RiskVA: A Visual Analytics System for Consumer Credit Risks Analysis

Xiaoyu Wang, Dong Jeong, Remco Chang, Arun Pinto and William Ribarsky

ABSTRACT

Consumer credit risk analysis plays a significant role in stabilizing a bank's investments and in maximizing its profits. As a large financial institution, Bank of America relies on effective risk analysis to minimize the net credit loss resulting from its credit products (e.g. mortgage and credit card loans). Due to the size and complexity of the data involved in risk analysis, risk analysts are facing challenges in monitoring large amounts of data, comparing its geospatial and temporal patterns, and developing appropriate management strategies based on the correlation from multiple analysis perspectives. To address these challenges, we present RiskVA, an interactive visual analytics system that is tailored to support credit risk analysis. RiskVA provides risk analysts with interactive data exploration and information correlation, and visually assists them in depicting market fluctuations and temporal trends of the targeted credit product. When evaluated by analysts from Bank of America, RiskVA was appreciated for its effectiveness in performing in-depth risk analysis, and is considered useful in facilitating the bank's risk management operations.

KEYWORDS: Risk management, visual analytics.

INDEX TERMS: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Life Cycle; K.7.m [The Computing Profession]: Miscellaneous—Ethics

1 INTRODUCTION

Consumer credit risk analysis plays a significant role in stabilizing a bank's investments and in maximizing its profits. Credit risk management, in general, refers to the process in which the investors assess the risk of loss arising from a consumer who does not make payments as promised [3]. Credit risk typically occurs in investing and in the allocation of capital.

For most banks and financial institutions, loans and credit products are the largest and most obvious source of risk. In order for a bank to profit from a large consumer base, it must invest in credit products (e.g. credit cards, mortgages) that are reasonable to customers. However, the bank must strike a balance between the investments and the substantial amount of capital in its reserve; so that investments would be profitable yet sustain the bank's financial stability.

Therefore, the assessment of credit risk is, on the one hand, crucial for banks to position themselves to profit through balancing credit investments and returns; on the other hand, it is critical for the stability of an entire financial market. Inadequate risk management can result in severe consequences for companies as well as individuals.

As shown in their study of the correlations between recessions and banking crisis, Bloom et al. [4] suggested that the credit investment strategies were directly associated with the stability of the entire financial market; an unexpected credit crunch would lead to the complete disarray of the financial markets. For example, the loose credit risk management of financial firms was determined to be one of the factors that triggered the recession in 2008.

As a large institution, Bank of America (BOA) constantly faces the challenge of managing its credit risk. Their stability relies on the effective risk analyses to minimize the net credit loss resulted from its credit products, and to determine profitable market strategies. Essentially, BOA emphasizes the use of risk management to quantify the potential losses in an investment and to take the appropriate action given investment objectives and risk tolerance.

However, given BOA's wide range of credit investments, analyzing risk in such diversified portfolio has become an overwhelming process. This demands that analysts evaluate credit risks both temporally (i.e. identifying market turning points before and after the recession) and across credit markets (e.g. comparing product performances in major cities). Current analysis practices and analytical tools can't meet the challenges to comprehend the trends and patterns of markets from multiple perspectives.

Exacerbating this challenge is the increasing size and complexity of the collected credit data that each analyst needs to examine. This places an extra burden on each individual analysis process in terms of the efforts needed to acquire the most appropriate information. This is exacerbated further by the need to gather information from heterogeneous data sources and bring it into a common picture.

To help address these crucial financial challenges, we formed a research partnership with Bank of America to investigate novel analysis technologies. One of our first actions within this partnership was to observe and characterize the risk analysts' analytical workflow. This domain characterization granted us the opportunity to closely identify the key perspectives in risk analysis, and helped us gain insights on the design elements that are needed to facilitate such analysis.

In this paper, we present RiskVA, an interactive visual analytics system that is tailored to support credit risk analysis. RiskVA addresses the aforementioned challenges by supporting interactive data exploration and information correlation over a large corpus of credit data. It aids the domain analysts in depicting and comparing the performance of the credit products by visually revealing market fluctuations and temporal trends of the targeted credit products. To support individual analysis workflow, RiskVA further allows the analysts to choose different combinations of visualizations and to customize the visual interface based on their own preferences.

To evaluate the efficacy of our system, we conducted expert evaluations with risk analysts from BOA, and found that most analysts considered RiskVA to be useful and complimentary to their existing analysis needs. We further identified analysis scenarios for which our system could provide analysts with insights to develop appropriate risk management strategies.

Given these results, this work presents substantial qualitative advances over current practices in consumer risk analysis.

- It provides a visual exploratory environment to handle consumer risk data that scales to hundreds of thousands of credit data over any given length of time.
- It provides highly coordinated interactive visualizations to enhance both tactical and strategic risk analyses that are essential in identifying emerging risks.
- It provides customizable workspaces that support the individual analyst's analysis routines.

The remainder of this paper is structured as follows: Section 2 characterizes the analysis processes of risk analysts and describes our system's targeted users. Section 3 provides more detail about the limitations of current risk analysis practices. Section 4

presents our visual analytics system, RiskVA. Section 5 provides several scenarios in which our system can facilitate risk analysis. In Section 6, we present our evaluation with risk analysts and our discussion for advancing from the current stage. We conclude the paper in Section 7.

2 Domain Characterization: consumer Credit Risk Analysis

In this section, we describe the nature of consumer credit and the current practice in conducing credit risk analysis. By discussing these in detail, we intend to shed light on the characteristic of these challenges and how an interactive visual analytics system can bring about a substantial improvement in risk management.

The importance of credit risk analysis, as part of financial risk analysis, comes from the New Basel Capital Accord (Basel II), published in 1999 and revised in 2004 by the Basel Committee on Banking Supervision (BCBS). In general, credit risk analysis is "*the process of establishing credit standards for investors and counterparties, assessing the portfolios of the existing credit products, and preventing deterioration in the credit standing of a bank's counterparties"* **[3]**. Financial companies typically use credit risk models to evaluate the insolvency risk caused by credits that enter into default,

According to BCBS, roughly speaking, the tradition in consumer credit has been to take performance data over a fixed time interval for a sample of consumers. Then each consumer is ranked by performance, where unsatisfactory performance is often equated with being 60 or 90 days overdue with repayments. This historical information is reviewed by risk analysts and is then used to model internal risk strategies based on the characteristics of each consumer. The modelled strategy is then applied to new consumer to determine whether they are above the cut-off level to

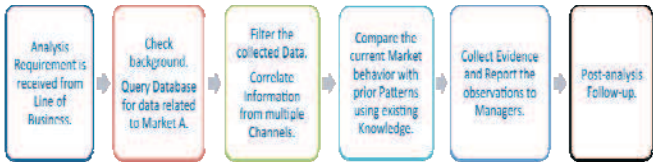

Figure 1. This is a typical risk analysis workflow observed and summarized from risk analysts in BOA.

be accepted to open a line of credit.

