Using Visual Analytics to Enhance Data Exploration and Knowledge Discovery in Financial Systemic Risk Analysis: The Multivariate Density Estimator

Victoria L. Lemieux[1,2], Benjamin W.K. Shieh[2], David Lau[2], Sun Hwan Jun[2], Thomas Dang[2], Johnathan Chu[2] and Geran Tam[2]

[1] iSchool, The University of British Columbia

[2] Media and Graphics Interdisciplinary Centre, The University of British Columbia

Abstract

Analyzing and managing the risks in financial systems is necessary to maintain healthy global financial systems and economic wellbeing. However, the complexity of the financial system and the heterogeneity and volume of data sources needed for financial systemic risk analysis are currently overwhelming. Visual Analytics tools can be used to provide macroprudential supervisors with greater visibility into the health of financial systems by augmenting their information processing capabilities. To this end, we present a novel prototype design for a visual analytics tool that implements the multivariate density estimator of financial systemic risk, explaining how it addresses macroprudential supervisors' need for enhanced data exploration and knowledge discovery capabilities.

Keywords: financial systemic risk analysis, financial transparency, visual analytics, treemap

Citation: Lemieux, V. L., Shieh, B. W. K., Lau, D., Jun, S. H., Dang, T., Chu, J., & Tam, G. (2014). Using Visual Analytics to Enhance Data Exploration and Knowledge Discovery in Financial Systemic Risk Analysis: The Multivariate Density Estimator. In *iConference 2014 Proceedings* (p. 649–653). doi:10.9776/14307

Acknowledgements: This research was supported by The Boeing Corporation and the Natural Sciences and Engineering Research Council of Canada.

Contact: vlemieux@mail.ubc.ca

1 Introduction

The financial crisis of 2007-2009 led to many bank failures and a global credit crunch that drew attention to the need for enhanced financial systemic risk analysis capabilities (Lemieux, 2013). A Financial Stability Board and International Monetary Fund report on "The Financial Crisis and Information Gaps" noted that, "Indeed, the recent crisis has reaffirmed an old lesson—good data and good analysis are the lifeblood of effective surveillance and policy responses at both the national and international levels" (FSB/IMF, 2009). Given this, it is worthwhile exploring how new approaches to the analysis of large and complex data sources might be used to enhance financial systemic risk analysis capabilities. One approach is to apply Visual Analytics (VA), defined as the science of analytical reasoning facilitated by interactive visual interfaces (Thomas and Cook, 2005). VA is particularly suited to addressing information processing challenges as it combines machine intelligence with the visual and cognitive intelligence of human analysts through the use of interactive visual interfaces.

This paper presents a design study of a VA system to support financial systemic risk analysis. The presentation of the study follows Munzner's Nested Model for Visualization Design and Validation (Munzner, 2009). The model divides visualization design into four levels: 1) characterize the problems of the real-world users; 2) abstract into operations on data types; 3) design visual encoding and interaction techniques; and 4) create algorithms to execute techniques efficiently. As our design study makes a contribution in only the first three areas, we will focus the discussion on these areas of the Nested Model.

2 Domain Problem Characterization

A VA tool designer must learn about the tasks and data of users in the target domain (Munzner, 2009). To do this, we reviewed literature in the field of financial systemic risk analysis to understand how it is characterized. To complement our literature review, we held an interdisciplinary workshop where financial systemic risk experts met with visual analytics experts to characterize the domain problems (Lemieux, 2012a). The workshop was supplemented by an interdisciplinary panel discussion on the application of VA to financial systemic risk analysis and further discussions of the topic at meetings of experts in the field of financial systemic risk analysis (Lemieux, 2012b).

Having conducted research to characterize the domain problems and to learn about the tasks and data of users in the target domain, we identified several key domain problems. Domain experts placed emphasis on one of these in particular, which was the need to understand counterparty networks and interconnectedness of financial institutions. There are many different measures of financial interconnectedness; however, we chose to rely upon the time-varying multivariate, distress dependency (MDE) approach in the design of our tool (Segoviano and Goodhart, 2009). The MDE approach consists of a set of measures to analyse stability from three complementary perspectives by allowing: 1) the quantification of common distress in the banks of a system, 2) distress between specific banks, and 3) distress in the system associated with a specific bank i.e. a cascade effect. We acknowledge that the MDE approach is but one of many measures of financial interconnectedness, and a robust approach to financial systemic risk analysis will benefit from tools that support a variety of approaches.

