

Evaluating the Relationship Between User Interaction and Financial Visual Analysis

Dong Hyun Jeong*
UNC Charlotte
Viscenter

Wenwen Dou†
UNC Charlotte
Viscenter
Heather Richter Lipford‡
UNC Charlotte
HCI Lab

Felesia Stukes‡
UNC Charlotte
HCI Lab
Remco Chang||
UNC Charlotte
Viscenter

William Ribarsky§
UNC Charlotte
Viscenter

ABSTRACT

It has been widely accepted that interactive visualization techniques enable users to more effectively form hypotheses and identify areas for more detailed investigation. There have been numerous empirical user studies testing the effectiveness of specific visual analytical tools. However, there has been limited effort in connecting a user's interaction with his reasoning for the purpose of extracting the relationship between the two. In this paper, we present an approach for capturing and analyzing user interactions in a financial visual analytical tool and describe an exploratory user study that examines these interaction strategies. To achieve this goal, we created two visual tools to analyze raw interaction data captured during the user session. The results of this study demonstrate one possible strategy for understanding the relationship between interaction and reasoning both operationally and strategically.

Index Terms: H.5.2 [Information Interfaces And Presentation (e.g., HCI)]: User Interfaces—Evaluation/methodology; H.5.2 [Information Interfaces And Presentation (e.g., HCI)]: User Interfaces—Graphical user interfaces (GUI)

1 INTRODUCTION

As the fields of information visualization and visual analytics mature, evaluation and verification become increasingly important. Recently, there have been tremendous efforts in such evaluation and verification, most notably the establishment and success of the BE-LIV workshops (Beyond time and errors: novel evaluation methods for Information Visualization) [1]. Some of the active research topics in this area include the evaluation of techniques, the measurement of insights, and metrics and methodologies for user studies. Together, these efforts are gradually validating the usefulness and power behind the science of information visualization and analysis.

In this paper, we take a slightly different approach to the evaluation of visual analytics. Instead of identifying the usefulness of a visual analytical tool or measuring the amount of insight a user gains through the tool, we assume that the tool we are testing is useful and that the user can discover the unknown using it. The questions we ask are: what is the user thinking while approaching the task, what do they actually do, and how do the two relate to one another in the context of the analytic process?

If we can answer these questions, we can extract an analyst's interactions in a meaningful way, discover the reasoning behind the

analysis process, and more importantly, relate the two to extract successful analysis strategies.

We performed a study that involved ten participants using a financial visual analytical tool called WireVis [2] to investigate a synthetic wire transaction dataset that contains known suspicious activities. WireVis is a tool developed and tested with financial analysts at Bank of America, and is now in deployment at the bank's wire investigation center WireWatch. The participants' interactions were logged into files and captured on video and audio. Through interviews and analysis of the captured session, we annotated each user's session with his or her reasoning process and discoveries.

The resulting annotations for each user are stored in a document containing both time stamps and unstructured text descriptions of the user's analytical steps using the WireVis tool. Although this document contains a great deal of information on the user's reasoning process, it does not explain what the user actually does in executing the reasoning process. Conversely, the log file of the user's interactions describes the user's interactions, but not the user's thinking during the investigation.

While the two sets of data depict two different aspects of a user's analysis process, neither is complete on its own. We believe that it is the combination of the two that gives us a holistic view of how the user interacts with the analytical tool for executing specific strategies in the investigation process. With the two sets of data combined, we can begin to fully understand the user's analytical process.

We combine the two sets of data through the use of visual analytical tools. We have implemented two separate tools for visualizing and analyzing the operational and strategic aspects of the user's investigation process. On the operational side, we use a time-based view to examine a user's interactions with the annotations. On the strategic side, we use a treemap approach that combines multiple users' data to identify similarities and differences in the users' foci during their investigations.

There are two main contributions in our work. First we introduce a methodology for capturing user interaction and reasoning in financial fraud investigation. We then apply a visual analytical approach to analyze and relate the user's interactions with this reasoning process. We present two visual analytical tools for discovering both the operational and the strategic aspects of a user's investigation.

The rest of the paper is divided into seven sections. First we discuss other research in user-interaction capturing and evaluation of visualization tools based on insight measurement. Then we describe how we generated the synthetic data and the setup of our user study. In section 4, we introduce the two visual analytical tools that we developed, followed by the results that we were able to obtain through the interviews and the use of the tools. Finally, we discuss some lessons learned in section 6, as well as future work and conclusion in the last two sections.

2 RELATED WORKS

As mentioned previously, there has been a notable increase in publications in the area of evaluation of visual analytics. Al-

*e-mail: dhjeong@uncc.edu

†e-mail: wdou1@uncc.edu

‡e-mail: fcartis@uncc.edu

§e-mail: ribarsky@uncc.edu

¶e-mail: richter@uncc.edu

||e-mail: rchang@uncc.edu

though these publications have different emphases and different approaches, most of them aim to measure and validate the effectiveness of visualization tools and visual analytics as a science. Among this growing literature, the areas most relevant to our work are efforts in evaluating visualization tools by measuring insights and in capturing and analyzing user interactions.

2.1 Evaluating Visualization by Measuring Insight

It has been proposed that the purpose of visualization is insight [8], and that the evaluation of any visualization is to determine the degree to which the visualization achieves this purpose. However, due to the elusive nature of insight, several varied definitions and measurement strategies have been proposed.

Using the definition that an insight is a unit of discovery, Saraiya et al. used a think-aloud method to evaluate bioinformatics visualizations [10, 11]. Rester and Pohl applied different methods to evaluate a visualization called Gravi++ that includes the use of insight reports [9]. More recently, Yi et al. performed extensive literature review and created a categorization of how people gain insight through information visualization [15].

