TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 07/11 pp440-451
Volume 17, Number 4, August 2012

RiskVA: A Visual Analytics System for Consumer Credit Risk Analysis

Xiaoyu Wang**, Dong Jeong, Remco Chang, William Ribarsky

Charlotte Visualization Center, UNC Charlotte, Charlotte, NC 28223, USA

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 analyses 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 this process, analysts are facing challenges in monitoring the data, comparing its geospatial and temporal patterns, and developing appropriate 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 interactive data exploration and correlation, and visually facilitates depictions of market fluctuations and temporal trends for a targeted credit product. When evaluated by analysts from Bank of America, RiskVA was appreciated for its effectiveness in facilitating the bank's risk management.

Key words: risk management; visual analytics; human computer interaction; design study

Introduction

Credit risk management, in general, refers to the process in which the investors assess the risk of loss arising from a borrower who does not make payments as promised[1]. 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 will be profitable yet sustain the bank's financial stability.

Therefore, the assessment of credit risk is crucial to banks in positioning themselves to profit through balancing credit investments and returns; it is also 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 crises, Bloom[2] 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 effective risk analyses to minimize the net credit loss resulting from its credit products, and to determine profitable market strategies. Given BOA's wide range of credit investments, however, analyzing risk in such diversified portfolios has become an overwhelming process. This demands that analysts evaluate credit risks both

Received: 2012-05-01; revised: 2012-06-27
**To whom correspondence should be addressed.
E-mail: xwang25@uncc.edu; Tel: 1-704-687-7741

temporally (i.e., identifying market turning points before and after the recession) and across credit markets (e.g., comparing product performances in major cities). 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 effort 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. Current analytical tools are not designed specifically to address these challenges and can't help analysts comprehend the trends and patterns of markets from the necessary analytical perspectives.

To alleviate these challenges, we formed a research partnership with Bank of America to investigate novel analysis technologies. One of our first actions under this partnership was to observe and characterize the analysts' workflow. Such 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.

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 analysts in depicting and comparing the performance of the credit products via visually revealing market fluctuations and temporal trends of the targeted credit products. RiskVA further supports individual analysis workflow and allows the analysts to customize the visual interface based on their preferences.

To evaluate the efficacy of the system, we have conducted expert evaluations with analysts from BOA. We found that most analysts considered RiskVA to be useful and complimentary to their analysis needs. We further identified analysis scenarios for which our system could provide analysts with insights for developing 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 scalable, multi-channel risk-related data over any given period of time.
- It provides highly coordinated interactive visualizations to enhance both tactical and strategic analyses that are essential in depicting emerging risks.
- It provides customizable workspaces that support the individual analyst's analysis routines.

1 Domain Characterization: Credit Risk Analysis and Its Challenges

In this section, we describe the current practices in credit risk analysis. By discussing these in detail, we intend to shed light on the characteristics of existing challenges and how an interactive visual analytics system can provide substantial improvements in this domain.

The importance of credit risk analysis, as part of financial risk analysis, originates from the New Basel Capital Accord. Credit risk analysis generally refers to the process of establishing 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[1]. Financial companies typically use risk models to evaluate the insolvency risk caused by credits that enter into default. 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 modeled strategy is then applied to new consumers to determine whether they are above the cut-off level to receive a line of credit.

As shown in Fig. 1, the credit risk analyses workflow is based on the following considerations:

- Requirement on the assessment of performances for individual credit product. For example, if the net loss of a credit product exceeds a certain threshold, the analysts need to evaluate the associated potential risk.
- Shifts in general credit investment strategy. For instance, in the change of market focus (i.e., focus shifts from high-end to mid-range consumers), the existing credit products would need to be re-evaluated in order to accommodate the changes.
- Optimization of the portfolio of consumers or markets: e.g., analysts need to minimize credit loss

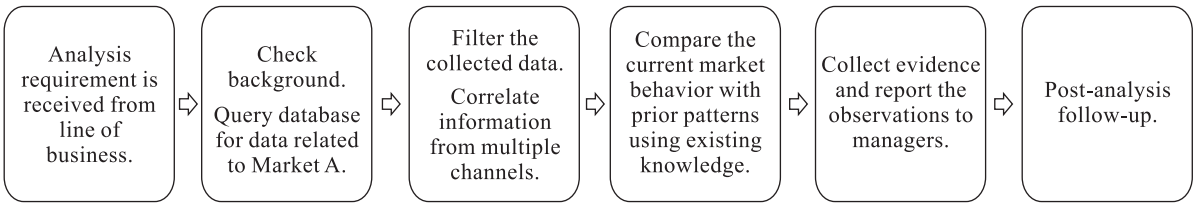

Fig. 1 A typical risk analysis workflow observed and summarized from risk analysts in BOA.

and increase the product revenues. The analysts need to operate under a sound credit-granting process.