3 Abstract into Operations on Data Types

Following Munzner's (2009) Nested Model, in the abstraction stage we mapped problems and data from the vocabulary of the domain of financial systemic risk into the more abstract and generic description that is in the vocabulary of information visualization and visual analytics. The output is 1) a description of operations (i.e., generic tasks) and 2) a description of data types, which are the input required for making visual encoding decisions at the next level.

To achieve a description of operations, we first extrapolated high-level domain specific tasks from the literature on financial network analysis. We then linked the higher analytic tasks to several relevant formative evaluation frameworks in the field of information visualization and visual analytics. For example, we used Amar and Stasko's (2004) knowledge task-based taxonomy aimed at addressing complex decision-making, especially under uncertainty. We also applied Yi *et al.* (2007), who have proposed a taxonomy of interaction techniques, which they define as "the features that provide users with the ability to directly or indirectly manipulate and interpret representations" (p.2). Finally, Lee *et al.* (2006) propose a task taxonomy specifically for graph exploration, which is suited for use in interactions with networks.

Raw data inputs for computation of the MDE measure are: (1) a list of banks denoted by their ticker tape symbols and (2) a table of Credit Default Swap (CDS) values for each bank, which we obtained from Bloomberg. CDS products are used because they act as signal to the market about the viability of the underlying financial institution (Credit Default Swaps, 2013). The raw data were held in two .csv files. Computation of the raw data, using the algorithms from Segoviano and Goodhart, produced derived data consisting of: (1) values for a Joint Probability of Distress (JPoD) and (2) values for a Distress Dependence Matrix (DDM) (Segoviano and Goodhart, 2009). The derived data were output to two .csv files. We supplemented this data with a table of market capitalization values (once again obtained from Bloomberg) per bank, also in .csv format, in order to be able to represent the size of each bank relative to the market as a whole.

4 Design Visual Encoding and Interaction Techniques

Once we had identified the basic operations and transformed the raw data into the derived data, it was possible to determine the type of visualization best suited to representing the data and likely to meet the task requirements. We had a number of options to choose from. Networks of financial relationships, like other types of networks, are typically abstracted as graphs. Graphs may be represented as matrices, as node-link graphs, or as trees. In much of the literature that we reviewed for this study, financial interconnections are represented as node-link graphs. Many of these graphs are over plotted and difficult to interpret.

We chose to visually encode our derived data using a Treemap (Johnson and Schneiderman, 1991). Technically, a tree is a connected, unweighted acyclic graph. There are two types of trees: space filling and non-space filling. A Treemap is of the space-filling variety and therefore uses screen real estate very efficiently compared to non-space-filling visualizations. For this reason, we judged that a Treemap would avoid the over-plotting and occlusions that make node-link graphs representing large networks difficult to read, though we acknowledge that Treemap also has limitations in the representation of large datasets. That is, banks that are small in size relative to the market as a whole may not be clearly visible in displays of large networks. Recalling our high-level domain tasks, we also saw the Treemap as an effective way to represent the topology of the network at different levels of granularity as it has been used to do in other financial application areas (Smartmoney™, 2013).