Similar to these efforts, our work relies on the users telling us when they have made a discovery or gained insight. However, instead of using the result to determine the usefulness of a visualization, our interest is in correlating the reasoning and discovery process directly with the users' interactions.

2.2 Capturing and Analyzing User Interactions

Capturing user interactions for the purpose of understanding user behavior is very common in both industry and academia to evaluate effectiveness of software. Commercially, numerous applications can be found and purchased online that capture anything from desktop user activity (such as Morae [12]), to interactions on a website (which is a common feature in most web servers).

From an academic perspective, the most notable system designed for capturing user activities is the Glass Box system by Greitzer at the Pacific Northwest National Lab [5]. The primary goal of the Glass Box is to capture, archive, and retrieve user interactions [3]. However, it has also been shown to be an effective tool for evaluating research products [4]. Hampson and Crowley further investigated the use of Glass Box for capturing specific types of interactions for the purpose of intelligence analysis [6].

Although all the above mentioned systems are similar to our work in spirit, the approaches and scope are drastically different. While Morae and Glass Box both aim to capture and analyze user interactions across all applications on a desktop computer, our focus is much narrower and only considers a specific application in the financial visual analytical domain. The difference in scope means that Morae and Glass Box can only capture low-level events that are shared across all applications (such as copy, paste, mouse click, window activation, etc.) while we are able to capture events with known semantic information (such as what transaction the user has clicked on). The types of events that we capture are described in more detail in section 3.3, and the advantage of our approach is discussed in section 4.

It is important to note that although there are sophisticated methods in measuring insights and elaborate systems for capturing user interactions, to the best of our knowledge, there has not been any work on understanding a user's reasoning process through analyzing the interactions with an analytical tool.

3 USER EXPERIMENT

In order to understand the user's reasoning process through his interactions, we conducted a qualitative, observational study of a financial visual analytical tool called WireVis [2] (Figure 1). WireVis was developed jointly with wire analysts at Bank of America for discovering suspicious wire transactions. It is currently installed



Figure 1: Overview of WireVis. It consists of four views including a Heatmap view (top left), a Strings and Beads view (bottom left), a Search by Example view (top right), and a Keyword Relation View (bottom right).

at Bank of America's wire monitoring group WireWatch for beta testing. Although it has not been officially deployed, WireVis has already shown capabilities in revealing aspects of wire activities that analysts were not previously capable of analyzing.

3.1 Synthetic Data with Embedded Threat Scenarios

To preserve the privacy of Bank of America and their individual account holders, we created a synthetic dataset for the purpose of this study. Although none of the transactions in the dataset are real, we captured as many characteristics and statistics from real financial transactions as we could and modeled the synthetic data as closely to the real one as possible. The approach we took is loosely based on the methods described by Whiting et al. in developing the 2006 VAST contest [13].

For the purpose of the user experiment, it is important that the dataset is simple enough that users are able to look for suspicious transactions within the time frame of a study, but is complex enough that interesting and complicated patterns can be found. The synthetic dataset therefore contains 300 financial transactions involving approximately 180 accounts. 29 keywords are used to characterize these transactions, with some of them representing geographical locales (such as Mexico, Canada), and some representing goods and services (such as Minerals, Electronics, Insurance, Transportation, etc.) Each record of a transaction consists of the transferred amount, the sender and receiver's names, date, and one or more keywords relating to the transaction.

Four different types of known threat scenarios were identified. Two cases of each of the four types were created and embedded into the synthetic dataset based on the approach proposed by Whiting et al. [14]:

Incompatible Keywords in a Transaction: Transactions with two or more keywords that do not belong together. For example, a transaction containing the keywords "car parts" and "baby food".

Accounts with Dual Roles: An account that has had transactions of different incompatible keywords is questionable. For example, an account that transacts on "gems" at one time and "pharmaceuticals" at another.

Keywords with Large Amounts: Transactions of certain keywords are expected to have corresponding dollar amounts. For example, a transactions from a local store on "arts and crafts" should not be in the millions.

Change in Amounts Over Time: An account with an established temporal and amounts pattern receiving a large sum outside of its norm should be examined further. For example, an account with a

steady deposit of paychecks of fixed amounts on regular intervals receiving a transaction of a large amount.

3.2 Participants

Ten users volunteered for our user experiment, seven male and three female. All of the study participants were students at the University of North Carolina Charlotte. These participants were chosen based on their graduate-level research experience in either visualization, human-computer interaction, or virtual reality. Some of the participants have seen demos of WireVis or read the publication, but none had ever used the WireVis tool for analysis. All participants were familiar with Infovis and visual analytics concepts.

The user experiment was conducted in the Human Computer Interaction (HCI) Usability Lab at UNC Charlotte. The goal of the experiment was to understand the users' behaviors and strategies as they interacted with the visualization while performing fraud detection analysis.

At the beginning of each study, each participant was trained on the concept of wire fraud and the use of WireVis for 10 minutes. We introduced each of the views and features, and then walked the participant through two scenarios of finding suspicious activities. The participant was provided a one-page overview of the functionality of WireVis and encouraged to ask questions. We also provided hints as to what kinds of patterns the participant might investigate in his analysis. In the real world, an analyst's training, experience, and actual world events would lead him to develop these strategies.

Following the training, the user was asked to spend 20 to 30 minutes using WireVis to look through the dataset to find other suspicious activities. We asked the participant to think-aloud to reveal his strategies. We specifically encouraged the participant to describe the steps he was taking, as well as the information used to locate the suspicious activities. Once the user identified a transaction as suspicious, he was encouraged to continue looking for others until the time limit was reached. At this point, a post-session interview was conducted for the participant to describe his experience and additional findings.

