Supporting hypothesis formation during asynchronous collaboration for visual analytics for text

Ankit Gupta
Simon Fraser University
Surrey, Canada
aga53@sfu.ca

Carman Neustaedter
Simon Fraser University
Surrey, Canada
carman_neustaedter@sfu.ca

ABSTRACT
Visual Analytics identifies the need to support analytical reasoning using interactive visual interfaces. As datasets get large, this analytical reasoning process is unlikely to be performed by a single analyst. To tackle the growing size of datasets, analysts collaborate as a team, often asynchronously to make sense of the data. They work independently on subsets of a large document collection to find new information. An important aspect of the analytical reasoning is hypothesis formation which includes proposing and evaluating hypotheses. The analysts collaboratively form hypotheses based on the information they find while working independently. Existing visual analytics tools do not provide support for hypothesis forming during asynchronous collaboration. In this paper, we present a visual analytics tool that supports hypothesis formation during asynchronous visual analytics for textual data.

Author Keywords
Visual Analytics, Hypothesis Formation, Text Analysis, Asynchronous Collaboration

INTRODUCTION
Intelligence analysts are constantly required to make sense of large document collections ranging from intelligence reports and telephone intercepts to blogs and newspaper articles. The analysts extract new information from text documents. This new information is put together to support or reject existing hypotheses or form new ones. However, this activity of making sense of text is not limited to the intelligence domain. The process of making sense of large document collections is similar in several areas such as journalism and digital humanities. Since the number of documents to be analyzed is large, the sense-making activity is often collaborative. In order to support effective analysis of large datasets, the Visual Analytics agenda [15] identifies "support for collaboration" as one of primary components of visual analytics systems. Collaborative visual analytics allows analysts to share their work as well as allows for better sense-making. Analysts working collaboratively, individually contribute new information. The new information found by analysts is then used to collaboratively support or reject existing hypotheses or form new hypotheses.

Several visual analytics (VA) tools have been implemented to support individual as well as collaborative sensemaking. Such tools include Jigsaw [14], CZSaw [9] and IN-Spire™. Jigsaw [14] and CZSaw[9] provide support for entity-based visual analytics where the analysts attempts at finding relationship between different named entities like person, location, organization etc. IN-Spire™[1], on the other hand provides interactive visualization based on document clustering and word frequencies. As mentioned, none of these tools are designed to support collaboration.

Research has also been done in the field of collaborative visual analytics. However, the efforts in this area have largely focused on collocated and synchronous remote collaboration. Cambiera [7] and [10] are tools designed for visual analytics in a collocated setting. Sense.us [6] is another tool designed to support visual analytics for text during asynchronous collaboration. However, none of the existing tools designed for collaboration provide support for analytical reasoning, which is another important requirement for visual analytics tools [15].

Another set of tools have been designed to support the reasoning process during visual analytics. For example, Shrinivasan et al. [13] implement a knowledge view in their system and Sanfillipino et al. [11] implement a hypothesis space in IN-Spire™. However, none of these tools are designed to support collaboration.

While both collaboration and analytical reasoning are important parts of the sensemaking process, existing tools have been designed to support either collaboration or the reasoning process. We recognized this gap and designed a tool that supports both collaboration and sensemaking.

The rest of the paper is organized as follows. Firs, we provide details about the research done in the field of visual analytics that focuses analytical reasoning and collaboration. Next, we discuss the design principles. We discuss the principles that we followed while designing our system. Next we discuss how our design evolved and how our tool can be used for sensemaking during asynchronous collaboration. Finally, we discuss the strengths and weaknesses of our design in section.

RELATED WORK
Several visual analytics systems have been created and studied in the past. The focus of these tools has largely been one of the recommendations provided by Thomas and Cook [15]. These include support for collaboration, reasoning, synthesis etc.In this paper, we present a system that focuses on asynchronous collaboration and support for analytical reasoning in the context of entity based text analysis. With this focus, we now discuss existing visual analytics tools that have pro-

vided support for collaboration, reasoning and entity- based text analysis.

Collaboration

While most of the visual analytics tools that have been designed for supporting collaboration have focused either on collocated collaboration or synchronous remote collaboration [16, 3], it is important to discuss these tools as we derive certain design principles from these tools and try to overcome some of their limitations in asynchronous settings.

Hajizadeh et al.[5] studied brushing techniques for providing awareness in a synchronous remote setting on tabular data. They compared three brushing techniques (brushing and linking, selection and persistent selection) for providing awareness. In their research, they identified awareness as the ability of collaborators to understand the brushing actions taken by their remote collaborators. They studied these techniques using a collaborative visualization of tabular data where two collaborators shared a visualization workspace. The results of the study indicated that persistent selections in which users saw their collaborators previous as well as current selection provided most awareness among the three techniques. While this research did not study asynchronous collaboration, its results indicate that persistent selections cause minimal interference. Since, we want analysts to work independently, we do not provide real time brushing and linking when multiple analysts are annotating the same document.