Thus, it is obvious that the determination of the credit risk relates to multiple intertwined factors. While a set of SQL queries can be used to retrieve delinquency patterns matching a limited set of existing hypotheses, inspecting risky behaviors or hidden investment opportunities are less straightforward. 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, risk analysts often need to construct elaborate queries over multiple facets of a large body of credit data, searching for certain statistical values that indicate high risk. Manually correlating these information channels is overwhelming: a single-focused analysis that only analyzes the known risks might hidden an investment opportunity; whereas only retaining portfolios with very low delinquency rates could harm their relationships with 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 incidences are aggregated monthly to represent the market behaviors. This includes information about the banks and their competitors, 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 analysis process in terms of the effort needed to acquire the most appropriate information. At present, analysts do not have the capability to investigate all the activities, not to mention conducting thorough temporal analyses.

In summary, credit risk analysis is an important yet challenging process, due to the increasing size and complexity of the data. Effective analysis tools are needed to address these challenges. Such tools must be tailored to specific analysis practices based upon the nature of the credit data. The tools must also enable comprehensive analysis of a large body of data to help verify hypotheses as well as allow discovery of hidden features.

2 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[3-5]. Particularly, recent work by Ziegler et al.[6] presented useful clustering techniques to help visualize the temporal financial changes.

To support market analysis, Keim et al.[7] have presented pixel-based visualization techniques showing the performance of individual stocks in high detail. In addition, the treemap-based Map of the Market by Wattenberg[8] introduced an effective way to examine and compare market and stock performances.

Finally, much research emphasized the integration of visualizations and mathematics in maximizing investment profits[9,10]. A representative work by Savikhin et al.[11] combines the Winner's and Loser's Curse to reach optimal investment decisions.

Given the complexity of the aforementioned risk analysis, these individual systems 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 sources, such as JigSaw[13], PatentVis[14], and IRSV[15], enhance analyses by providing an integrated visual environment for direct data manipulation.

Moreover, another significant aspect in consumer risk analysis is its time-critical nature, which requires a concise yet dynamic visual interface to effectively perform both tactical and strategic analyses on emerging risks. Examples for systems that intend to address such analysis are GTDVis[16] and WireVis[17]. Compared to GTDVis, which reveals global terrorist attack patterns based on individual attributes, our system emphasizes revealing the combined impact 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 dynamic rule generation and workspace creation.

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.

3 Identifying Analytical Requirements in Risk Analysis Processes

To design a visual analytics system that is tailored to the risk analytical workflow at Bank of America, we established a long-term collaboration with its Consumer Credit Risk Solution division and grounded our investigation on a field study conducted within the bank. During the design phase, we communicated with 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 team. The interviewees held a broad range of positions, including analysts who focused on analyzing credit products and risk managers who were in charge of business plans and management strategies.

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

As specified in Fig. 1, the analytical tasks of conducting risk analysis often include requirement specification, data aggregation, information organization and correlation, and result sharing. To analyze the 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 (late-payment) rating for a product, but also data from credit rating agencies, such as credit scores. To improve their own 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, like SQL databases and Excel, to produce and communicate analysis outcomes. In the process, their prior experiences (i.e., knowledge of a potentially deteriorating credit product) are used, and further task actions are taken 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 help them effectively perform necessary tasks.

Therefore, the 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, is described as follows.

3.1 Support the identification of emerging risks

Consumer risk analysis requires analysts to "know your 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., loans as they are unsecured debt without collateral) are inherently riskier. A financial institution must know where and how its credit investment is

being spent in order to accurately assess the emerging risk. Yet, performing this comprehensive risk analysis with the less tailored tools is not easy. Specifically, the analysts require tools that focus on the following aspects:

Market analysis The trends of a credit market have significant influence on a bank's investment strategies. Typical markets are metropolitan areas where the bank focuses its credit products. The analyses of these markets concentrate more on the performance across geospatial regions and the health of a credit market.

Temporal analysis The temporal changes of a credit product are of great value in revealing the deterioration rate of a credit product and its related net loss. Risk analysts rely on this analysis to adjust their future product investments. However, temporal analysis in existing tools is limited to a per-product basis. Having a complete picture of the temporal changes of products in all markets will help spot abnormal investment behaviors and is very crucial to analysts.