As shown in the analytical workflow (see Figure 1.), the risk analysts perform the credit analyses based on the following considerations:

- Requires for assessing the performance of credit products. For example, if the net loss of a credit product exceeds a certain threshold, the analysts need to evaluate the potential risk caused by that product.
- Shifts in general credit investment strategy. For instance, in the change of market focuses (i.e. focus shifts from high-end customer to mid-range consumers), most of the existing credit products would need to be re-evaluated in order to accommodate the changes. The analysts need to maintain an appropriate credit administration, measurement and monitoring process.
- Optimization of the portfolio of consumers or markets: e.g. analysts need to minimize credit loss and increase the product revenues. The analysts need to operate under a sound credit-granting process.

Thus, determining the risk in a consumer credit product relates to multiple factors that are intertwined with each other. While a set of queries can be easily used to retrieve delinquency patterns matching a limited set of existing hypotheses, risky market behaviors or hidden investment opportunities are more implicit and elusive. They are determined by a dynamic analysis context influenced by various factors, such as the markets' geospatial distributions, the prior investment strategies, and the competitors' behaviors. Hence, in the current practice, risk analysts often need to construct queries over multiple facets of a large body of credit data, searching for certain statistical values that may be indicative of high risk. Manually correlating these channels of information can be challenging and overwhelming: one type of analysis (e.g. only analyzing the known risks) would miss the hidden investment opportunity that could potentially lead to lost profits for the bank; whereas another type of analysis (e.g. only retaining portfolios with very low delinquency rates) could harm their relationships with their clients or regulatory agencies.

Exacerbating this challenge is the increasing size and complexity of the collected credit data that each analyst needs to examine. Millions of detailed consumer risk incidences are aggregated monthly to indicate the market behaviors, including information about the banks and their competitors, the state of the local economic environment, the consumer demographics, and the third party credit ratings. Thus, to comprehend the trends and patterns of a market and identify the investment opportunities, risk analysts are responsible for developing methods to utilize that large data corpus and predict the likelihood that a client who borrows money from a financial institution will default or fall behind on a loan payment. This places an extra burden on the individual analysis process in terms of the efforts needed to acquire the most appropriate information. At present, risk analysts do not have the capability to investigate all the patterns and activities, not to mention conducting thorough analyses of data over time.

In summary, credit risk analysis is an important yet challenging analysis process, due to the increasing size and complexity of consumer risk data. Effective analysis tools are needed to address these challenges. On the one hand, such tools must be tailored to specific credit risk management practices depending upon the nature and complexity of the business's credit analysis activities. On the other hand, they must enable comprehensive analysis of a large body of credit data to help verify known hypotheses, as well as allow discovery of hidden features in the dataset.

3 Related Work

In recent years the idea of using visualization to support financial analysis has gained a lot of interest. The forefront of visual analysis within the financial market can be categorized in several main sectors, namely temporal analysis, market analysis, and investment analysis. For temporal analysis, several efforts have previously investigated representing and analyzing changes in financial data [1, 2, 8]. Particularly, recent work by Ziegler et al. [20] presented useful clustering techniques to help visualize the temporal financial changes.

To support market analysis, Keim et al.[9] have presented pixel-based visualization techniques showing the performance of individual stocks in high detail. In addition, the tree-map-based "Map of the Market" by Wattenberg[17] had introduced an effective way to examine and compare different market and stock performances online.

Finally, many researchers have emphasized the integration of visualizations and mathematics in maximizing investment profits [11, 16]. A representative work in this sector is the work by Maciejewski et al. [13], which combined the Winner's and Loser's Curse to reach optimal investment decisions.

Given the complexity of risk analysis described in the previous section, these individual visualizations are not particularly suited for depicting multidimensional data on their own. However, these visualizations can be integrated into systems handling multidimensional data by using multiple coordinated views [12] in conjunction with highly interactive exploratory techniques. Systems for integrating large data source, such as JigSaw[15], PatentVis[10] and IRSV[19], enhance analysis tasks by providing an integrated visual environments for direct data manipulation.

While those systems provided feasible approaches for flexible data explorations, the situation in consumer risk analysis differs from such scenarios due to its time-critical nature, which requires a concise yet dynamic structural visual interface to effectively performance both tactical and strategic analyses on emerging risks. Examples for systems that intend to address such analysis are GTDVis [18] and WireVis[7]. Compared to GTDVis, which reveals global terrorist attacks patterns based on individual attributes, our system emphasizes revealing the combined impacts to the credit markets based on associating multi-channel risk information. While both RiskVA and WireVis present concise visualizations for large financial datasets, RiskVA also supports a customization mechanism to support the analysts to dynamically generate rules and to create analysis workspaces.

In the following sections, we detail the challenges of the current risk analysis process, and describe our visual analytics system, which is designed to address these challenges by encoding the essential analyses into a cohesive visual analysis environment.

4 IDENTIFYING CHALLENGES AND ANALYTICAL REQUIREMENTS IN CURRENT RISK ANALYSIS PROCESS

To design a visual analytics system that is tailored to the risk analytical workflow in Bank of America, we established a long-term collaboration with its Consumer Credit Risk Solution divisions and grounded our investigation on a field study conducted within the bank.

During the design phase of this project, we communicated with risk analysts from this division on their current practices as well as their needs for good analysis of consumer credit risks. We carried out multiple interviews and discussions to observe the day-to-day operations performed by the risk analysis team. The interviewees held a broad range of positions, including risk analysts who focused on analyzing the consumer credit products, managers who were in charge of business planning and crafting risk management strategies, and risk management architects who emphasized the identification of novel technologies that could be useful for risk analysis.

The interview data collected was used to characterize these analysts' task activities, and further used to develop the design requirements for a visual analytics system. Interviews with representatives from this team revealed the analytical needs of the risk analysts, including fusing multiple streams of data, retrieving information for context-dependent tasks, and analyzing their findings.

In general, the risk analysis team constantly needs to respond to market changes and conduct analyses involving the assessment of asset quality, the adequacy of provisions and reserves, and the balance of delinquency and investment. In addition, they are required to generate shared results effectively (e.g., a report of delinquency analysis or a summary of a market performance).