In another work [8], the researchers studied the use of a collaborative awareness technique called ”collaborative brushing and linking” in which the collaborators are aware of each other’s selection via brushing and linking. In their exploratory study, they studied a system called Cambiera [7]. They found that Cambiera’s implementation of collaborative brushing and linking was quite effective in providing awareness in a collocated collaborative environment. The techniques implemented is specific to collocated settings and does not apply in our case. However, the study highlights the importance of awareness during text analysis, which we consider as an important design principle of our tool.

Sense.us [6] is another tool for collaborative visualization. The tool is designed to support asynchronous collaboration during analysis of tabular data. The authors of the tool studied the system to provide design recommendations for encouraging social interaction in asynchronous collaborative visualization. Their goal is different from ours. In our system, we do not consider social interaction as an important factor as our tool is designed for professional analysts for whom analysis is part of their job. Yet, some of the recommendations provided by the authors is quite useful. They recognize awareness and provenance[1] as important goals of asynchronous collaborative visualization. We base our design principles on these findings.

We discussed several tools in this section that provide support for collaboration. Some of them do not provide support for asynchronous collaboration. The one that do provide support for asynchronous collaboration do not provide support for analytical reasoning.

Support for Sensemaking

There are several visual analytics tools that support the reasoning process explicitly or implicitly. In this discussion, we do not review systems that provide Artificial Intelligence or Machine Learning enabled support for automated reasoning. We limit our discussion to systems that provide support reasoning by human analysts[13, **?**].

Shrinivasan et al. [13] implemented a prototype called Aruvi, which was designed for individual analysis. In Aruvi, the authors implemented a view called knowledge view. The knowledge view in Aruvi allowed an analyst to create notes and then organize them into groups and drawing edges between them to connect related notes. The result was a graph of notes. This was the initial design that we created for the hypothesis view. However, we soon realized that this representation is not suitable for collaborative reasoning. We present our design choice later in this paper.

In another work Sanfillipino et. al. implemented a hypothesis space in IN- Spire™[1]. The hypothesis space was designed to allow an analyst to form hypotheses. The visualization was similar to argument maps. It had a hypothesis and the analyst added related evidence to a hypothesis. While our hypothesis view is different from the hypothesis space developed by Sanfillipino et al., we use their implementation as the basis of our hypothesis view.

Entity-based Analysis for Text

Entity-based analysis is a kind of text analytics in which the goal of the analyst is to find important entities in a document collection and to find relationship between different named entities like person, location, organization etc. Our system is designed to support entity-based text analytics.

Entity Workspace [2] was one of first VA tools to be developed to support entity-based analysis. Entity Workspace provided automatic extraction of named entities using Natural Language Processing (NLP) algorithms. It also allowed the analysts to find important entities and find entities related to any given entity. The authors showed how entity-extraction allowed the analysts in finding documents that are important for analysis. In our tool, we consider support for entity-based collaboration to be important for the same reasons.

Another important tool is Jigsaw [14] which also provided entity-based collaboration. However, Jigsaw added several visualizations for visualization of relationships among entities as well as several NLP algorithms including sentiment analysis. Jigsaw allowed an analyst to look at the importance of an entity (based on its frequency) using a list view in which entity names were displayed with a bar to indicate the frequency of the entity. In addition, the list view also allowed the analyst is looking at related entities. The related entities were found based on the concept of bibliographic coupling. According to this concept, two entities are considered related if they appear in at least one document. A limitation of this

[1]The authors do not use this word. However, their notion of doubly linked discussions is similar to provenance.

approach is that when visualizing important and related entities in a list view, the context of the entities is lost. In other words, the analyst no longer views the document where an entity has been mentioned. We realize this limitation and attempt to overcome it in our design. Instead of using a separate list view, we show important entities using opacity of highlights within the document itself. We discuss it in more detail in a later section.

CZSaw [9], in addition to providing entity-based collaboration and entity-based visualizations provided by Jigsaw, provided a history mechanism, an editable script and a dependency graph. The history view allowed the analyst to track his actions and jump back to any previous state. CZSaw also captured user actions and inputs in the form of an editable script called CZScript. It allowed the analyst to revise his analysis by changing the inputs at some previous point in the script. On changing the inputs in the script, CZSaw would perform the analysis with new inputs without requiring the analyst to manually performing the intermediate steps. A dependency graph provided the analyst with a visualization of the several dependencies among different components of the analysis. While our system does not implement a reusable history, we do provide information about creation date of artifacts. A history mechanism and replayable analysis is a design principle that we intent to implement as part of our future work.

