualization; this visualization is shown back in Figure 5. The stronger themes are labeled in the figure and represent the following:

- discussion of the embargo and Jordan's role
- a trio of themes relating to a summit meeting between the superpowers
- the "soviet" and "iraq" themes themselves.

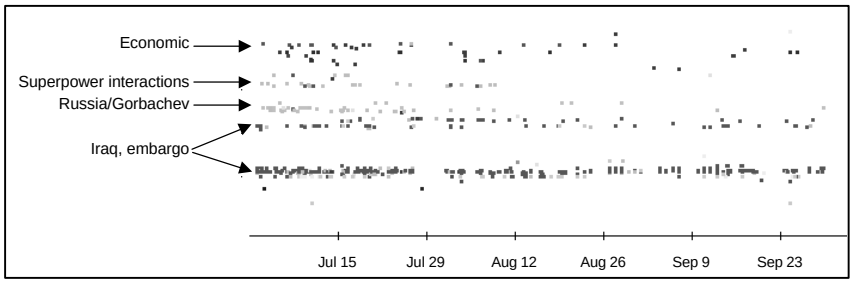

Figure 9: Distribution of selected documents over time

To examine correspondences between the visualizations, we also loaded data for three of the visualizations into DataDesk, a commercial data analysis package. Figure 10 shows an example. On the left is a SPIRE Galaxy view, showing dots for individual documents. Recall that in the SPIRE visualizations, proximity indicates relatedness, but location along the axes has no inherent meaning. On the right is a Hypercube view where each axis is a sorted list of themes, giving an explicit meaning to location. In this figure, a portion of the Galaxy visualization has been selected, representing documents that were projected close together, because of connections between their thematic contents. The same documents are also highlighted in the Hypercube view, indicating specific combinations of dominant themes represented. Explorations such as this gave us a much better sense of the relationships among the visualizations, a good orientation to the various layouts, and a picture of how the thematic combinations in the SPIRE and Tumbleweed views related to the strictly ordered themes in Hypercube.

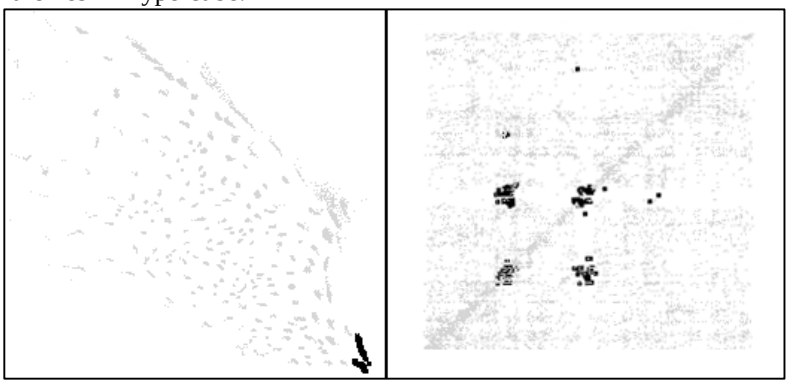

Figure 10: Selected documents highlighted in Galaxy display, with selection reflected in Hypercube display

4 Discussion

This section summarizes the results of the analysis both in terms of information useful to our scenario analyst and feedback useful to the tool design and development team.

By examining the AP data with the combined visualization tools, the analyst discovered broad thematic content and relationships between the themes. Next, the analyst observed changes in topic between July and August. An intriguing relationship between "soviet" and "iraq" topics was observed. Focusing in on these documents with the combined tools, the analyst was able to detect relationships between themes and examine thematic changes over time.

Combining visualization tools revealed several specific items of interest in this analysis scenario. The following are key results of the type that an information analyst might focus on:

- Abrupt changes in news topics occur from July to August of 1990. Detailed examination of the data during that time shows that the changes in the data are associated with Iraq invasion of Kuwait.
- There is an unusually wide duration between Iraqi related stories surrounding the time of the invasion. The abrupt lack of data related to a particular topic is potentially informative to an analyst.

Additionally, the following were observed:

- Broad similarities are seen in the AP news across 1990: there are large collections of economic/ financial news and large collections of political news.
- There is a weekly periodicity in new-story volume: fewer news stories are time-stamped with weekend dates.
- Themes about super-power interactions die off as lead themes. Likely explanations include that the particular event reached a natural conclusion (e.g., negotiations completed) or that the news was subsumed in articles with other high themes.