As specified in Figure 1, the analytical tasks of conducting risk analysis often include requirement specification, data aggregation, information organization and correlation, and result sharing. To analyze risk of a credit product, an analyst often starts by gathering relevant content from multiple data sources for a comprehensive view of that product. This aggregated dataset not only includes BOA's own data, such as the delinquency rating for that product (e.g. 30-day payments), but also data from credit rating agencies, such as credit scores. To improve their own

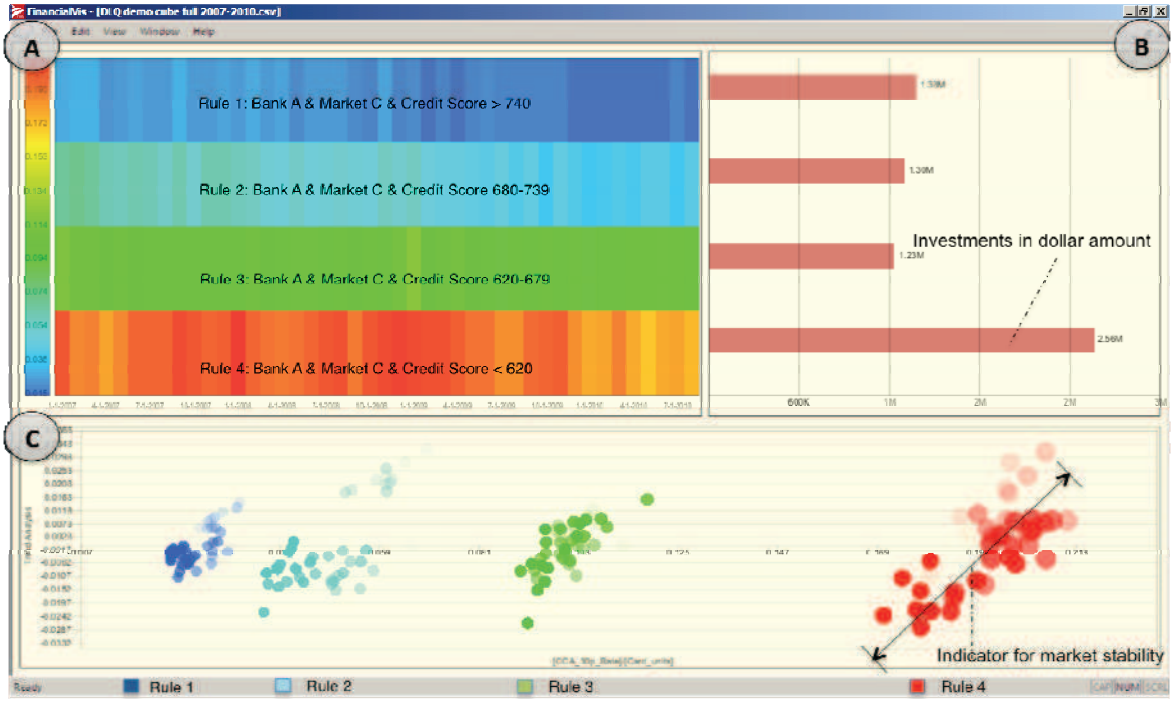

Figure 2.This is an overview of RiskVA, including the *Entity Heatmap* view (A), the *Product Comparison* view (B) and the *Trend Analysis* view (C). Here, analysts examine the 30-day delinquency rate (late payment) in Market C across a wide range of consumers.

assessments of risk, the analysts then filter this large collection of data and attempt to organize it in a clear and consistent manner to support the awareness and sense-making process.

Tools, in this context, are considered as a means to transform their hypotheses into desired task actions. Currently, the risk analysts primarily use tools, such as SQL databases, Excel and emails, to produce and communicate analysis related contents. In the process, the analysts' prior experiences (i.e. knowledge of a potentially deteriorating credit product) are used, and further task actions are taken to be used in their analytical process. Although these analysts currently use a number of different tools, we found that they were lacking tools actually designed to support their analysis workflows and provide the detailed information they need. This finding demonstrates the need for a tool that supports the users' analytical workflows and helps them effectively perform necessary analysis actions.

Therefore, the primary goal of our system is to address these challenges in accordance with the analytical requirements of the risk analysis team. A detailed characterization of these three challenges, as well as how they are addressed in our system are described in the following sections.

4.1 Support the Identification of Emerging Risks

Consumer risk analysis requires analysts to "know you customers". This includes knowing the performance of credit products for the customers individually, in commercial markets, and statistically. An important part of knowing the customers is the assessment of their overall activities in terms of risk. Certain credit products (e.g., credit loans, because they are unsecured debt without collateral) are inherently riskier than others. The corollary to "know your customers" is "know your investments." A financial institution must know where and how its credit investment is being spent in order to accurately assess the emerging credit risk. With the limited analytical tools available today, performing this comprehensive risk analysis is not easy.

A synthesis of credit information and investment knowledge from all relevant sources is therefore needed for risk analysis. As described earlier, risk analysts are demanding tools to support analysis of multi-channels of credit information. These tools must support not only data integration from multiple information channels, but also information correlation that brings together end products such as knowledge of delinquency ratings that represent the down shifting of a credit market.

Especially, the analysts require tools that focuses on the following aspects:

- **Market Analysis:** The changes and trends of a credit market have significant influences on the investment strategies for a bank. Typical markets are metropolitan areas where the bank focuses its credit products. The analyses of these markets concentrate more on performance than physical locations. According to risk analysts, examining the health of a credit market is a primary task for risk analysis.
- **Temporal Analysis:** By thoroughly analyzing the temporal changes of a credit product, the analysts can compute the deterioration rate of that product and its related net loss. In addition, a risk analyst can adjust their future product investments by assessing the outcomes under changing credit conditions or from previous behavior. Therefore, the ability to capture temporal information is of great value to risk analysis when assessing the risk of a credit product. However, temporal analysis in existing tools is limited to a per product basis. Having a complete picture of the fluctuation of products in all markets that could help risk analysts spot abnormal investment behaviors could be very beneficial.
- **Product Comparison Analysis:** Typically, the bank invests in a wide range of credit products to maximize its investment profit. While current query-based analysis could help risk analysts follow the changes of a credit product and typically on known relations, it only allows risk analyst to focus on a limited set of credit products. Given the diversity of the products, tools that could assist risk analysts comparing and comprehending the behaviors of these products while uncovering hidden relations would be helpful.

In practice, the risk analysts often examine a mixture of these analyses, such as depicting the trends in the markets through both market and temporal analysis, or examining a product's impacts to the markets using both market and product comparison analysis.

4.2 Support More Strategic than Tactical Analysis

Due to the burdensome necessity of analyzing multi-channel credit information, analysts often think narrowly about their investigative tasks. In particular, they think in terms of known patterns or in terms of activities that have been identified externally (e.g., from line of existing business or client) rather than in terms of what the patterns in the data are revealing to them.