3.3 Capturing User Interaction and Reasoning

Several methods were used to capture each participant's session as thoroughly as possible. A commercial usability software Morae was used to capture the screen, mouse clicks, and keystrokes of the user's interactions. A separate webcam was used to record the user's audio and actions during the session. Lastly, functions built into the WireVis system captured the user's interaction with the tool itself.

Unlike Morae, the interactions captured directly in the WireVis system contain semantic information relevant only to the WireVis system. Instead of recording every mouse movement or keystroke, WireVis only captures events that generate a visual change in the system. For example, a mouse movement that results in highlighting a keyword in the Heatmap view will generate a time-stamped event noting that the user has highlighted a specific keyword.

Since the participants were encouraged to think-aloud during their sessions, we were able to review the video and audio recordings of the participants to recreate their reasoning and thinking processes. We ("the observers") used recordings from each participant's think-aloud session and post interview to construct a hierarchical annotation file that described in text different levels of their strategic and operational approaches in investigating fraudulent activities.

4 VISUAL ANALYSIS OF USER INTERACTION AND REASONING

The interactions and reasoning processes of our participants were very complex. Similar to intelligence analysts, our participants had

varied preconceived knowledge about the keywords and their meanings. In interacting with WireVis, the participants also had disparate interpretations of the visualization in determining what appeared to be suspicious. Even when the participants identified the same embedded threat scenarios, their methods and approaches were decidedly dissimilar.

The complexity of a participant's interaction and reasoning process makes it very difficult to describe his behavior and intent quantitatively. We therefore propose approaching this complex problem using visual analytical methods. Two visual tools are created to investigate the operational and strategic aspects of a participant's investigation process. Using the operation analysis tool, we can analyze a participant's interactions in accordance to his reasoning process; while using the strategy analysis tool, we can compare the strategies of different participants or groups of participants. Together, these two tools provide the means to discover the relationship between user interactions and reasoning processes at all levels.

4.1 Operation Analysis Tool

Our operation analysis tool is designed to support the analysis of a participant's operational interactions in relation to his annotations (Figure 2). The tool is implemented in OpenGL, and is fully zoomable and pannable and supports selections of interaction elements for detailed inspection. The x-axis of the main view represents time, with a striped background indicating the length of a fixed time duration (defaulted to 60 seconds per strip). The y-axis is divided into 5 sections, with each section supporting one aspect of the participant's investigation process. Figure 2 (A)-(E) show the 5 perspectives, which are the participant's annotations (A), the participant's interactions with the three views in WireVis (B), the depths of a participant's investigation (C), the areas of the participant's investigation (D), and the time range of the investigation (E).

The sliders in (F) allow the user to scale time, while checkboxes in (G) control various visualization parameters. The detail view in (H) depicts detailed information of a specific user-interaction element.

Annotation View As mentioned in section 3.3, the results of our participants' think-aloud during the experiment are recorded into separate files. These annotations to the participant's investigation are shown in this view. As can be seen, our participants often exhibit a hierarchical structure in their reasoning process, with the highest level of reasoning depicting strategies they employ such as "seek all keywords related to the keyword *food*" The lower levels depict specific operations to execute those strategies, ranging from "search for keywords other than *food* relating to account 154" to "identify the receiver of a transaction (account 64) of account 154."

The hierarchical nature of the participants' reasoning are represented in the annotation view, with the higher level annotations shown above the lower ones as interactive floating text boxes [7]. The time range of each annotation is drawn as nested boxes using different colors. The user can select any particular annotation, and its corresponding time range is highlighted across all the other views (Figure 2).

WireVis Interaction View WireVis uses multiple coordinated views to visualize different relationships within the data. In the WireVis Interaction view, we look to display the participant's usage pattern of the WireVis tool. The three rows in this view correspond to the three main interactive views in WireVis: Heatmap, Strings and Beads, and Search by Example. In each view, we can choose two different attributes of the participant's selection. In Figure 2 (B), the two attributes are keywords (shown as red dots) and accounts (shown as green dots).

On first glance, it is easy to see which views in WireVis the participant interacts with over time. On closer inspection, the distribution of the red (keywords) and green dots (account) also reveal high-level patterns in the investigation. Scattered red dots could