Product-comparison analysis The bank invests in a wide range of credit products to maximize its investment profit. While current query-based analysis can help follow the changes of a single product, it has difficulties in scaling up the analysis. Given the diversity of the products, tools that could assist in comparing the behaviors of these products would be very significant.

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 impact on the markets using both market and product comparison analysis.

3.2 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 activities that have been identified previously rather than in terms of what the patterns in the data are revealing to them. In addition, credit analysis requires knowing the patterns of market changes for individual clients as well as the entire market. However, the analyses are often hindered by the tools, the time, or the sufficiency of evidences in revealing the more broad nature of credit products.

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 effectively sift through noise and irrelevant data, explore a large body of credit data, and combine information from multi-channel information sources into a strategic risk analysis.

4 RiskVA: A Visual Analytics System for Analyzing Credit Risks

In response to these identified analytical requirements, 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.

4.1 Data integration

At the heart of RiskVA is a data cube structure[18] that is customized to handle the large-size and complex credit data. Shown in Fig. 2, our design of a three-dimensional cube structure provided our collaborators with a means to correlate multiple credit information.

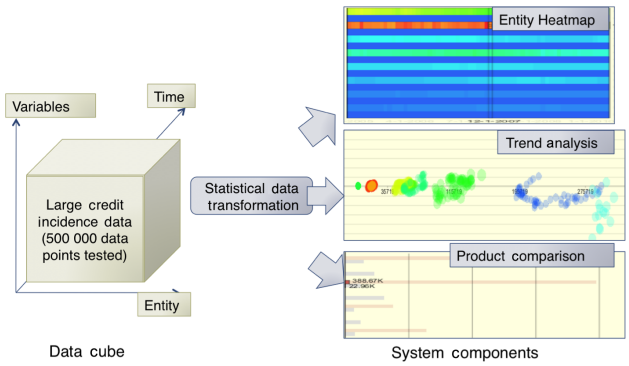

Fig. 2 The system pipeline for RiskVA.

On the conceptual level, our cube structure is specified to accommodate the risk analysts' tasks, presenting the rules of the desired credit information. Particularly, such rules are constructed around the key analysis elements, including credit entity (e.g., FICO scores, wealth level, and market ID), temporal information, and credit products (e.g., credit loans and mortgage).

The use of rules (e.g., market A & bank B & time C) in the cube structure can effectively help analysts navigate through a large body of high-dimensional data and locate the desired information. For example, the rules help answer questions like, "What's the consumer behavior in market M, with credit score below 690?" In addition, rules can be recorded and reused in different analysis scenarios, making it possible for RiskVA to trace an individual's analytical process and share the outcomes between analysts. The ease of creating rules provides the flexibility to customize the credit information and fit it with an individual analyst's workflows. 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 the most relevant information. RiskVA further uses rules to coordinate and update the visualizations.

On the implementation level, this data cube follows the previous work[18] on creating an effective structure for slice-and-dice data from multiple aspects. As illustrated in Fig. 2, to optimize the memory usage, the cube structure is disseminated into three parts: the meta-cubes that store 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.

4.2 Design interactive visualizations to support the identification of emerging risks

Given the complexity of the related information, no single view could fulfill 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 with a set of visualizations, each of which corresponds to facilitate market analysis, temporal analysis, and comparison analysis.

4.2.1 Depict market behaviors with Entity Heatmap view

RiskVA utilizes an 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, as the former influences the behaviors of the latter. As shown in Fig. 3, 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 a direct sense of how the markets' perform 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 and credit card information in that time-period. Such values are associated with a particular credit performance indicator (e.g., late-payment rates). Depending on the choice of mathematical measurements (e.g., difference value granularity), RiskVA enables the user to interactively apply various color schemes (e.g., colors based on the min/max values) to depict some significant market trends, informing the overall strategies about the targeted markets.

Using the heatmap view, risk analysts can now spot the markets that perform worse than the others, relate to a given credit product. As shown in Fig. 4, this view is also enhanced with user-configurable rules that make it possible for analysts to visually compare patterns of product behaviors across different markets. For example, Fig. 3 (right) shows the comparison of overall market fluctuations between two financial institutions in the period of 2007-2010. The heatmap view clearly indicates the performance variation between two institutions, where the lower late-payment rate (blue) suggests one has more stable products than the other. The heatmap view is also highly interactive and enables interactive selection, highlights, and sorting of credit information. It further allows analysts to apply statistical analysis over a particular time period or investment market for vertical and horizontal analysis.