In addition, consumer credit analysis requires the analysts to know the patterns of market changes for the customers individually, in different markets, and statistically. However, the analysis team is often hindered by the analytical tools, the time, or the sufficiency of evidence in thinking more broadly about the meaning of the credit products in terms of larger strategies to determine risks (especially previously unknown risks) and uncover hidden investment opportunities or their benefits to the overall banking strategy.

Therefore, careful thought must be given to gathering information to assist the analyst in providing objective, fully reasoned assessments backed by evidence and avoiding inevitable pitfalls and biases. Hence, analytical tools must be able to help the analyst effectively cut through noise and irrelevant data, explore the large body of credit data, and combine information from multi-channel information sources into a strategic risk analysis.

4.3 Support Individual Analysis Routine

Risk analysis is currently an art that is learned through long apprenticeship and then practiced. It is embedded in the creative reasoning processes of the practitioners. While Figure 1 illustrates some general analytical workflows in risk management **[3]**, specific credit risk management practices may differ among analysts depending upon the nature and complexity of their analysis goals and expertise. Sometimes, even the same analysts need to take alternative analytical practices just to accommodate the changes in focuses and priorities. Therefore, it is important for tools to support such diversified analytical needs and provide risk analysts with the flexibility to combine and sequence the analytical components to fit their own workflow.

5 RISKVA: A VISUAL ANALYTICS SYSTEM FOR ANALYZING CONSUMER CREDIT RISKS

In response to these identified analytical requirements for an integrated, efficient, analytics tool tuned to the consumer credit risk analysis environment, we designed RiskVA, an interactive visual analytics system that helps domain analysts in depicting and comparing the performance of the credit products by visually revealing market fluctuations and temporal trends of the targeted

credit products. Throughout the system design and implementation phase, we maintained close communication with the consumer credit risk solution group and routinely showed our progress and received feedback for our prototypes.

In the following sections, we first explain the data integration process that enables RiskVA to effectively combine multiple information channels, and then describe how each of the analytical requirements is depicted in our system. For privacy and proprietary reasons, details about the market information, bank associations and consumer information have all been anonymized in the following sections and in the figures. Of course, when the bank analysts use RiskVA, all this information is depicted.

5.1 Data Integration

At the heart of RiskVA is a data cube structure [14] that is customized to handle the large-size and complex credit data. Shown in Figure 3, the design of a three-dimensional cube structure was determined collectively with inputs from the risk team and provided them with a means to correlate multiple-channels of credit risks.

On the conceptual level, this data cube structure is specified to accommodate the risk analysts' tasks, presenting the rules of the desired credit information. In particular, such rules are constructed around the key elements in risk analysis, including the *entity* (e.g. FICO scores, wealth level, market ID, and etc.), *the temporal information*, and the *credit product variables* (e.g. credit loans, mortgage, and etc.).

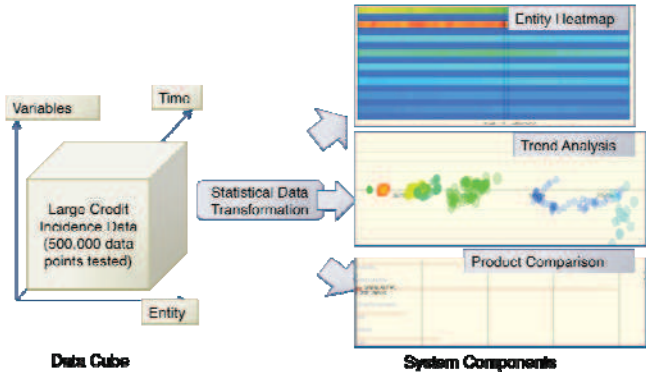

Figure 3. This is the system pipeline for RiskVA

The use of rule (e.g. market A & bank B & Time C) in the cube structure can effectively help risk analysis to navigate through the large body of high-dimensional credit data, and locate the desired information. For example, the rules can help to answer specific questions like, "What's the behavior of consumers with FICO score below 690 in market M?" In addition, the rules can be recorded and reused in different analysis scenarios, making it possible for RiskVA to trace individual's analytical process and share the analytical evidences between groups of analysts (section 5.3). The ease of creating rules provides the risk analysts with the flexibility to customize the credit information to fit with their own analytical workflow. RiskVA utilizes the rules to support the risk analysts in customizing their individual analytical environments (section 5.3).

Finally, the rules are an essential part in the constructions of the visual representation. They inform the views to filter unnecessary data elements and to present the analyst with most relevant information. RiskVA further uses these rules to coordinate and update the visualizations (section 5.4).

On the implementation level, this data cube follows the previous work [14] on creating an effective structure for slice-and-dice data from multiple aspects. As illustrated in Figure 3, to optimize the memory usage, the cube structure is disseminated into three parts: the meta-cubes that stores the rules, the virtual-cubes, which enable the comparison of credit products over a large dataset, and the physical data cube that points to the actual credit data. Due to page limits, the details of the implementation are beyond the scope of this paper.

5.2 Design Interactive Visualizations to Support the Identification of Emerging Risks

Since one cannot usually depict diverse rules, we need to visualize the activities of the corresponding markets in order to reveal the behaviors of the targeted credit products. Given the complexity of the related information, no single view could fulfil all the analytical requirements and show all the necessary data. Therefore, RiskVA is designed as system of coordinated views that would allow the analysts to see different data, while being able to understand the connections between the views easily. In particular, RiskVA encodes the three essential analyses as described in the following with a set of visualizations, each of which correspond to facilitate market analysis, temporal analysis and comparison analysis, respectively.

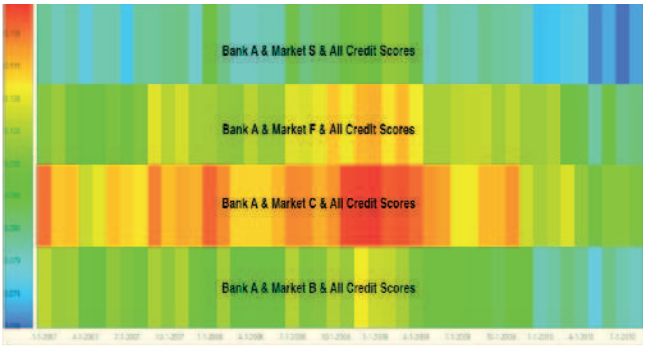

Figure 4. This is the overview of the *Entity Heatmap* view. Each Horizontal bar corresponds to a user-created entity rule.

5.2.1 Depict market behaviors with *Entity Heatmap* view

RiskVA utilizes entity heatmap to display the statistical measurements associating investment markets (e.g. population in the US West with credit scores less than 690) and credit products (e.g. mortgage or credit loans), as the former influences the behaviors of the latter. As shown in Figure 4, the heatmap is based on a grid where columns are the timestamps, and whose rows are rules that indicate the market performance of different credit products. This design aims to provide the analysts are direct sense of how the markets' performances over time.