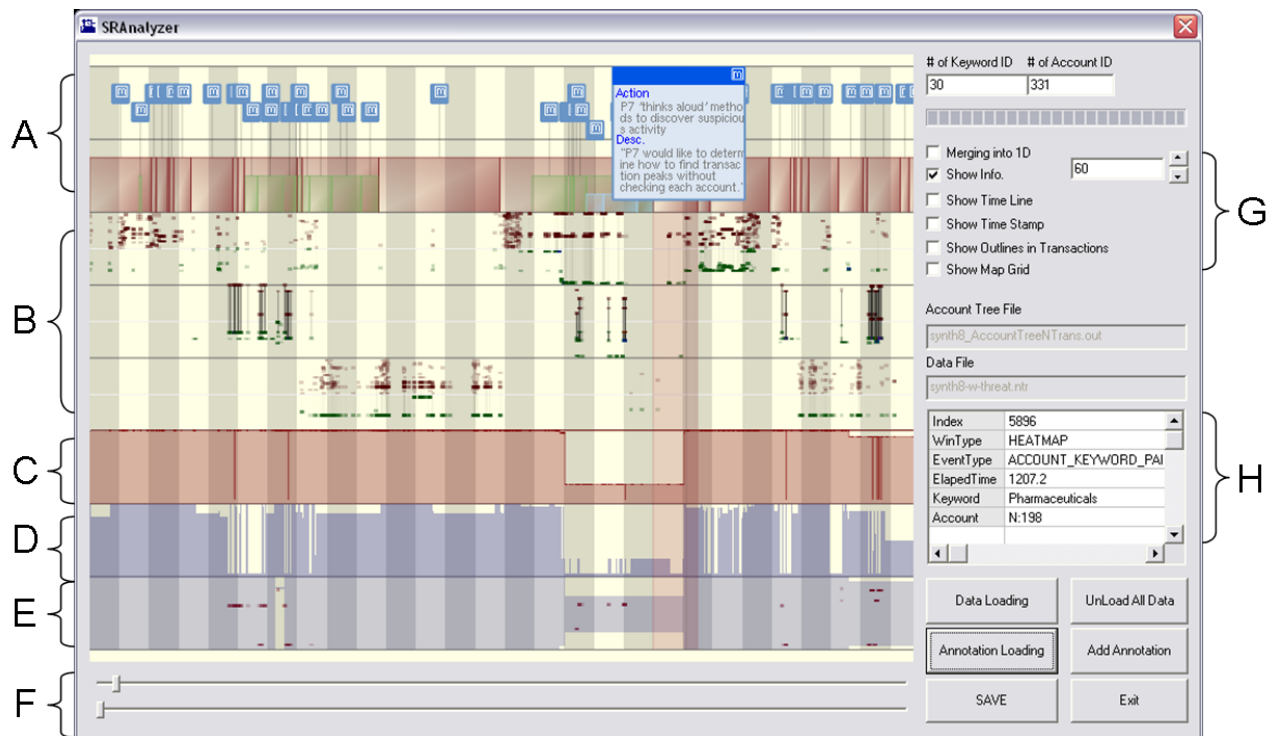


Figure 2: Overview of the operation analysis tool. (A) shows the participant's annotations. (B) shows the participant's interactions with the three views in WireVis. (C) represents the depths of a participant's investigation. (D) shows the areas of the participant's investigation, and (E) the time range. Sliders in (F) control the time scale, while checkboxes in (G) change various visualization parameters. (H) shows the detail information of a participant's selected interaction element.

indicate an exploration of keywords, whereas concentrated green dots (e.g., if the green dots are aligned horizontally) could reveal the participant's interest in a specific account. When both red and green dots appear together and are connected by a line, it denotes that the participant is investigating the relationship between the two (such as a cell in the Heatmap view in WireVis).

Depth View On top of visualizing a participant's direct interactions with WireVis, it is also important to see some semantic information regarding the participant's investigation process. In this view, we visualize the "depth" of a participant's investigation by displaying the number of visible transactions in WireVis. For example, when the participant is looking at the overview in WireVis, our Depth view will be completely filled, indicating that all transactions are visible. As the participant zooms in to investigate specific keywords or accounts, the Depth view will show a drop in visible transactions (Figure 2 (C)).

The Depth view also indicates when a participant requests detailed information for a specific account or transaction (such as double-clicking on a bead in the Strings and Beads view). These interactions show up as a vertical line, which is easily distinguishable from a participant's operations for zooming in or focusing on a specific area in the data.

Areas View While the Depth view shows the number of visible transactions in WireVis, it is also relevant to indicate interactions that highlight areas that the user has shown interests. These interactions are commonly used in WireVis through the "mousing-over" operation. As the participant mouses-over keywords, accounts, or transactions in WireVis, the system displays information about the highlighted data without requiring the user to change the zoom level or focus.

Using the mouse-over operation in WireVis is common, and of-

ten indicates an exploration process in which the participant is looking for suspicious areas for further investigation. In the Areas view, a high variation in a short amount of time could indicate such an exploration process, while a more leveled section suggests that the participant is investigating specific activities (Figure 2 (D)).

Time Range View Time is an important aspect in discovering financial fraud, and WireVis provides views to explore the temporal dimension. In the Time Range view, we look to capture the participant's time-based investigation. The y-axis of the Time Range view denotes the dates represented in the data from more recent to least. A fully colored section indicates that the participant's investigation spans the entire time range, whereas a change would denote that the participant has zoomed in to a specific time period (Figure 2 (E)).

The dots in the Time Range view indicate selections of transactions of a specific date. In WireVis, this is done by either mousing-over or double-clicking on a bead in the Strings and Beads view. A high concentration of the appearance of these dots often suggests that the participant has found some specific transactions and is looking to find out the details of these transactions.

4.2 Strategy Analysis Tool

As opposed to operation, strategy is a long term plan of action designed to achieve a particular goal. As shown in the Annotation view of the operation analysis tool (section 4.1), most of our participants exhibit both strategic and operational reasoning when investigating fraud. So besides addressing the question "what do the participants actually do" using our operation analysis approach, we also look to investigate the high level strategies that the participants employ while approaching the tasks. Through the use of our strategy analysis tool, we can identify each participant's areas of interest as well as comparing different participants' strategies.