Once general understandings about a particular credit product are established, RiskVA further facilitates in deepening the understanding of the impact of that credit product on the market. It enables analysts to interactively parameterize the heatmap view with more focused rules, and helps them depict the market impacts

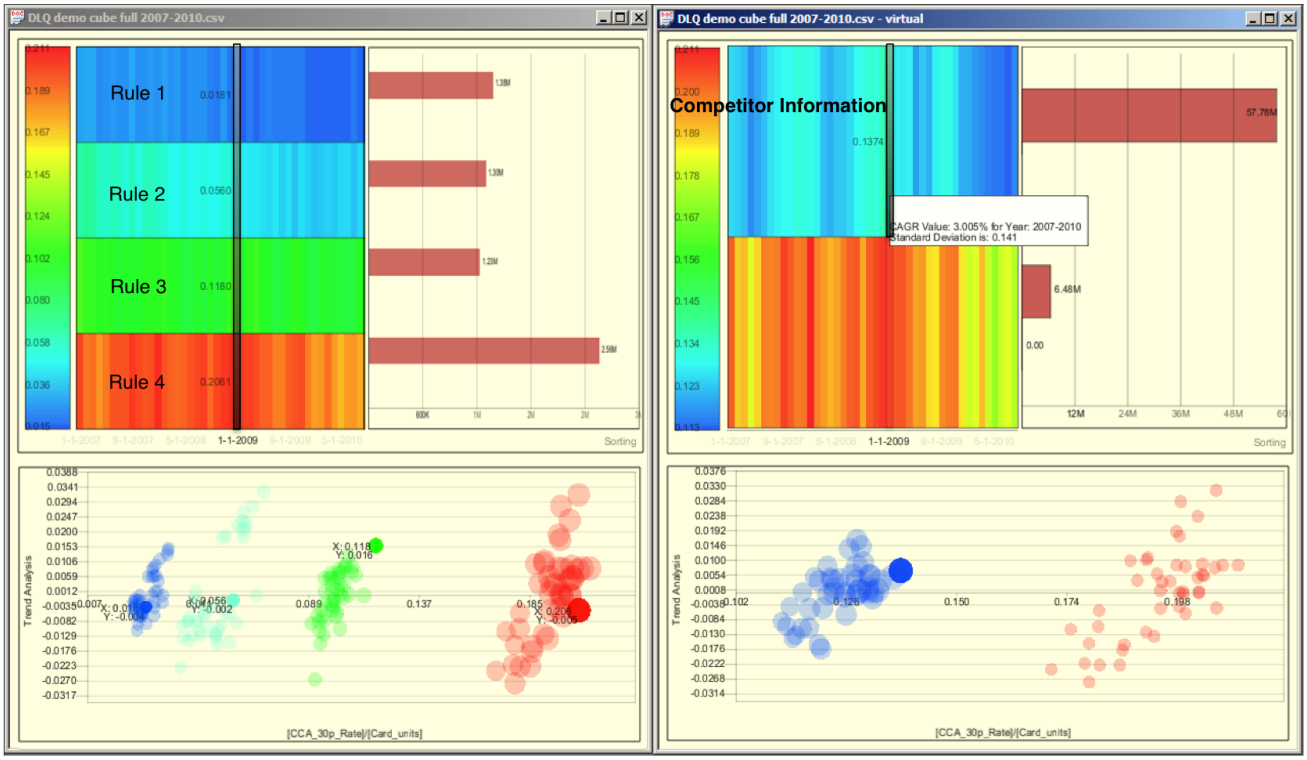

Fig. 3 Comparison analysis between information from two customized workspaces.