At the intersection of a particular column and row, the cell is color-coded with a value derived from the combination of market/credit card in that time-period. Such values are associated with the market performance of a particular credit performance indicator, such as the 30-day delinquency rates or credit loans. Depending on the users' focus on the measurement displayed in the grid (e.g., difference value range or granularity of increments), RiskVA enables the user to interactively apply various color schemes to the visualization. At the beginning of each market analysis, the heatmap presents the overview of the targeted credit performance indicator, using a simple scheme where the color is computed based on the min/max values. In doing so, RiskVA aims to provide analysts a common picture that helps them

quickly recognize some significant market trends and changers, informing the overall strategies about the targeted markets.

Using the entity heatmap view, the risk analyst can then spot at a glance the markets that perform worse than the other ones, relating to a given credit product. As shown in Figure 4, the heatmap view is enhanced with a user-configurable rule capability that makes it possible for analysts to visually compare patterns of credit product behavior across different markets. For example, Figure 7 (right) shows the comparison of overall markets fluctuation between two financial institutions in the period of the 2007-2010. The color-coded heatmap (Figure 7 (right)) clearly indicates the performance differences between two institutions, where the lower delinquency rate (blue) suggests one has more stable credit products than the other.

In addition, once general understandings about a particular credit product are established, RiskVA would further facilitate the analyst in deepening his understanding of the impact of that credit product to the market; it enables the analysts to interactively parameterize the heatmap view with more focused rules, and helps them depict the market impacts with finer analytical context. In many cases, this helps analysts to disseminate the general market trends into multiple populations, and to adjust their investment strategies based on the risks associated with each group. As shown in Figure 7 (left), a further analysis of the lesser-performing institution indicates that the population with lower than 620 credit score and larger debt tends to pose higher risks than other populations.

We designed the heatmap view to be highly interactive. It enables the analysts to interactively select, highlight, and sort credit information; each of which is accompanied by detail tooltips. It also allows the analysts to apply statistical analysis over a particular time period or investment market for vertical and horizontal analysis (e.g. Compound Annual Growth Rate (CAGR) and standard deviation).

5.2.2 Reveal temporal patterns by *Trend Analysis* view

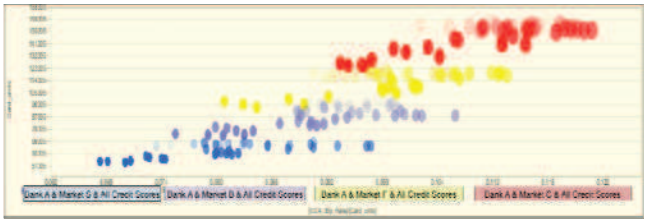

Figure 5. The overview of Trend analysis.

The ability to examine the stability of a credit market over a period of time is of great value to the analysts. Especially, when a market is turbulent, the analysts are required to identify the most vulnerable consumer credit products, and to report the possible cause of the market fluctuation. Through manually examining the changes of all the credit products prior to that time frame, the analysts may eventually be able to identify the weak line of the products line and report a plausible cause. However, current practices and tools still limit such crucial temporal analysis to analysis on a per product basis. Our coordinated visual interface enables the analysts to effectively identify these temporal patterns.

RiskVA utilizes the trend analysis view (Figure 5) to support the needed visualization of the market behaviors over time. It shows the overall temporal performances of each credit product, and allows the analysts to compare the stability of that product in different markets. The x-axis of this view shows the progression of time, and the y-axis shows the performance of each credit product in a particular market. Such performance can be actual investment amount of the product (e.g. total/average investment of mortgage loan), or it can be the trend that is calculated to indicate the stability of that product (i.e. the numerical differences between the current investment cycle and the last one).

As illustrated in Figure 2 (C), each dot represents the 30-day credit delinquency rate, and its transparency shows the temporal trail for that market. The more opaque, the closer that market is to the current time. If a market shares less drastic changes, such as the green dots to the left, it would reside in a more clustered group. On the other hand, if the market is like the one shown on the right in Figure 2 (C), where the red dots a less grouped, this pattern suggest the delinquency ratings (i.e. the dots) in that market have changed drastically over the years, indicating a larger fluctuation over the previous investment cycle.

To facilitate efficient interactions with the temporal analysis, the analysts can quickly select a specific market or product by hovering the labels. The analysts can further examine the details of a specific time period or certain product/market, through interactively filtering to the desired analysis items.

5.2.3 Compare products in *Product Comparison* view

Comparing performance between credit products plays an essential role in determining the bank's investment strategies. If a credit product continues to pose net loss in a particular market, this should be quickly identified and further inspected by risk analysts.

Figure 6. This is the *Product Comparison* View. Here, the analysts is comparing the investments distributions in two groups of consumers.

To present the comparisons between credit products, we use an interactive bar chart view as shown in Figure 6. In this view, the credit products involved in a particular market are grouped and represented in horizontal bars, which are directly associated with the rule in the heatmap view (see section 5.2.1). Different products are distinguished by the assigned colors (e.g. blue for mortgage and red for credit loans). The length of each bar corresponds to the actual investment values of that product, and can be further customized based on aggregation methods such as summation or average.

By placing the different credit products side-by-side, the analyst can directly compare the performances of these investments for different consumer populations. As shown in Figure 6, the analysts can clearly verify the well being of the targeted market, given the fact that the credit investments in that market are distributed proportionally to the consumers' credit scores (e.g. the bank invests more on consumers with higher credit scores and less on consumers with lower scores). To further examine the details of these credit products, the analysts can mouse over each bar and correlate it with the information in the heatmap.

5.3 Create Customizable Workspaces to Support individual analysis routines

To support the diversified analysis goals and user preferences, RiskVA presents the analysts with customizable analysis workspaces. Workspace is the main analysis environment of the RiskVA system. We define workspace as a user-configurable combination of the above three visualizations that provides a particular analysis context. Much like a probe system [5], each workspace builds upon the data cube structure and empowers the risk analysts to depict multiple-channels risk information at the same time. For example, the analysts can now examine and compare two different credit datasets for market performances in different analysis contexts. See Figure 7.) Moreover, RiskVA also handles workspace coordination. As detailed in the next section, it compares individual entities in different analysis context (e.g. credit load performance in US West v.s. US East). If there are matching entities (e.g. date, market, or rules), RiskVA can then link the multiple workspaces using this information and provide the analysts a coherent understanding about the risks that are embedded in different contexts.