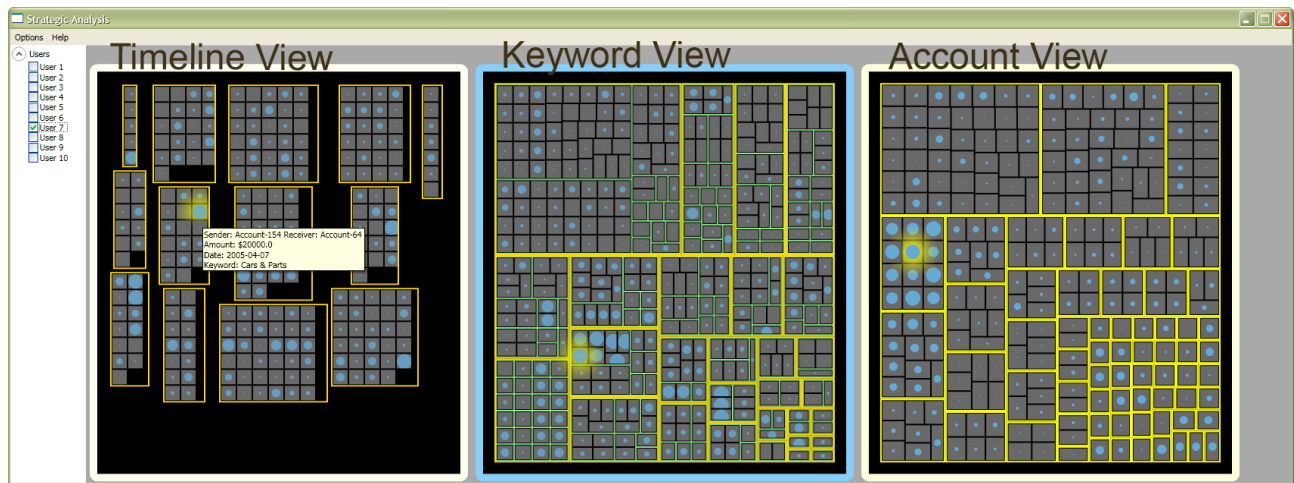


Figure 3: The same participant's interactions in Figure 2 are shown in using the strategy analysis tool. The left view shows transactions grouped by time, middle view shows grouping by keywords, and the right view shows grouping by accounts. The patterns in the account view indicate that the primary strategy employed by this participant was to look for accounts with similar activities.

We adopt treemap as the basis of our strategy analysis tool. The treemap visualization allows us to investigate similarities between our participants' strategies without considering the flow or speed in which our participants execute their strategies. As mentioned in section 4), our participants had varied preconceived knowledge about the keywords and their meanings, and therefore approached the investigation tasks differently. Many of them identified the same embedded fraud scenarios, but none of them shared the same path in discovering these activities. Using our modified treemap visualization, we can identify the participants' strategies without regard to the paths they have chosen.

The initial layout of the strategy analysis tool shows three different treemap views classifying the transaction data based on three attributes: time, keywords, and accounts (Figure 3). The three views are coordinated such that highlighting a transaction in one view also highlights the same transaction in the other two views. We choose transactions to represent the lowest level of the treemaps because they represent the lowest granularity of the data. A colored circle is displayed on each cell, and the size depicts the amount of time the participant's investigation has included that transaction. When comparing two participants or two groups of participants, the color of the circle indicates which of the two participants spent more time on the transaction (Figure 6).

Timeline View Transactions in this view are classified based by their date. Each grouping contains transactions of the same month. As shown in Figure 3, the transactions in our synthetic data set span a 13 month period. Note that the participant depicted in this view did not perform his investigation based on the transaction dates as the circles appear fairly evenly through all 13 months.

Keyword View This view applies two different classification criteria. On the top level, transactions are grouped based on keywords (shown as yellow cells in Figure 3). Each cell is then further subdivided by individual accounts (shown as green cells).

Since a transaction often contains multiple keywords, the same transaction could appear in more than one keyword cells. Similarly, every transaction contains two accounts, a sender and a receiver, so a transaction will always appear at least twice, once for each account. Due to these two reasons, the total number of transaction cells in this view are greater than those in the Time view. However, we find this layout more intuitive for understanding a user's strategy involving keywords. For example, in Figure 3, it is easy to see that

the participant focused on a few specific keywords, but even more specifically on a few accounts relating to those keywords.

Account View The Account View orders the transactions based on their corresponding sending accounts. As shown in Figure 3, this view makes clear that this participant's strategy in discovering financial fraud using WireVis is almost entirely based on the detailed investigation of one or two accounts. Time and keywords appear to be secondary considerations during his investigation.

5 ANALYZING USER INTERACTIONS AND REASONING

From the annotations and from processing the audio and videos of the interviews from the user experiment, we know that all participants identified several transactions they felt were suspicious. While the training and the hints we provided influenced partici-

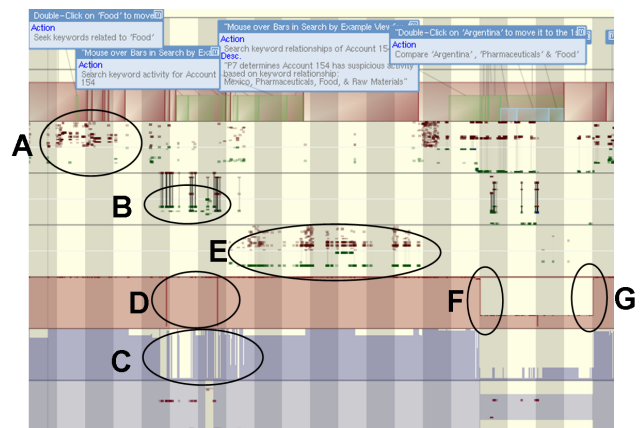


Figure 4: A participant's interactions and annotations are shown in the operation analysis view. (A) shows the participant's activities in searching for keywords in the Heatmap view. (B) shows that the participant utilized the Strings and Beads view for exploration (C), but did not perform extensive detailed inspections on any keyword or account (D). The participant then used the Search by Example view to discover accounts of similar behaviors (E), and drilled down into a specific account for further investigation (F). However, no suspicious activities were found, and the participant zoomed out again (G).