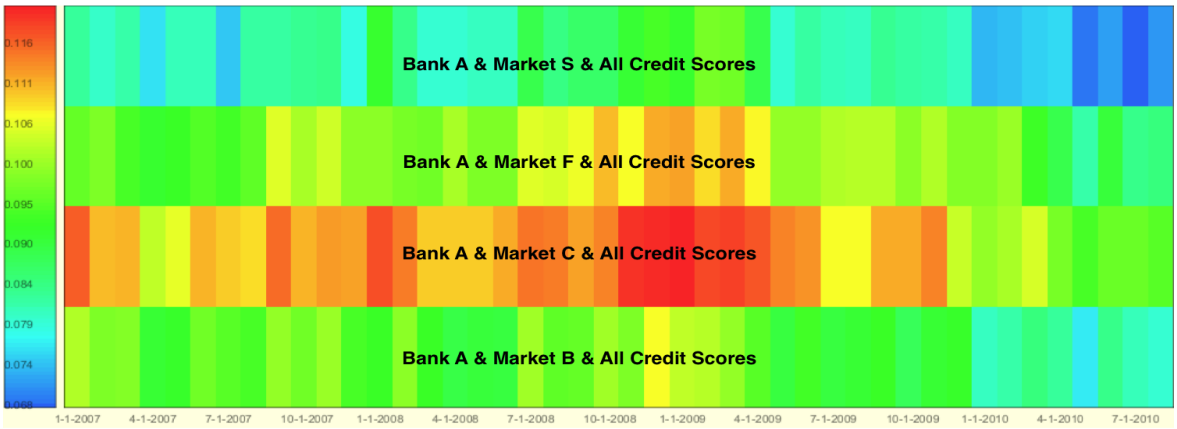

Fig. 4 The overview of the Entity Heatmap view. Each horizontal bar corresponds to a user-created rule.

with finer analytical context. In many cases, this enables the dissemination of market trends into multiple populations, revealing the strategies based on the risks associated with each population. Figure 3 (left) shows a further analysis of the lesser-performing institution that indicates the population with less than a 620 score and a larger debt tends to pose higher risks than others.

4.2.2 Reveal temporal patterns by trend analysis view

The ability to examine the stability of a credit market over a period of time is of great value to the risk analysts, especially during a market turbulence. This can help identify the most vulnerable consumer credit products and depict the possible causes. Through manually examining the changes of all the credit products prior to that time frame, the analysts may eventually be able to identify a weak product. However, current practices still limit such crucial temporal analysis to a per-product basis. RiskVA utilizes a coordinated interface to effectively reveal these temporal patterns.

The trend analysis view (Fig. 5) is designed to support such analysis. 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 shows the progression of time, and the y-axis shows the performance of a product in a particular market. Such performance can be actual investment amount (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 investment cycles).

In particular, each bubble in Fig. 6C represents a 30-day credit late-payment rate; its transparency shows the temporal trail for that market. The more opaque, the closer that data is to the current time. If a market has less drastic changes, such as the green bubbles to the left, the pattern resides in a more clustered group. However, if the market is like the one shown on the red dots in Fig. 6C, as it is less grouped, this market is less stable. Specifically, the late-payment ratings (i.e., the dots) in that market changed drastically over the years.

To facilitate efficient interactions with the temporal analysis, RiskVA enables quick selection of a specific market or product by hovering over the labels. Analysts can further examine the details of a specific time period and/or certain product/market, through interactively filtering to the desired analysis items.

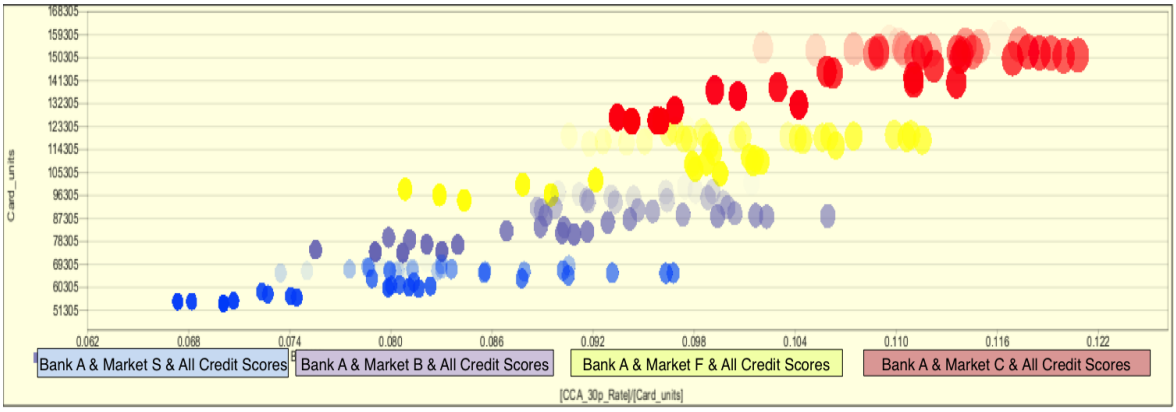

Fig. 5 The overview of trend analysis.