The experiment of using the various tools to analyze this collection in partnership with information analysts helped the team reach a better understanding of the analysts' needs and priorities. For example, while the usefulness of flexible and responsive interactivity is commonly assumed, this kind of exercise demonstrates the point extremely well. The experiment also suggested a number of specific interaction features that would be helpful. For example, it turned out to be extremely useful to examine Hypercube coordinates and Galaxy coordinates

vs. time. More explicit support for viewing time would be useful in each of the visualizations. Methods for better highlighting differences between two visualizations would also be helpful, such as the difference between two months' THEMESCAPES. Also, many visualization-specific interactions would be useful, such as the ability to pull on one of the Cosmic Tumbleweed vertices and deform the framework to explore the connections.

5 Related work

The individual tools described in this paper compare broadly with other published tools as follows:

Cosmic Tumbleweed—The potential for viewing words and documents in the same space is available in latent semantic indexing [3] and in the tools described in Lin [10] and more recently in Rushall [15]. However, the underlying methodology and the visualization used in the Cosmic Tumbleweed are intrinsically different from these.

32D Hypercube—The tools described in Lin [10] and Rushall [15] include the ability to represent documents with respect to a few concepts. The Cat-a-Cone [6] interface lets users search and browse a hierarchical ontology. However, the authors are not aware of published visualizations of large document collections with respect to an ontology.

Starstruck—The ideas used in the Starstruck are strongly related to glyphs and parallel coordinate representations; there are numerous references for these types of displays applied to diverse data types: Seber [16], Wong [19], and Inselberg [9] provide general, historic, and a recent specific reference, respectively. The work of Pickett [13] is also related, in that it uses a small glyph to represent individual items, while a combination of many shows a very different global pattern.

THEMESCAPE—The THEMESCAPE shows a concentration of topics on the same base space as the 2D space on which SPIRE projects the documents. MDS-like projection plays a key role in this representation; see Ebert [4], Fairchild [5], and Brodbeck [1] for other applications of MDS in related settings.

Segmenting and visualizing the data by time (or other variables) is not new; indeed, it is a standard and extremely valuable concept in data analysis. This capability seems necessary for an exploratory information analysis system.

Finally, other systems that integrate across different views of documents (as data) are described in Rushall [15] and Risch [14]. A systematic approach to integrating multiple visualizations is described in MULTISPREAD [7].

6 Conclusions and future work

In this paper, we proposed an approach of multiple interoperable visualizations. We described a range of information analyst strategies and presented a complementary set of visualizations, together with some traditional techniques, targeted at supporting these strategies. Finally, we presented an example scenario illustrating this approach.

The scenario demonstrated the usefulness of interacting among multiple visual paradigms to gain a variety of insights from a document collection. No single visualization, paradigm, or method would have been sufficient for this kind of exploration. Also critical to this analysis were interactions *across* tools, such as the ability to interactively see the placement of document groups across multiple visualizations or the ability to select subsets in one visualization and view them in another.

In future work, we will focus on some of the lessons learned from this experiment. Speed, especially with larger collections, is a critical issue. In some cases, simply revisiting the implementation strategy may improve the speed significantly. In other cases, such as the Hypercube, we may look at alternative methods to scale the collection while reducing the complexity of the visualization. Example techniques include

- collapsing the ontology represented along the axes into higher-level terms
- smoothing the visualization
- using visual summary techniques, such as those suggested in Huang [8].

For this experiment, the analysis and comparison among the different visualizations were performed by a combination of manual data transfers and importing visualizations into a standard data analysis package. We are beginning work on a research framework for combining these and other prototype visualizations together so that such experiments and multi-visualization analyses are more easily done. The goal is to put an integrated combination of such prototype visualizations into the hands of users so that they can evaluate the various visualizations and paradigms in the context of their work.

Acknowledgements

We wish to thank the Defense Advanced Research Projects Agency (DARPA) for funding this work; we particularly acknowledge John Silva, MD (Col. USAF) for his encouragement and guidance in our visualization work. We also wish to thank the CIA's Office of Research and Development and the U.S. Department of Energy's NN-60 for their assistance in initial and continuing development of the SPIRE system. We gratefully acknowledge the rest of the Battelle and Pacific Northwest National Laboratory

Information Visualization Team, including Gus Calapristi, Kris Cook, Vern Crow, Sharon Eaton, Shelly Harris, Sue Havre, Tonya Martin, Dennis McQuerry, Renie McVeety, Nancy Miller, Grant Nakamura, and Pak Wong.