Existing tools have been designed either to support collaboration, particularly in collocated and synchronous remote settings or to support the reasoning process. However, none of the existing tools support both. This presents us with an opportunity for developing a system that does both and to study collaboration during hypothesis formation using this system.

DESIGN PRINCIPLES

The main goal of our system is to provide collaboration during the analytical reasoning process during text analysis. While designing the system, we had four major design principles: *Awareness*, *Collaborative Hypothesis Formation & Evaluation*, *Provenance* and *Entity-based Analytics*.

Awareness

Awareness of collaborators' activities is an important element of collaboration. Information about awareness helps an analyst in assessing what has been done and where more effort needs to be put. The term 'awareness' is used in different contexts to mean different things [12].It is therefore important to explain what we mean by awareness.

When we refer to awareness in the context of our tool, we refer to the ability for a collaborator in a distributed team of analysts to know what has happened since they last logged. In addition, the analysts should be aware of when the new information has been added and who are the contributors of the new information.

Collaborative Hypothesis Formation & Evaluation

Since most visual analytics tools do not support the process of hypothesis formation and evaluation, this was another important requirement of our tool. In asynchronous collaborative settings, analysts would work individually to find new information. This new information is used by all team members to validate or reject a hypothesis. We wanted to build a system in which analysts can work independently to find new evidence and use information found by themselves as well as other team members to create hypotheses and add supporting and opposing evidence for it.

Provenance

As discussed in the previous section, we want to support collaborative hypothesis formation. When analysts look at new information created by other analysts, they might often feel the need to look at the source of the evidence. We assume this based on personal experience while working on VAST challenges and the recommendations given by Thomas and Cook [15]. We designed our tool keeping this requirement as another important principle in our design. When looking at a piece of information, we want to allow the analyst to jump back to the source of the information as well as know about who contributed the new piece of information.

Entity-based Analytics

Another important design principle that we identified is the need to support entity-based analytics. In entity-based analytics, the analysts are interested in finding important entities and possibly the relationship between the entities. It is therefore important that we provide the analysts with the entities within a document and their importance. While this requirement does not contribute in supporting collaboration, yet, it is an important requirement given the context of intelligence analysis that we have focused our system on.

DESIGN EVOLUTION

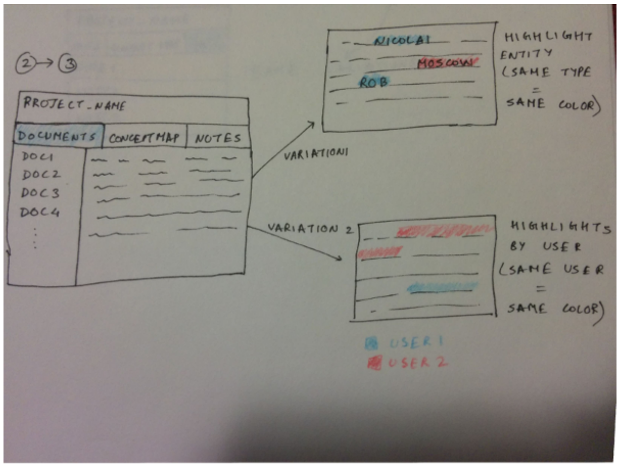
We identified two primary visual spaces for supporting asynchronous collaboration during hypothesis formation during entity-based text analytics. First, a document view that allows an analyst to read document, view the entities contained within the documents and attach notes. Second, a visual space for forming hypothesis and evaluating them. We now discuss the design evolution of each of the views.

Figure 1 shows the initial design of our system. Figure 1a shows the document view with a list of documents and a text area for showing text of a selected document and the contained entities highlighted. Figure 1b shows the initial design of the Hypothesis view.

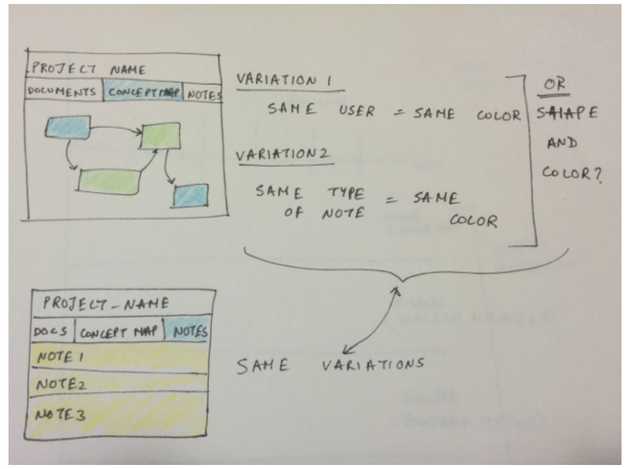
After the initial sketches, we evaluated the designs using paper prototypes. Once, we finished evaluation using paper prototypes, we implemented the system using *Ruby on Rails* web development framework and iteratively modified the design. We now discuss the design evolution for each of the two views.