During the analysis process, RiskVA allows the analysts to have full control of the creation of a workspace. Given their analytical needs, the analysts can directly branch off their analysis by duplicating the existing workspaces or construct a completely new workspace with a different set of data. Such a branching mechanism enables risk analysts to easily shift their focus while still maintaining their train of thought embodied in the branching structure. In addition, RiskVA also helps the analysts to customize their workspaces by providing them the flexibility to combine and sequence the analytical components to fit their own workflow (i.e. remove or hide certain visualizations or apply specific statistical analysis equations).

To help analysts maintain their train of thought and resume or repeat their previous analysis processes, RiskVA automatically records the history of the users' workspace usages. It logs the time when a workspace is created, branched off, and closed. It also records the above system customizations (user's view preferences during a particular analysis). Thus, at any given time, the analysts can use RiskVA to revisit their previous workspaces and to continue their analyses.

5.4 Coordinate Views to Support Strategic Analysis

The utilization of entity rules provides a foundation for RiskVA to coordinate between the three views and workspaces. Since the underlying structure for data cubes (e.g. meta-cubes and virtual-cubes) is the same, the information passing between these view representations becomes straightforward. Two levels of coordination mechanism are implemented in RiskVA currently, namely, within workspace coordination and between workspace coordination.

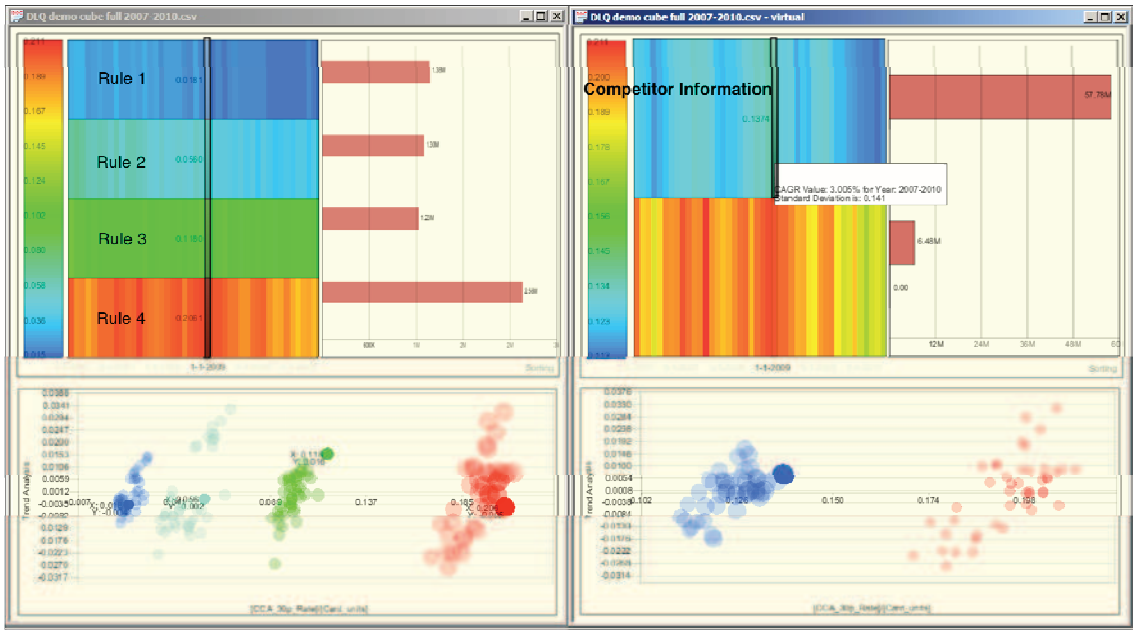

Figure 7. Two workspaces placed side-by-side. Here, analysts examine both workspaces together, looking for correlations between banks.

For within workspace coordination, RiskVA tightly coordinates the three views together so that updates in one view are immediately reflected in others. The analysts can now simultaneously interact with credit products, markets, consumers and etc., and monitor the correlation between all the information channels in a cohesive way.

For example, when an analyst mouses over a particular market in the trend analysis view, all related cells in the heatmap are highlighted to illustrate the consumer base for that market. At the same time, all the credit products that are associated with this market are also highlighted in the product comparison view, showing their performances and correlations.

For between workspace coordination, the analyst can hover over the dates (or another entity) in the heatmap view in a workspace, highlighting all the consumer information of that particular month (Figure 7.). All the views in the other workspace then highlight items involving the same entity in their own analysis context: the heatmap will highlight the related markets; the trend analysis will present the appearance dates; and the comparison view will show the credit product that may associate with these items.

This coordination is especially useful validating investment strategies in multiple market environments and verifying results in different analysis contexts. This coordination is made possible largely because of the shared properties that generally exist in many risk datasets at banks.

Both levels of coordination bring the risk analysts a cohesive depiction of the consumer risk market with coordination among

important views and correlation between different significant contexts.

6 ANALYSIS SCENARIOS

Identifying and understanding the cause of market deterioration is a key step for risk analysts to develop corresponding investment strategies. To conduct an unbiased risk analysis, it is necessary for analysts to monitor and compare credit products for both the competitors and their own. Based on our discussions with the risk analysis team at BOA (2 risk managers and 5 analysts), we have observed that the analysis of a market typically follows three analysis stages, namely identifying risky markets, comparing product performance between banks, and identifying potential causes for the fluctuation. The following scenario was identified with the risk analysis team at their regular strategy meeting. The agenda for this meeting was to discuss the bank's credit product performances in several major investment markets using RiskVA. As observers, we documented their analysis processes and helped them become familiar with the system. We also provided explanations about certain features in our system during their exploration.

To pursue this scenario, as shown in Figure 2, RiskVA was initialized with consumer credit data from year 2007 to 2010. To depict the performances of individual markets, the risk analysis team utilized the entity heatmap view to check if any interesting fluctuation pattern could be identified. The team found a set of markets with high 30-day delinquency rates (warmer colors in Figure 2 (A)) over the entire time span, which indicated a large body of late payments and the potential net losses for the bank. To get a clearer picture of the behavior of those markets, the analyst turned to the trend analysis view to examine the development of these markets. As illustrated by the cluster of red dots in Figure 2 (C), the analysts noticed that one investment market (red dots) was particularly vulnerable (i.e. the cluster scattered over the time), suggesting an unstable product performance since the beginning of the credit crunch in late 2007 until recently. Given that the general investment market is recovering since 2009, the analysis team decided to first take a closer look at unusual market behavior.

Instead of drilling down to that market, the team utilized the built-in workspaces to branch off their current analysis to keep track of their analysis processes. As shown in Figure 7, the team created a new workspace to compare the performance of the bank's own credit products with other competitors in that market. A quick glance at Figure 7 indicates that the competing banks on average invested more in that market and maintained a quite healthy performance. This finding immediately raised several questions: could the client base affect the market performance? Or was it caused by the unbalanced or sudden increase of investments in that market (e.g. mortgage v.s. credit loans)? Although these were all possible causes of the market deterioration, the risk analysts had no definitive answers or evidence to confirm their hypotheses by looking at the product comparison view alone.