References

1. Dominique Brodbeck, Matthew Chalmers, Aran Lunzer, Pamela Cotture. Domesticating Bead: Adapting an Information Visualization System to a Financial Institution. In *Proceedings of IEEE Symposium on Information Visualization*, InfoVis ’97, 1997. IEEE.

2. Vern Crow, Kelly Pennock, Marc Pottier, Anne Schur, James J. Thomas, James Wise, David Lantrip, T. Fiegel, C. Struble, J. York. Multidimensional Visualization & Browsing for Intelligence Analysis, Gviz ’94: Graphics & Visualization Conference, Richland, WA, Sept. 1994.

3. S. Deerwester, S.T. Dumais, T.K. Landauer, G.W. Furnas, R.A. Harshman. Indexing by Latent Semantic Analysis. *Journal of the Society for Information Science*, 41(6):391-407, 1990.

4. David Ebert, Chris Shaw, Amen Zwa, Ethan Miller, D. Aaron Roberts. Minimally Immersive Interactive Volumetric Information Visualization. In *Proceedings of IEEE Symposium on Information Visualization*, InfoVis ’96. IEEE.

5. K. Fairchild, S. Poltrock, G. Furnas. SemNet: Three-Dimensional Graphic Representation of Large Knowledge Bases. In R. Guindon (Ed.), *Cognitive Science and Its Applications for Human-Computer Interaction*. Erlbaum, 1988.

6. M.A. Hearst and C. Karadi. Cat-a-Cone: An Interactive Interface for Specifying Searches and Viewing Retrieval Results Using a Large Category Hierarchy. In *Proceedings of ACM SIGIR* ’97, Philadelphia, PA, July 27-31, 1997. ACM Press.

7. Ed Huai-hsin, Phillip Barry, John Riedl, Joseph Konstan. A Spreadsheet Approach to Information Visualization. In *Proceedings of IEEE Symposium on Information Visualization, InfoVis ’97*. IEEE.

8. Chisheng Huang, J.A. McDonald and W. Stuetzle. Variable Resolution Bivariate Plots. *Journal of Computational and Graphical Statistics*. 6(4):383-396, 1997.

9. Alfred Inselberg. Multidimensional Detective. In *Proceedings of IEEE Symposium on Information Visualization*, InfoVis ’97. IEEE.

10. X. Lin, D. Soergel, G. Marchionini. A Self-Organizing Semantic Map for Information Retrieval, in Proc. SIGIR'91, published as a special issue of SIGIR Forum, ACM Press, pp. 262-269, Oct. 1991.

11. Nancy Miller, Beth Hetzler, Grant Nakamura, Paul Whitney. The Need for Metrics in Visual Information Analysis, in *Workshop on New Paradigms in Information Visualization and Manipulation*, Nov. 1997, Las Vegas, NV.

12. P1000 Science and Technology for Information Visualization. Sept. 1996.

13. Ronald M. Pickett, Georges G. Grinstein. Iconographs Displays for Visualizing Multidimensional Data, in *Proceedings IEEE Conference on Systems, Man, and Cybernetics*, PRC, Beijing & Shenyang, May 1988, pp. 514–519.

14. J.S. Risch, D.B. Rex, S.T. Dowson, T.B. Walters, R.A. May, BD Moon. The STARLIGHT Information Visualization System. In *Proceedings International Conference on Information Visualization*, IV97, London. IEEE.

15. David Rushall, Marc Ilgen. DEPICT: Documents Evaluated as Pictures, Visualizing Information Using Context Vectors and Self-Organizing Maps. In *Proceedings of IEEE Symposium on Information Visualization*, InfoVis ’96. IEEE, 1996.

16. G.A.F. Seber. *Multivariate Observations*. John Wiley & Sons, 1984.

17. Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, in *Proceedings for IEEE Symposium on Visual Languages*, pp. 336-343, Sept. 3-6, 1996. IEEE.

18. James A. Wise, James J. Thomas, Kelly Pennock, David Lantrip, Marc Pottier, Anne Schur, Vern Crow. Visualizing the Non-Visual: Spatial Analysis & Interaction with Information from Text Documents, in *Proceedings of IEEE ’95 Information Visualization*, pp. 51-58, Atlanta, GA, Oct. 1995. IEEE.

19. Pak Chung Wong, R. Daniel Bergeron. 30 Years of Multidimensional Multivariate Visualization. In Gregory M. Nielson, Heinrich Mueller, Hans Hagen, editors, *Scientific Visualization: Overviews, Methodologies & Techniques*, Los Alamitos, CA, 1997. IEEE.