Document View

We started with two primary requirements for the document view. We wanted to show the named entities in a document to the analyst and show who added what note and where in a document. It was a challenge to show both these pieces of information at the same. In the initial design, we considered



(a) Document View



(b) Hypothesis View

Figure 1: First Iteration

providing the analyst with the option to choose the visual encoding – whether he/she wants to highlight the text based on the entities or highlight the text that is being annotated. In the former case, the color of the text will represent the entity type. In the latter case, the color indicated the collaborator who annotated the text. However, we were able to show both piece of information in the same view as we will discuss later in this section.

While evaluating our design using paper prototype, we realized the need for text-search to find documents containing a keyword or an entity, as well as the need to filter documents by their title. We then implemented the document view including the new requirements we identified during paper prototype evaluation.

When evaluating our design against the design principles, we realized that the current design was lacking in two areas. It did not provide sufficient awareness. We wanted the analysts to know by just a quick glance, the documents that have been annotated heavily as well as the ones that still required attention. Second, the current design showed all entities to be of equal importance and therefore it was no better than no entity-highlights at all. We added these two metrics into the design.

The final document view is shown in figure 2. It consists of a list of documents and a search box to filter the list by their title (figure 2a). In the document list, the title of the documents are prefixed with a small colored rectangle (see figure 2e). In the figure the documents **CIA03.txt** and **ArmyCID01.txt** have red rectangles as a prefix. The color of the rectangle indicates the number of annotations that were added to the document by any analyst such that more saturation indicates more annotations. This helps the analyst quickly identify the documents that have contain a large number of annotations and those that do not. In this case **CIA03.txt** contain more notes than **ArmyCID01.txt**. An analyst can use this information to decide that he needs to annotate other unannotated documents or he/she may decide to take a look at the documents that are highly annotated to find out the important pieces of information in those documents.

Figure 2b shows the content area of the document view. The content area highlights the entities within the text. The entities are extracted using Named Entity Recognition algorithm [4]. When highlighting, the entities are color-coded. Each color represents a different type of named entity such as location, organization etc. We use opacity of the color to represent the importance of an entity within a document collection. An darker highlight (i.e. higher opacity) indicates that the entity has been mentioned by more documents as compared to an entity with lower opacity. This helps the analyst in searching the document collection based on entities that are important. In the content view an analyst can select a piece of text and add a note to it. When the analyst adds a note, the annotated text gets underlined with the color corresponding to the analyst. Here one can see that the word "manpads" was annotated by Joe (joe@example.com). In addition, the analyst can find all documents containing a highlighted entity by just clicking on that entity. This helps the analyst in filtering the documents that are of interest at a given time. While this feature does not help collaboration, it is an important part of the process as it helps the analyst in quickly filtering the documents and what new piece of information can he/she get about a particular entity of interest.

Figure 2c show a legend that shows the analysts collaborating on the project and the notes that are added to the current document and the list of notes that have been added corresponding to the selected document. In figure 2c, one can see that there is only one annotation "Manpads theft indicates to a possible air attack." created by Joe.

The notes are color coded by their authors. This information is important for an analyst to decide the importance of a note. Different analysts have expertise in different areas. This means that an analyst can look at a note and decide whether or not to trust an annotation based on the author or to alternatively ask for some clarification by asking for more details

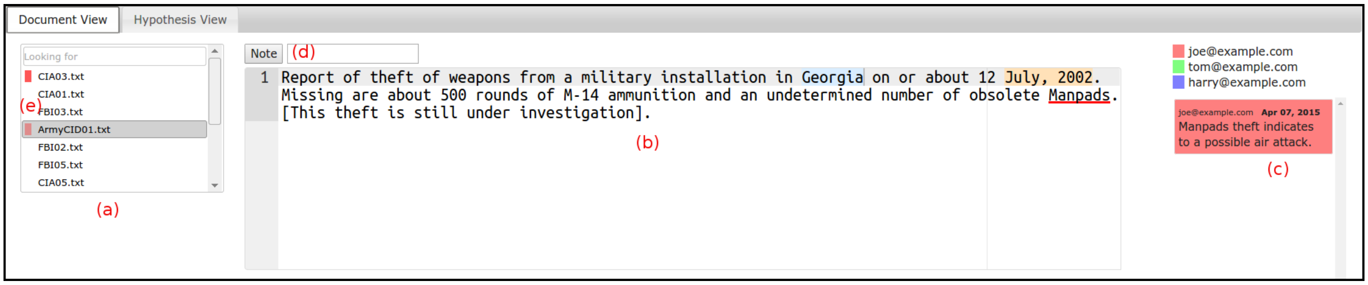


Figure 2: The Document View. (a)Document List; (b) Content View; (c) Legend and Notes