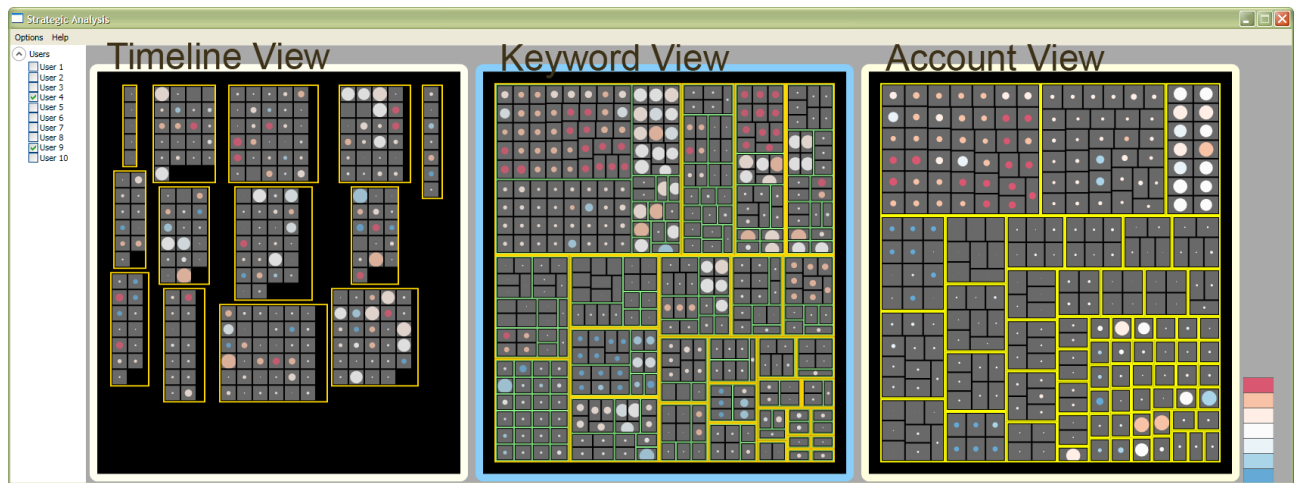


Figure 5: Two participants' interactions are compared using the strategy analysis tool. The red circles indicate a dominant inspection from one participant, and the blue circles indicate the dominance from the other. A white circle denotes that the two participants spent equal amount of time investigating the transaction. The size of each circle represents the total amount of time spent on investigating each transaction. In this case, the two participants exhibited similar strategies as most circles are white.

participants' early fraud detection strategies, we noticed that the users quickly developed their own strategies to locate fraudulent transactions. In the following sections, we show examples of using our operation and strategy analysis tools to analyze the participants' interactions within WireVis and relate the interactions to the participants' annotations.

5.1 Operation Analysis Result

Figure 4 shows a participant's interactions. The participant started by examining keywords in the Heatmap (Figure 4 (A)), and identified a specific account for examination as evident by the subsequent appearance of that account in the Strings and Beads section of the WireVis Interaction view (shown as green dots in (Figure 4 (B))). The Areas view (Figure 4 (C)) indicate that while the participant explored that account extensively in Strings and Beads, the lack of activities in the Depth view (Figure 4 (D)) show that the participant did not drill down for detailed investigation. This series of interactions suggest that the participant did not identify any suspicious behavior during this time.

The participant's annotations for this period corroborates with our analysis. He indicated that he was identifying a specific account (*account 154*) based on the keyword *food* that he thought was interesting. However, upon further investigation, he determined that the transaction involving *food* from *account 154* is not fraudulent because the amount is too low.

Nonetheless, the participant continued to believe that *account 154* was of interest. As shown in Figure 4 (E), he used Search by Example to identify other accounts that were similar to *account 154*. After he identified one (*account 128*), he used the Heatmap view (Figure 4 (F)) to drill down to examine the details of that account. Unfortunately, upon further investigation, no fraudulent activities were found, and the participant zoomed back out in the Heatmap view (Figure 4 (G)).

This observation is again corroborated by the participant's annotations. The only point of ambiguity in our analysis came at the end of the process when the participant zoomed out in the Heatmap view. We believe that although we had analyzed the participant's activities correctly, it was possible that the participant had indeed identified fraudulent activities while zoomed in on *account 128*. His subsequent zooming out of the Heatmap view could have been to start a different part of the investigation or to execute a new strategy.

Without explicit annotations, we currently have no way to differentiate the two scenarios using only the participant's interactions.

This simple example shows that using the operation analysis tool to analyze a participant's behavior indeed corresponds to the participant's annotations describing his reasoning process. Although we found that different strategies could generate similar patterns in the tool, we believe that our tool can still provide a reasonable overview to the participant's intent.

5.2 Strategy Analysis Results

The strategy analysis tool can be applied either to analyze a single participant's strategy, or to compare the difference between the strategies of two participants. In this section, we demonstrate the use of the tool, and discuss how the participants' strategies can be discovered.

Analyzing a Single Participant Visualizing a single participant's interactions in the strategy analysis tool reveals patterns of high-level reasoning processes, from which we can extract the participant's strategy without examining each interaction event. Since the patterns are independent of the order of task performed, or the time spent on each task, our tool can be applied to analyze investigation process of varying lengths.

Since we have examined the operational aspect of a particular participant using the operation analysis tool, we go on to identify the high level strategies employed by the same participant using the strategy analysis tool. As shown in Figure 3, the participant focused his investigation based on a specific account over a keyword or time-based approach. By highlighting that account, we can identify that the name of the account is *account 154*. This finding is consistent with the operation analysis result of the participant. While the two tools eventually reveal the same high-level strategy employed by the user, the strategy view reveals this account-centric method easily and intuitively, while using the operation analysis tool required a long process of reasoning and correlating with annotations.