Trying to verify these hypotheses, the team started to search for clues from the investment history of that market. By using fine-grain rules, they found that the investment in that market had always been a steady amount, and a reasonable proportion between secured and unsecured credit products. This therefore rules out the possibility if investments patterns being the cause of this fluctuation. However, a closer examination of client bases in the heatmap view suggested a different story. Figure 2 (B) showed that, compared to the typical strategy of pursuing consumers with higher credit scores, these markets invested on a fair amount of population with lower credit scores but with long credit histories. A quick check on the trend analysis view (Figure 5) further indicated that, the trend of this consumer group has peaked since 2008. A quick reference to the recent financial news around that market confirmed that there was an increasing amount of unemployment in market, which gave the risk analysts reasons to conclude that the changing in client bases may be a key factor in causing the fluctuation of this market.

Given the unsatisfactory market performance and concerns for losing more investment, the risk analyst team indicated that this market needed more attentions to bring its performance back on track. They also decided they needed some strategies to alleviate the pressure imposed on the customers and to help revive the market. After this exercise, the risk analyst commented on the effectiveness of RiskVA in helping them to explore the credit markets, as well as in identifying possible cause of market fluctuations. Although simple, the scenario has demonstrated the usefulness of RiskVA in support of strategic risk analysis.

7 EXPERT EVALUATION

To assess the efficacy of our system, we conducted expert evaluations with consumer risk analyst and risk managers from Bank of America. The goal for this process is to perform summative evaluation to measure how well RiskVA could facilitate the actual risk analyses.

During several on-site visits, we demonstrated design of the system and the utilities of the visualization to a total number of 8 risk managers and analysts from or related to the consumer risk analysis team. We invited risk analysts to perform in-depth analyses using the system, in a think-aloud manner. We observed and documented the details about their analysis processes. Finally, we concluded the evaluation by gathering their feedback and comments about our system. Since RiskVA has been deployed to this team, we also conducted email follow-ups to see if there were additional comments they would like to share with us.

7.1 Visual Facilitation on Tactical and Strategic Risk Analyses

One of the benefits of RiskVA that was noted by all risk analysts was its visual exploration environment that enables them to perform more strategic rather than tactical analysis. All the risk analysts consider being able to interactively perform all three analyses (i.e. market, temporal, and comparison analyses) at the same time to be powerful in portraying the detailed, dynamic nature of the emerging risks.

To support tactical analysis, RiskVA allowed the analysts to utilize their prior knowledge about the market to efficiently verify known risk patterns, and helped them to pursue their tactical goals using resources at hand. In RiskVA, rules were used to facilitate the analysts to interactively filter and analyze the credit information at different granularities.

When this capability was presented to the risk analysts, they spontaneously formulated a variety of rules to find credit information in the current market. All analysts were generally satisfied with the efficiency of using RiskVA and appreciated the flexibility to perform customized analysis. Specifically, one of the seasoned analysts pointed out that the ability to do such interactive analysis served two roles in supporting his tactical analysis. On the one hand, he considered RiskVA as an efficient method to "slice and dice" information to monitor market conditions and to test hypotheses. On the other hand, he thought the current implementation of RiskVA addressed another important aspect in risk analysis: the ability to verify and validate the accuracy of this new technology. Being able to interactively construct rule helped him match and confirm his prior

expectations and the visual representations, and gain confidence in using the system in his daily analysis.

In assessing its efficacy in handle strategic risk analysis, many analysts considered RiskVA to have the advantage to let the data tell the story about where risks are emerging. In particular, it reduced the amount of noise they have to sift through in order to see the broader picture and home in on suspicious outliers, enabling the analysts to explore possible risk patterns that were previously unidentified. They agreed that our tool assisted this analytical process by visually providing a global pattern as well as details on demand. As demonstrated in the scenario (see section 6), RiskVA helped risk analysts to effectively analyze their data across multiple dimensions and assisted them in determining the cause of market deterioration. All the analysts found the system practical, and believed that the system would be useful in helping them to perform more strategic risk analysis. As summarized in one of the analysts' comments, "[RiskVA] first provides me the general idea of what's going on with the market. And quickly and interactively let me navigate into a specific interested analysis segments. It allows me to get more hands-on analysis, and to check what I might have missed in analyzing the data".

One suggestion was to provide additional geospatial analysis, and display more regional risk entities. Analyst would like to see incorporated information like distances, densities, and areas in analysis as possible explanatory variables. One participant suggested that "some higher resolution geospatial view can be used to drill down below the census Bureau's Core Based Statistical Area (CBSA) level, or create a geographical segmentation of the US that's independent of CBSA, ZIP, or districts. Then we can compare our performance in these new geographic entities."

7.2 Customizable Workspaces and History Tracking

Using the workspace metaphor, RiskVA enabled the analysts to perform their tasks in a flexible and customizable environment. It provided the analysts with the flexibility to interactively combine and sequence different visualizations, customizing the workspace to fit their individual analysis routines. RiskVA utilized the workspace structures to sustain a dynamic analysis environment; it enabled the analysts to branch off their analysis at any time by duplicating any existing analysis workspaces. All the created workspaces were coordinated through the identification of similar market entities.

All participants appreciated the flexibility of the interface, finding it useful for customizing the system to only utilize the necessary visualizations in their analyses. In particular, they liked the ability to construct different workspaces to simultaneously analyze different markets groups. They thought this would be quite helpful in understanding the relationships between these markets. One of the managers commented that, "[RiskVA] brings the analyst at the center of the analysis, with their subjective attitudes, to interact with the data. This is where I see visual analytics can improve our risk management process."

Furthermore, RiskVA logged the analysis workspace history to enable the analysts to capture and revisit their previous analysis states. Most analysts found the idea of tracking analytical trails intriguing. While this was still a preliminary feature, the analysts had already noted its effectiveness in managing their diverse analysis practices.

8 DISCUSSION AND FUTURE WORK

We undertook this research to design a visual analytics system that facilitates the risk analysts' tactical and strategic consumer credit risk analyses. To this end, we presented RiskVA, an interactive visual analytics system that demonstrated unique and effective capabilities for a class of problems that involve complex consumer risk analyses.

The design of RiskVA is grounded in the task analysis results of a group of risk analysts from Bank of America. These results provide us clear identification of general domain analysis process, including fine-grain task activities, task flows, and overall analysis objectives. We further disseminated this general analysis process into individual analytical requirements (section 4), and transformed them into the specific system implementations through iterative prototyping with the risk analysts (section 5). Given the positive feedback from risk analysts, we found such design process to be particularly effective in designing an analysis-rich visual analytics system.