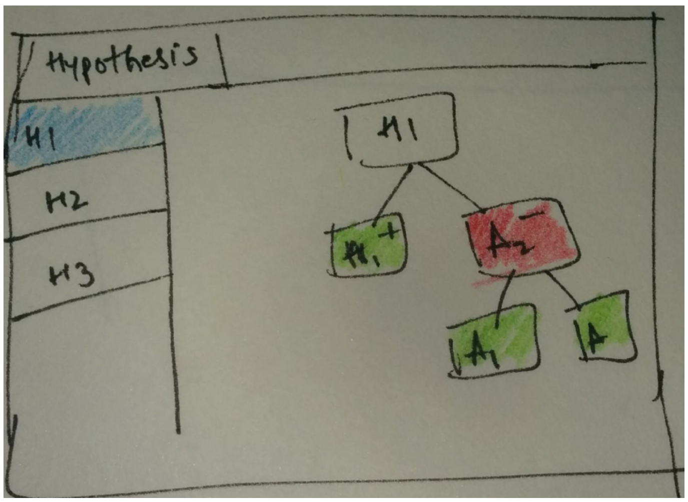


Figure 3: Second iteration of hypothesis view

.

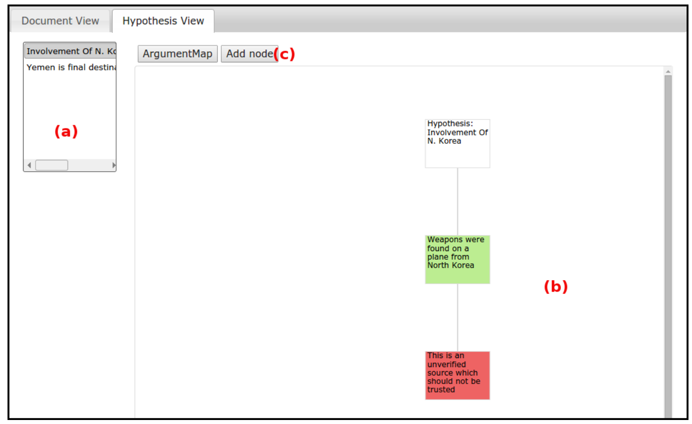


Figure 4: Hypothesis view with two argument maps (only one visible in the image). (a) The list of arguments and (b) the argument map visualization

.

from the author of the note. Currently, this happens outside the system as we do not have support for conversation among analysts within the tool.

The notes also show the date and time of creation of a note. This is an important piece of information for providing awareness. As analysts analyze more and more documents, their understanding about the document collection grows. This means that the annotations that were made during the initial stages of analysis might not be relevant anymore. An analyst can look at the date of the annotation created by another analyst and decide on its importance by looking at its creation date.

Hypothesis View

Figure 1b shows the initial design for the hypothesis view in which we started with a concept map [2]. The goal of the hypothesis view/concept map view was to support the analytical reasoning process. This design was inspired by the knowledge view of Aruvi [13]. In the sketch, we considered a single graph of entities and notes. The motivation behind this design was that when making sense of data, analysts are finding relationships between entities and information mentioning those entities. The sketch of figure 1b shows a concept map, which consists of color coded nodes containing text. A node in this concept map can either represent a named entity from the document collection or a note created by an analyst. When designing the concept map view, we were faced with the similar design choices as for the document view. We wanted to show information about the author of any particular node as well as the information about the role that the node served. A node in the concept map can be an entity or a note, or it can be pose a question, suggest a hypothesis or provide evidence to a hypothesis.

We evaluated the design with a paper prototype and realized several shortcomings with the design. First, we found that having a graph to support reasoning was not suitable. Graphs can contain circular links which creates the possibility of circular reasoning, which we did want. Another shortcoming was the fact that graphs tend to get unstructured. The unstructured layout of graphs can create cognitive load and make the reasoning process very hard. This was in dissonance with our requirement to support the reasoning process and ability to suggest and evaluate hypothesis.

We reiterated on the design to create the design shown in sketch 3. It consists of a list of hypotheses created by the analysts. Each hypothesis consists of a tree with the hypothesis

[2]The figure also shows a notes tab. In the initial design, we decided to have a separate notes tab to contain all the notes created by the analyst. However the idea was soon dropped as we realized that having a separate notes view disrupted the interaction.