Comparing Two Participants Since our strategy analysis tool is not bound by the time or sequence of a participant's analysis, we can compare the two easily by displaying them simultaneously in the same view. Figures 5 and 6 show two comparisons of two different pairs of participants.

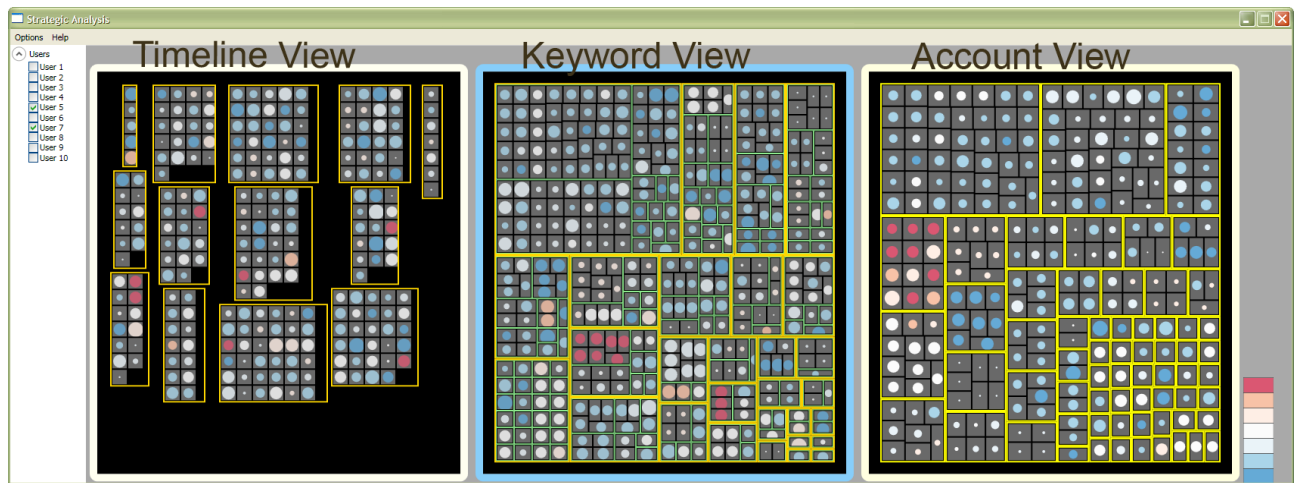


Figure 6: Contrary to Figure 5, this image depicts two participants with different strategies. It can be seen that one participant (red) employed an account-focused approach, while the other (blue) performed the investigation based on keywords.

Since the color of a circle indicates which participant spent more time investigating a transaction, and the size of a circle shows the combined time for both participants, we can identify transactions that are of interest to both participants by identifying large white circles. In Figure 5, we can easily see two participants sharing a similar strategy as the circles are mostly white and large with very few bright red or blue circles. From this figure, we can conclude that both participants were interested in a few specific accounts, and they both spent most of their time investigating them.

On the other hand, Figure 6 tells a story of two participants with different strategies. The differences in colors and sizes of circles indicate that the two participants rarely share the same focus. The participant represented as blue utilized a more keyword-centric approach, spending most of his time looking for keyword-related possibilities; while the other participant focused more on searching for suspicious accounts.

5.3 Combining the Two Analysis Tools

We believe that the two tools presented in this section adequately capture both the operational and strategic aspects of a user's interaction and reasoning process. However, it is the combined use of the two tools that truly reveals a participant's intent and approach to investigating fraud. With the use of the operation analysis tool, we can identify localized patterns for specific investigation processes. These patterns can then be visualized in the strategy analysis tool to uncover the high-level strategy behind them, as well as compared against other participants to discover different approaches to similar tasks. Although our two tools are not tightly integrated at present time, we look forward to merging them into one single application in the near future.

6 DISCUSSION

While the user experiment described in this paper achieved the goal of connecting the user's interaction with his reasoning process, there were some unexpected results that are worthy of further investigation. First of all, most of our participants expressed that they lacked domain expertise in financial fraud detection. However, many of them identified what they believed to be suspicious activities that were not part of the embedded threats scenarios and yet made very convincing arguments on why the activities could have been. These discoveries reaffirm the capabilities of WireVis in discovering suspicious trends and patterns, but they also bring to our attention that our synthetic data set is far from the "ground

truth" we hoped it could be, and that the definition of suspicious activities is indeed subjective.

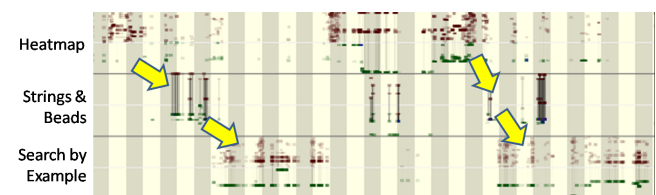


Figure 7: A participant's interactions in the operation analysis tool. Each yellow arrow represents how the sequence of operations while searching for suspicious activities are created.

Furthermore, the ways in which our participants interacted with WireVis were sometimes unexpected. Through the operation analysis, we find that almost all of our participants exhibit the flow of using Heatmap first, followed by Strings and Beads, and eventually Search by Example (Figure 7). This behavior could be a result of the sequence of the views presented during the training sessions (the three views were introduced in that order), or a reflection of the inherent functionality of each view (overview vs. detail presentation). We currently do not have any concrete arguments to support either possibility, but we are looking into better understanding this behavior.