There are limitations to our research that should be addressed. Specifically, our research characterized the domain analytical workflow through interviews, which generally are self-reported by participants. Our research could also be limited, in that it modelled the analytical workflow from a retrospective viewpoint, whereas Brows et al. [21] demonstrated that problem spaces and solutions are established and changed dynamically in interactions with people and the environment. Therefore, our understanding of domain task flow may be constrained to the risk analyst' general way of performing tasks. One effective way we used to alleviate this constrain is enable the risk analysts to customize RiskVA to fit their own analytical workflows (section 5.3).

In addition, our research is limited by its evaluations with domain experts. Given the privacy and proprietary considerations in BOA, we only evaluated RiskVA through expert evaluations. While the results are positive, we believe much can be learnt if alternative methods were available. In particular, we would like to evaluate the risk analysts' knowledge gain from using our system. However, developing evaluation strategies to accurately assess the effectiveness of a visual analytics system is challenging yet beyond the scope of this paper. At this point we do not have a clear outline on the best evaluation approach; the design of guidelines for systematically evaluating a visual analytic system would be one interesting future direction for our research.

9 CONCLUSION

In this paper, we presented RiskVA, an interactive visual analytics system that demonstrated unique and effective capabilities for a class of problems that includes certain complex consumer risk analyses. RiskVA supports a thorough analysis of a financial institution's own data with data from other sources, including competitors, for a comprehensive view will permit these institutions to better make their own assessments of risk, independent of, and more focused than, assessments they get from ratings agencies that may not be suitable for the their situations.

By placing risk analysts in the center of their analytical processes, RiskVA provides analysts with customizable analysis workspaces, interactive data exploration, and the capability to correlate information over a large corpus of credit data. In our expert evaluations, risk analysts confirmed the novelty and utility of RiskVA to facilitate them in performing in-depth risk assessments, and further expressed interests in using it in their daily tasks. With such encouraging feedback, we are current deploying RiskVA to the consumer credit risk division in BOA.

These results indicate the efficacy of the cognitive task analysis process we undertook at the beginning. It is essential that this task analysis be carried out and also essential that the task analysis be made flexible to support exploration and unforeseen analyses. RiskVA and other tools we have developed indicate that this process is general. It is certainly clear that RiskVA is a tool

for emerging risk analysis that is applicable across financial institutions.

REFERENCES

[1] J. Allebach. Binary display of images when spot size exceeds step size. *Applied Optics*, 15:2513–2519, August 1980.

[2] M. Ankerst, D.A. Keim, H.-P. Kriegel, Circle Segments: A Technique for Visually Exploring Large Multidimensional Data Sets, in Proc. of Visualization 96, Hot Topics Session, 1996

[3] Basel Committee for Banking Supervision. www.bis.org/bcbs

[4] Nick Bloom (2007),'The Impact of Uncertainty Shocks',National Bureau of Economic Research Working Paper No. W13385,also available as CEP Discussion

[5] T. Butkiewicz ; W. Dou; Z. Wartell; W. Ribarsky; R. Chang, "Multi-Focused Geospatial Analysis Using Probes," Visualization and Computer Graphics, IEEE Transactions on , vol.14, no.6, pp.1165-1172, Nov.-Dec. 2008

[6] E. Catmull. A tutorial on compensation tables. In *Computer Graphics*, volume 13, pages 1–7. ACM SIGGRAPH, 1979.

[7] R. Chang.; M. Ghoniem.; R. Kosara; W. Ribarsky,; Y. Jing ; E. Suma, C. Ziemkiewicz,; D. Kern; A. Sudjianto, "WireVis: Visualization of Categorical, Time-Varying Data From Financial Transactions," Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on, vol., no., pp.155-162, Oct. 30 2007-Nov. 1 2007

[8] D.A. Keim, Pixel-Oriented Visualization Techniques for Exploring Very Large Databases, Journal of Computational and Graphical Statistics, Vol 5, pp. 58 -77, 1996

[9] D.A. Keim, T. Nietzschmann, N. Schelwies, J. Schneidewind, T. Schreck, and H. Ziegler. A Spectral Visualization System for Analyzing Financial Time Series Data. in Proc. of Eurographics/IEEE-VGTC Symposium on Visualization (EUROVIS'06), pp. 195-200, 2006

[10] S. Koch, H. Bosch; M. Giereth, T. Ertl; , "Iterative integration of visual insights during patent search and analysis," Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on , vol., no., pp.203-210, 12-13 Oct. 2009

[11] J. Lin, M. Vlachos, E. Keogh, and D. Gunopulos, Iterative Incremental Clustering of Time Series, Advances in Database Technology - EDBT 2004, 9th Intl. Conference on Extending Database Technology, 2004

[12] J. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV '07, pages 61–71, July 2007.

[13] A. Savikhin, R. Maciejewski and David S. Ebert. Applied Visual Analytics for Economic Decision-Making. IEEE Symposium on Visual Analytics Science and Technology (VAST), pp. 107-114, 2008.

[14] C. Stolte, D. Tang, and P. Hanrahan. 2003. Multiscale Visualization Using Data Cubes. IEEE Transactions on Visualization and Computer Graphics 9, 2 (April 2003), 176-187.

[15] John Stasko, Carsten Gorg, and Zhicheng Liu, "Jigsaw: Supporting Investigative Analysis through Interactive Visualization", Information Visualization, Vol. 7, No. 2, Summer 2008, pp. 118-132

[16] S. Rudolpn, A. Savikhin, and D.S. Ebert,; FinVis: Applied visual analytics for personal financial planning, Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on , vol., no., pp.195-202, 12-13 Oct. 2009

[17] M. Wattenberg. Visualizing the stock market. CHI Extended Abstracts on Human Factors in Computing Systems, pp. 188 189, 1999

[18] X. Wang, E. Miller., K. Smarick, W. Ribarsky, and R. Chang (2008), Investigative Visual Analytsis of Global Terrorism. Computer Graphics Forum, 27:919-926

[19] X. Wang, W. Dou, S. Chen, W. Ribarsky, and R. Chang, "An Interactive Visual Analytics System for Bridge Management", presented at Computer. Graph. Forum, 2010, pp.1033-1042.

[20] H. Ziegler; M. Jenny; T. Gruse; D.A. Keim.; Visual market sector analysis for financial time series data, Visual Analytics Science and Technology (VAST), 2010

[21] J. S. Brown and P. Duguid. Organizational Learning and Communities-of-Practice: Toward a Unified View of Working Learning, and Innovation. Organization Science, 2(1):40–57, 1991.