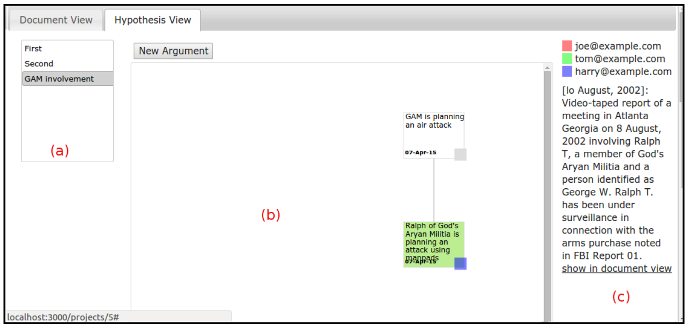


Figure 5: Hypothesis view with two argument maps (only one visible in the image). (a) The list of arguments and (b) the argument map visualization

.

Figure 6: Harry finds three documents containing the word "weapons" and adds three notes

.

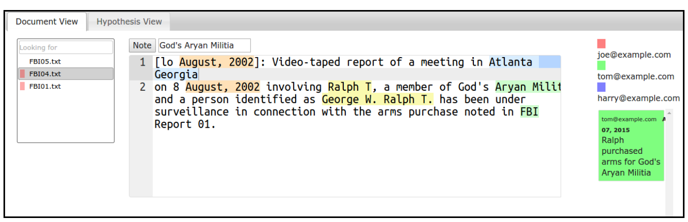


Figure 7: Tom believes Ralph purchased arms

.

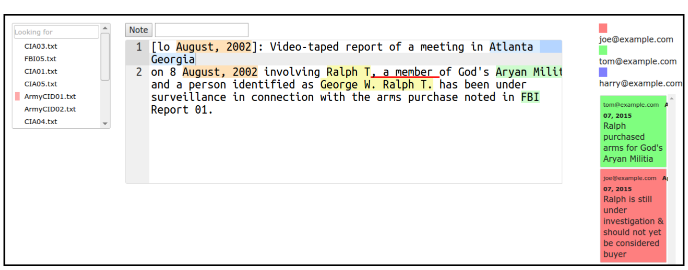


Figure 8: Joe adds a correcting comment

.

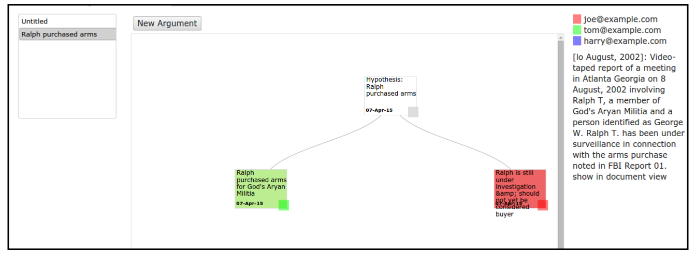


Figure 9: Joe corrects Tom's assumption.

.

at the root and supporting and/or opposing arguments as the internal or leaf nodes. We call this visualization an *argument map*. It should be noted that when we refer to argument maps in this paper, we refer to our implementation, which is a simplistic version of argument maps[3]. When the analyst selects a hypothesis from the list, the corresponding argument map visualization appears. An analyst can add supporting or opposing arguments. Every node supports or rejects the original hypothesis and not any internal node of the tree.

We implemented the second design into our *Rails* application (see figure 4). When evaluating the Rails prototype, we found some more issues with the visualization. The nodes of the argument map did not contain any information about when the note was created or the analyst who added a node. Also, it did not allow the analyst to look at the source document, where the note was created. All these pieces of information are important for creating awareness and provenance during asynchronous collaboration. During asynchronous collaboration, an analyst looking at a concept map will want to know who added an evidence. This helps the analyst in deciding on whether to trust or not any given piece of information, based on the expertise of an analyst in a given area. For example, a comment by Joe, who is an expert in analyzing large graphs and geography cannot be trusted about his comments about capabilities of a weapon. The date at which an evidence was added in support or opposition of a hypothesis is also important. Consider the scenario where an analyst working on a project logs into the system some substantial time, during which other analysts have added several hypothesis and updated argument maps. In such a scenario, information about when the arguments and hypotheses were added is important. It helps the analyst in getting aware about the activity that has been done recently. By looking at the creation date and time, the analyst can visualize how an argument progressed. Finally, the ability to go back to the source of a note helps the analyst in determining whether or not to trust a piece of information.