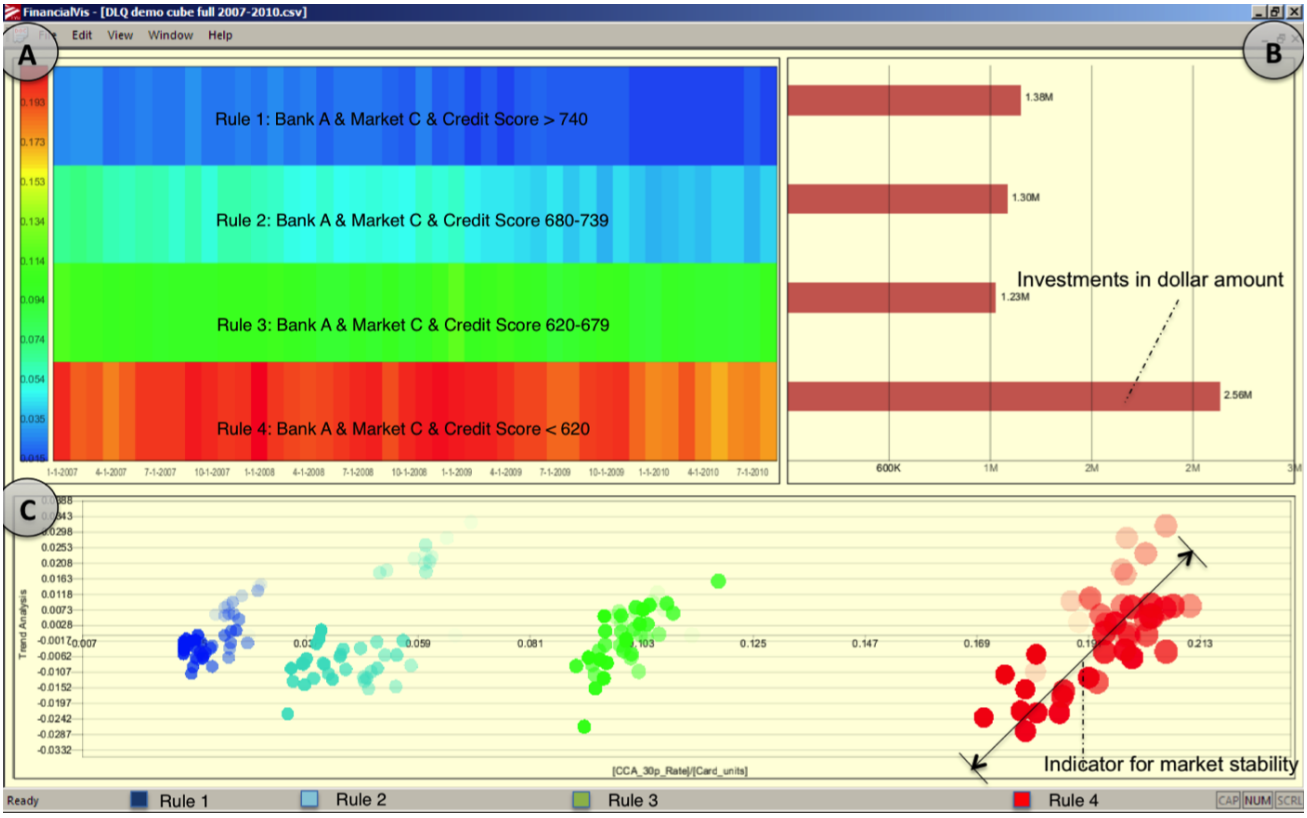

Fig. 6 An overview of RiskVA, including the Entity Heatmap view (A), the Product Comparison view (B), and the Trend Analysis view (C). Here, analysts examine the 30-day delinquency rate (late payment) in Market C across a range of consumers.

4.2.3 Compare products in product-comparison view

RiskVA uses an interactive bar chart view to present the comparisons between credit products (Fig. 7). In this view, the credit products are grouped and represented in horizontal bars, which are directly associated with the rule in the heatmap view. Different products are distinguished by the assigned colors. 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 products side-by-side, RiskVA enables direct comparison between the performances of these investments for various populations. As shown above, 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.

4.3 Create customizable workspaces to support individual analysis routines

RiskVA further presents customizable workspaces to support the diversified analysis goals. We define a workspace as a user-configurable combination of the above three visualizations that provides a particular analysis context. Much like a probe system[19], each workspace builds upon the data cube structure and empowers the ability to depict multiple-channel information simultaneously. For example, the analysts can examine and compare two market performance datasets via different analysis contexts (see Fig. 3). In addition, 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. RiskVA handles two levels of coordination mechanism, within workspace coordination and between view coordination. This helps the cross-comparison of individual products in different analysis contexts (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 further link the multiple workspaces using this information and provide analysts a coherent understanding about the risks that are embedded in different analysis contexts.