We are still a long way away from being able to fully understand the relationship between a user's interactions and his reasoning process for all visualizations, but we believe that we have taken a small step in the right direction. Our confidence in our approach in monitoring and analyzing a user's behavior is supported by financial fraud analysts at Bank of America, who, at this point, have little or no way to review or understand other analysts' reasoning process in investigating fraud. By using our tool, high-level semantic user events in WireVis can be captured such that each analyst's actions can be replayed without the storage of massive amount of video (screen-capture) data. When these events are reviewed using our operation and strategy analysis tools, the analysts' interaction patterns and reasoning processes at different levels are revealed. Together, we believe our methodology and tools provide the foundation for transforming analysis processes into visualizations that are sharable, repeatable, and accountable.

7 FUTURE WORK

Although we were capturing as many different types of events in our user experiment as we could, we felt that we could have strengthened our analysis of user activities with the use of an eye tracker. Similar to most multiple-coordinated view visualizations, interactions with one view in WireVis cause updates to all other views on screen. Often, it is the combined information depicted in all coordinated views that the user is basing his reasoning on. Unfortunately, without an eye tracker, it is difficult to obtain where the user is focusing his attention even with the use of the think-aloud method. So while we know the user is mousing over a set of keywords, we cannot be certain if the user is interested in the amounts associated with these keywords, or the dates in which the keywords appear. With the aid of an eye tracker, we look forward to capturing not just which data items the user is interacting with, but the relationships associated with those data items that are of interest to the user as well.

In regard to the user experiment, the participants of our experiment presented in this paper were not our intended users of WireVis. To truly understand the use of WireVis in executing strategies in fraud detection, we need to perform user experiments on real financial analysts. We believe that our experiment will have to be slightly modified to allow for more training of the WireVis tool to analysts without experience with visual analytical tools. Nonetheless, we believe that the general methodology and the specific tools presented in this paper can be directly applied without change. While we are still investigating necessary modifications to the details of the user experiment, we have already received some positive feedback from members of Bank of America on performing the user experiment with their analysts. We are hopeful that we will be able to analyze real financial analysts' interactions and reasoning processes using WireVis in the near future.

8 CONCLUSION

Understanding a user's reasoning process within visual analytical tools is an essential part of the process of evaluating and verifying these tools. In this paper, we present a methodology and visual tools to tackle the complex relationship between user interaction and financial visual analysis. The questions we seek to answer are: what is the user thinking while approaching the task of discovering fraud, what do they actually do while pursuing this task, and how do the two relate to one another in the context of the analytical process? We employ a think-aloud method while capturing both low-level and high-level user interactions to capture the investigative process of an analyst using the WireVis system. The operation and strategic analysis tools that we present in this paper are shown to be effective in discovering a user's reasoning process at all levels. Although we are still a long way from being able to fully understand the relationship between user interaction and visual analysis, we believe that our approach has taken us one step closer.

ACKNOWLEDGEMENTS

The authors wish to thank Isaac Thompson for his help in performing the experiment and Caroline Ziemkiewicz, Evan Suma, and Lane Harrison for their help in editing the paper.

REFERENCES

- [1] BELIV '06: *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, 2006. Conference Chair-Enrico Bertini and Conference Chair-Catherine Plaisant and Conference Chair-Giuseppe Santucci.
- [2] R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto. Wirevis: Visualization of categorical, time-varying data from financial transactions. In *IEEE Symposium of Visual Analytics Science and Technology*, pages 155–162. IEEE, October 2007.
- [3] P. Cowley, J. Haack, R. Littlefield, and E. Hampson. Glass box: capturing, archiving, and retrieving workstation activities. In *CARPE '06: Proceedings of the 3rd ACM workshop on Continuous archival and retrieval of personal experiences*, pages 13–18, 2006.
- [4] P. Cowley, L. Nowell, and J. Scholtz. Glass box: An instrumented infrastructure for supporting human interaction with information. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS '05*, pages 296c–296c, January 2005.
- [5] F. Greitzer. Methodology, metrics and measures for testing and evaluation of intelligence analysis tools. PNWD-3550, Battelle-Pacific Northwest Division, Richland, WA, 2005.
- [6] E. Hampson and P. Cowley. Instrumenting the intelligence analysis process. In *Proceedings of the First International Conference on Intelligence Analysis Methods and Tools*, 2005.
- [7] J. Hong, D. H. Jeong, C. D. Shaw, W. Ribarsky, M. Borodovsky, and C. Song. Gvis: A scalable visualization framework for genomic data. In *EuroVis 2005: Eurographics / IEEE VGTC Symposium on Visualization*, pages 191–198, 2005.
- [8] C. North. Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, 26(3):6–9, 2006.
- [9] M. Rester and M. Pohl. Methods for the evaluation of an interactive infovis tool supporting exploratory reasoning processes. In *BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, pages 1–6, 2006.
- [10] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456, 2005.
- [11] P. Saraiya, C. North, and K. Duca. Visualizing biological pathways: requirements analysis, systems evaluation and research agenda. *Information Visualization*, 4(3):191–205, 2005.
- [12] TechSmith Corp. Morae. <http://www.techsmith.com/morae.asp>.
- [13] M. A. Whiting, W. Cowley, J. Haack, D. Love, S. Tratz, C. Varley, and K. Wiessner. Threat stream data generator: creating the known unknowns for test and evaluation of visual analytics tools. In *BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, pages 1–3, 2006.
- [14] M. A. Whiting, J. Haack, and C. Varley. Creating realistic, scenario-based synthetic data for test and evaluation of information analytics software. In *BELIV '08: Proceedings of the 2008 CHI workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, pages 74–81, 2008.
- [15] J. S. Yi, Y. ah Kang, J. Stasko, and J. Jacko. Understanding and characterizing insights: How do people gain insights using information visualization? In *BELIV '08: Proceedings of the 2008 CHI workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, pages 39–44, 2008.