Figure 5 shows the final implementation of the hypothesis view. Figure 5a shows the list of hypothesis created by any of the collaborating analysts. Figure 5b shows a small argument map. In the argument map, analysts find evidence that supports or rejects a given hypothesis and add it to the map. The background color of any node in the argument indicates whether the node is supporting or opposing the hy-

[3] http://en.wikipedia.org/wiki/Argument_map

pothesis. Green nodes are in support of the hypothesis and the red nodes are opposing the hypothesis. In addition to the text of the source note, a node in the argument map also contains information about its creation time as well as its author. In figure 5b, the creation time of the note can be seen on the bottom left corner of the nodes. The bottom right corner of the node, contains a small square icon. The color of the icon indicates the analyst who added the given node. The analyst for a given color can be seen in the analyst legend (see figure 5c) just like the document view. When the user clicks on this icon, the text of the source document gets displayed in the right side bar of the hypothesis view. An analyst can then click on the link "show in document view" to see the contents of the document with the entities highlighted and all the contained notes, as described in the document view. In the figure 5b an analyst Harry(harry@example.com) has added evidence supporting the hypothesis.

Usage Scenario using JMIC dataset

JMIC dataset is a collection of fabricated reports about a possible threat to US security. It involves a collection of documents from different intelligence agencies in US. In this section, we demonstrate how we imagine our system to be used.

Consider a scenario with three analysts: Joe, Tom and Harry. They are collaborating to gain insights about the JMIC documents and thereby finding possible threats. Harry logs into the system and finds documents that mention the keyword weapons in it. He finds three documents. Based on these documents, Joe identifies, three important entities: Atlanta Georgia, Denver Colorado and God's Aryan Militia. He adds three notes, one to each entity (see figure 6). Joe signs out.

Some time later at a different office, Tom logs into the system. He finds that Joe has annotated three documents. He reads one of them starts exploring God's Aryan Militia (GAM). When he clicks on the entity, God's Aryan Militia, he finds about a person called Ralph who is involved in purchase of arms. He makes a note that "Ralph", a member of GAM is involved in purchase of arms. He also creates a corresponding argument map.

Later, Joe logs into the system to find out that the other two analysts have made some progress. By looking at the annotated documents and corresponding notes in the document view, he gains awareness about what Tom and Harry are exploring. He also looks at the hypothesis. He notices that the hypothesis is supported by Tom. He knows from working before with Tom that Tom is known for using uncertain evidence as certain information when proving his hypothesis. To check if Tom has made a mistake. He clicks on the source button to realize that Tom has indeed made a mistake. He notices that the report mentions that Ralph is still under investigation. Joe adds a comment that Ralph is under investigation and hence it might be inappropriate to consider him as the buyer of arms. He add an argument opposing Tom's claim (see figure **??** and **??**).

The example above is intended to be a simplistic demonstration of how we imagine it to be used. A complete analysis of the document collection, will not end at a single finding and will consist of many more annotations and hypothesis.

DISCUSSION

In this paper, we discussed the design principles and implementation of a tool that we designed to support hypothesis formation during asynchronous collaboration in the context of entity-based text analysis. The process of hypothesis formation was assumed to be one of the means to support collaborative analytical reasoning.

While designing the tool, we made several tradeoffs. First, in the document view, we considered the ability to view context for an entity to be more important than looking at related entities. This is a design decision that we made. We are still looking at ways to include information about related entities in the document view which will provide for better exploration.

Another decision we made was the level of highlights. We used a step function to indicate the amount of annotations made on any given document. Similarly, we used another step function to indicate the importance of an entity. The rationale behind these decisions was that it will be hard for analysts to distinguish between slight variances in color. Whether or not, this design decision is a valid one remains unexplored until we conduct a study to find how effective is the use of step function.

It is also important to mention a limitation of the design. We visually encode members of a team using colors. The range of colors that a person can distinguish is limited. In the present case, the number of users we imagined was not very high and color as a choice for distinguishing the users was suitable. We do not know if this visual encoding will work in cases where the number of collaborators is very large.

We believe that argument maps provide a structured to the analytical reasoning process. We found it to be quite effective in supporting collaborative reasoning and argumentation. However, we do recognize that there might be other possible visual representations that provide better support for collaborative reasoning. We also recognize a need for a user study to find the effectiveness of our system in promoting collaborative reasoning. In absence of a confirmatory study, our design choices are still based on personal preferences.

Another limitation of our system is the lack of any information about trust. When analysts read documents, they create notes which they use as evidence. Not all notes are correct. Some notes might have been created due to incorrect inference or limited knowledge by an analyst. Our system does not include any mechanism for users to rate evidence. We see this as an important feature to be implemented in the next version of the system.

In the beginning, we started with four design principles: *Awareness*, *Collaborative Hypothesis Formation & Evaluation*, *Provenance* and *Entity-based Analytics*. We implemented two views and designed them with these principles in mind. We made some tradeoffs while designing the view, often as we found some of our design principles requiring a tradeoff between each other. Yet, we were able to provide good, if not optimal support for all our design principles.