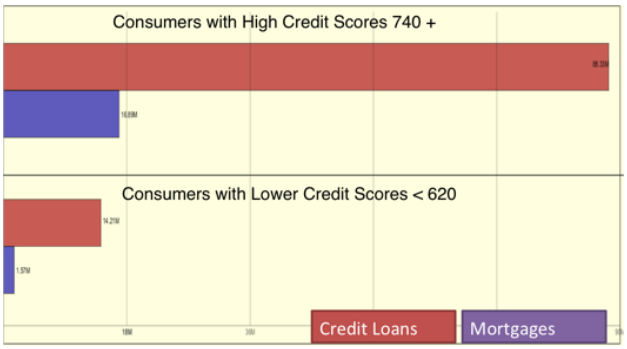

Fig. 7 The product-comparison view. This shows the comparison between two invested consumer groups.

RiskVA also enables the analysts to have full control of the creation of a workspace. Given their analytical needs, analysts can directly branch off their analysis by duplicating the existing workspaces or by constructing a completely new workspace with a different set of data. Such a branching mechanism enables them to easily shift focus while still maintaining the 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 Analysis Scenarios

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 from both the competitors and from themselves. Based on our discussions at BOA, 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 Fig. 6, RiskVA was initialized with 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 Fig. 6A) 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 clusters in Fig. 6C, the analysts noticed that one investment market (red bubbles) 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 directly 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 Fig. 3, 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 Fig. 3 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 of investments patterns being the cause of this fluctuation. However, a closer examination of client bases in the heatmap view suggested a different story. Figure 6B showed that, compared to the typical strategy of pursuing consumers with higher credit scores, these markets invested in a fair amount of population with lower credit scores but with long credit histories. A quick check on the trend analysis view (Fig. 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 the market, which gave the risk analysts reason 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 attention 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 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.

6 Expert Evaluation: Visual Facilitation on Tactical and Strategic Analyses

To assess the efficacy of our system, we conducted expert evaluations with end-users from Bank of America. Our goal was to perform summative evaluation to measure how well RiskVA could facilitate their analyses. During several on-site visits, we demonstrated the design of RiskVA and the utilities of each visualization to a total number of 8 managers and analysts from or related to the risk analysis team. We invited analysts to perform their standard analyses in a think-aloud manner using RiskVA. We observed and documented the details about their analysis

processes. Finally, we concluded the evaluation by gathering their feedback and comments. 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.

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 analyses. All analysts considered the ability to interactively perform the aforementioned three analyses simultaneously is powerful in portraying the detailed, dynamic nature of emerging risks. To support tactical analysis, RiskVA allows the analysts to utilize their prior knowledge about the market to verify known risk patterns. It further helps pursue analysts' tactical goals using resources at hand. In RiskVA, rules are used to facilitate the interactive filtration and analysis of the credit information at different granularities.

When this capability was presented to the risk analysts, they spontaneously formulated a variety of rules to identify credit information. They were generally satisfied with the efficiency and appreciated the flexibility to perform customized analysis. Specifically, one of the seasoned analysts commented that the ability of such interactive analysis served two roles in supporting his tactical analysis. On the one hand, he considered RiskVA an efficient environment to slice and dice market information and to test hypotheses. On the other hand, he thought the ability to interactively construct rules helped him match and confirm his expectations with the visual representations, and gain confidence in using the system. This helped addressing another important aspect in risk analysis practices: the need to verify and validate the accuracy of new technologies.

In assessing its efficacy in handling strategic risk analysis, many analysts considered RiskVA to be advantageous in identifying the emerging risks. Particularly, it reduced the amount of data noise and helped analysts to see a broader picture and home in on suspicious outliers. It further enables them to explore possible risk patterns that were previously unidentified. As demonstrated in the above scenario, RiskVA helped effectively analyze sizable credit data and lead to the depiction of the cause of market deterioration. All the analysts found the system practical and believed it would be useful in performing more strategic analysis. As summarized in one comment, "RiskVA" first provides me the general idea of what's going on with the market. And quickly let me navigate into a specific 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 with more regional risk entities. Analysts would like to see incorporated information like 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 or ZIP. Then we can compare our performance in these new entities."

7 Discussion and Future Work

We undertook this research to design a system that facilitates the risk analysts' tactical and strategic analyses. The design of RiskVA is grounded in the task analysis results of a group of analysts from Bank of America. These results provide us clear identification of general domain analysis processes, including fine-grain task activities, task flows, and overall analysis objectives. We further disseminated this general analysis process into individual analytical requirements, and transformed them into the specific system implementations through iterative prototyping with the analysts.

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. This method could be limited, in that it modeled the workflow from a retrospective viewpoint, whereas Brows and Duguid[20] 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 analysts' general way of performing tasks. One way we used to alleviate this constraint is enabling the risk analysts to customize RiskVA to fit their own workflows.

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 analysts' knowledge gain from using our system. However, developing evaluation strategies to accurately assess the effectiveness of a visual analytics system is challenging and 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.

8 Conclusions

In this paper, we presented RiskVA, an interactive visual analytics system that demonstrated effective capabilities for credit risk analyses. RiskVA supports a thorough analysis of sizable credit data from financial institutions and permits them 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 their situations. By placing analysts in the center of the processes, RiskVA provides an interactive environment to perform data exploration and to correlate information over a large corpus of data. In our expert evaluations, analysts confirmed the novelty and utility of RiskVA in performing risk assessments, and further expressed interest in using it in their daily tasks. 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 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 and is applicable across financial institutions.

References

[1] Basel Committee for Banking Supervision. http://www.bis.org/bcb, 2012.

[2] Bloom N. The impact of uncertainty shocks'. National Bureau of Economic Research working paper No. W13385, CEP Discussion, 2007.

[3] Allebach J. Binary display of images when spot size exceeds step size. *Journal of Applied Optics*, 1980, **15**: 2513-2519.

[4] Ankerst M, Keim D, Kriegel H. Circle segments: A technique for visually exploring large multidimensional data sets. In: Proc. of Visualization 96, Hot Topics Session. San Francisco, CA, USA, 1996.

[5] Keim D A. Pixel-oriented visualization techniques for exploring very large databases. *Journal of Computational and Graphical Statistics*, 1996, **5**: 58-77.

[6] Ziegler H, Jenny M, Gruse T, et al. Visual market sector analysis for financial time series data. In: IEEE Symposium on Visual Analytics Science and Technology (VAST), 2010: 83-90.

[7] Keim D A, Nietzschmann T, Schelwies N, et al. A spectral visualization system for analyzing financial time series data. In: Proc. of Eurographics/IEEE-VGTC Symposium on Visualization (EUROVIS'06). Lisboa, Portugal, 2006: 195-200.

[8] Wattenberg M. Visualizing the stock market, CHI extended abstracts on human factors in computing systems. In: Proc. of CHI. Pittsburg, PA, USA, 1999: 188-189.

[9] Lin J, Vlachos M, Keogh E, et al. Iterative incremental clustering of time series. In: Advances in Database Technology–9th Intl. Conference on Extending Database Technology (EDBT). Heraklion, Crete, Greece, 2004.

[10] Rudolpn S, Savikhin A, Ebert D S. FinVis: Applied visual analytics for personal financial planning. In: IEEE Symposium on Visual Analytics Science and Technology. Atlantic City, NJ, USA, 2009: 195-202.

[11] Savikhin A, Maciejewski R, Ebert D S. Applied visual analytics for economic decision-making. In: IEEE Symposium on Visual Analytics Science and Technology (VAST), 2008: 107-114.

[12] Roberts J. State of the art: Coordinated & multiple views in exploratory visualization. In: Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization. Zurich, Switzerland, 2007: 61-71.

[13] Stasko J, Gorg C, Liu Z. Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization*, 2008, **7**(2): 118-132.

[14] Koch S, Bosch H, Giereth M, et al. Iterative integration of visual insights during patent search and analysis. In: IEEE Symposium on Visual Analytics Science and Technology, 2009: 203-210.

[15] Wang X, Dou W, Chen S, et al. An interactive visual analytics system for bridge management. *Computer Graphics Forum*, 2010, **29**: 1033-1042.

[16] Wang X, Miller E, Smarick K, et al. Investigative visual analysis of global terrorism. *Computer Graphics Forum*, 2008, **27**: 919-926.

[17] Chang R, Ghoniem M, Kosara R, et al. WireVis: Visualization of categorical, time-varying data from financial transactions. In: IEEE Symposium on Visual Analytics Science and Technology. Sacramento, CA, USA, 2007: 155-162.

[18] Stolte C, Tang D, Hanrahan P. Multiscale visualization using data cubes. *IEEE Transactions on Visualization and Computer Graphics*, 2003, **9**(2): 176-187.

[19] Butkiewicz T, Dou W, Wartell Z, et al. Multi-focused geospatial analysis using probes. *IEEE Transactions on Visualization and Computer Graphics*, 2008, **14**(6): 1165-1172.

[20] Brown J, Duguid P. Organizational learning and communities-of-practice: Toward a unied view of working learning, and innovation. *Organization Science*, 1991, **2**(1): 4057.