Furthermore, we believe that our system provides a suitable platform for studying asynchronous collaboration among professional analysts, especially during the analytical reasoning process. We presented a possible set of design principles for supporting asynchronous collaboration during hypothesis formation or vice versa. What remains unexplored is how appropriate are these design principles and what other design principles might be relevant.

CONCLUSION AND FUTURE WORK

In this paper, we presented a system for supporting hypothesis formation as part of the collaborative reasoning process during text analytics. In our discussion, we also discussed how our design evolved to present our design rationale. The final design may not be optimal and more work needs to be done to provide better support for analytical reasoning during asynchronous collaboration. We identified hypothesis formation using argument as one of the important visualizations for collaboration. There might be better alternatives that still need to be explored.

As an important step in this direction, we plan to conduct an exploratory study using our system to find the strength and weakness of our approach and recommend design principles for supporting reasoning in an asynchronous setting.

REFERENCES

1. IN-SPIRE™. http://in-spire.pnnl.gov/. (????).
2. Eric A Bier, Edward W Ishak, and Ed Chi. 2006. Entity Workspace : an evidence file that aids memory, inference, and reading. In *Intelligence and Security Informatics*, Sharad Mehrotra, Daniel D. Zeng, Hsinchun Chen, Bhavani Thuraisingham, and Fei-Yue Wang (Eds.). Springer Berlin Heidelberg, 466–472. DOI:http://dx.doi.org/10.1007/11760146_42
3. Ken W Brodlie, David A Duce, Julian R Gallop, Jeremy PRB Walton, and Jason D Wood. 2004. Distributed and collaborative visualization. In *Computer graphics forum*, Vol. 23. Wiley Online Library, 223–251.
4. Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL '05)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 363–370. DOI: http://dx.doi.org/10.3115/1219840.1219885
5. Amir Hossein Hajizadeh, Melanie Tory, and Rock Leung. 2013. Supporting awareness through collaborative brushing and linking of tabular data. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2189–2197. DOI: http://dx.doi.org/10.1109/TVCG.2013.197
6. By Jeffrey Heer, Fernanda B Viégas, and Martin Wattenberg. 2007. Voyagers and Voyeurs : Supporting Asynchronous Collaborative Visualization. April (2007), 87–97.
7. Petra Isenberg and Danyel Fisher. 2011. Cambiera: Collaborative Tabletop Visual Analytics. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. ACM, New York, NY, USA, 581–582. DOI: http://dx.doi.org/10.1145/1958824.1958916
8. Petra Isenberg, Danyel Fisher, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. 2010. An exploratory study of co-located collaborative visual analytics around a tabletop display. *2010 IEEE Symposium on Visual Analytics Science and Technology* (Oct. 2010), 179–186. DOI: http://dx.doi.org/10.1109/VAST.2010.5652880
9. N Kadivar, V Chen, D Dunsmuir, E Lee, C Qian, J Dill, C Shaw, and R Woodbury. 2009. Capturing and supporting the analysis process. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*. 131–138.
10. Paul E Keel. 2007. EWall: A visual analytics environment for collaborative sense-making. *Information Visualization* 6, October 2006 (2007), 48–63. DOI: http://dx.doi.org/10.1057/palgrave.ivs.9500142
11. A. Sanfilippo, B. Baddeley, A. J. Cowell, M. L. Gregory, R. Hohimer, and S. Tratz. 1999. Building a Human Information Discourse Interface to Uncover Scenario Content. *Military Intelligence* (1999), 1–6.
12. Kjeld Schmidt. 2002. The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW'. *Computer Supported Cooperative Work (CSCW)* 11, 3-4 (2002), 285–298. DOI: http://dx.doi.org/10.1023/A:1021272909573
13. Yedendra Babu Shrinivasan and Jarke J. van Wijk. 2008. Supporting the analytical reasoning process in information visualization. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, New York, New York, USA, 1237. DOI: http://dx.doi.org/10.1145/1357054.1357247
14. John Stasko, Carsten Görg, and Robert Spence. 2008. Jigsaw: supporting investigative analysis through interactive visualization. *Information Visualization* 7, 2 (2008), 118–132. DOI: http://dx.doi.org/10.1057/palgrave.ivs.9500180
15. J J Thomas and K A Cook (Eds.). 2005. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society.
16. Fernanda B Viégas, Martin Wattenberg, Frank Van Ham, Jesse Kriss, and Matt Mckeon. 2007. Many Eyes : A Site for Visualization at Internet Scale. August (2007), 1121–1128.

17. Kristan Wheaton, Kathleen Moore, Stephanie Williams, JoEllen Marsh, and Kevin Flanagan. 2012. *Wikis and Intelligence Analysis*. MCIIS Press.