In our system, a quick glance at the Treemap visualization provides financial systemic risk analysts with an instant overview of the status of the financial system and the number of banks that may be in danger of default. The first Treemap visualization, the PoD view, shows the joint probability of distress among all banks in the system. The Treemap visualization maps the size of the rectangle to the market capitalization of each bank, so that an analyst can quickly infer the impact of a bank's failure on the market as a whole i.e. whether it might be "too big to fail". Colour is used to represent whether the bank's probability of default value is above (red) or below (blue) a user-defined threshold. The analyst is able to interact with the Treemap visualization by changing the default threshold value in a separate input box. This provides the analyst with the ability to conduct "what if" scenario analysis to instantly see the impact on the financial system's stability as the threshold value is changed. The Treemap visualization is reinforced by a separate Bullet graph visualization that shows the relative probability of distress values for each bank using colour (black) and bar length. The user defined default threshold is represented by a background display of colour (blue) and bar length. Users are able to change the default threshold in the Bullet graph by pointing and clicking on the blue slider bar to move the threshold value up or down. Linking techniques automatically update the Treemap visualization to reflect how changes in the threshold value affect stability of the financial system. Having understood the status of the financial system as a whole, an analyst might wish to understand the affect that a default of a particular bank would have on other banks in the system. The second visualization in our multiview system – the Distress Dependency Matrix (DDM) - supports this analytic task. By switching to the DDM view, the analyst is able to see the probability of default of all connected banks in the event that a particular bank defaults. In this view, the bank of interest is greyed out, while all other banks are either blue (safe) or red (at risk of default). The bullet graph view is retained in this view to indicate the distress level of each bank against the adjustable user-defined threshold to support "what if" scenario analysis.

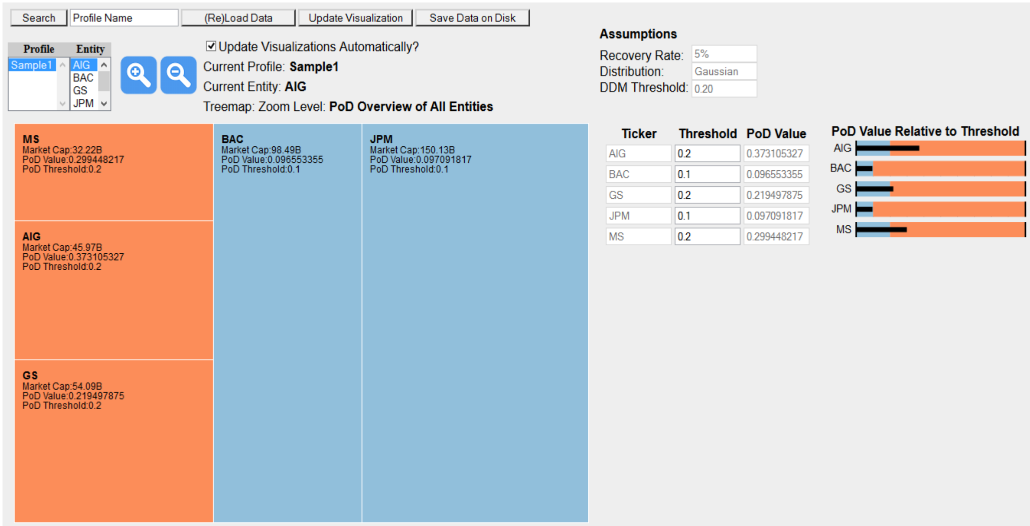

Figure 1: A screenshot of the MDEV tool showing the PoD view for five banks

5 Validation and Evaluation

In this section we discuss what has been done to date in regard to validation and evaluation of our system. Following Isenberg *et al.* (2008), we used a grounded evaluation approach to validate our characterization of the domain. We validated that visual analytics was a legitimate approach to solving the information processing problems in the domain of interest through conducting the workshop and panel discussions on the application of visual analytics for financial systemic risk analysis. After we developed our first prototype, we validated this assumption again by demonstrating the system to two groups of experts in the field. Both groups confirmed that the system was useful and would support the ability to more readily see and understand financial interconnections in relation to financial distress. We acknowledge; however, that ethnographic and cognitive work analysis would deepen our understanding of the domain problems and high-level analytic tasks.

At the visual encoding and interaction design level, we undertook a formative evaluation with one expert group at which we received feedback that the way we had chosen to present the multiple views in our system was confusing to users. Initially, we had users click on a particular bank in the PoD Treemap visualization to semantically zoom into the DDM visualization. Users found this confusing because they expected details about the particular bank, rather than about the effect of the bank on other banks. This feedback led to our changing how we presented the PoD and DDM views, and to reserve semantic zooming from one level of the Treemap visualization to another for use in providing greater detail on that bank's balance sheet structure. Users also expressed a desire to see a view of the interconnections in the more traditional node-link graph representation, which we will incorporate into a later release of our system.

6 Conclusion

Addressing the information processing challenges that contributed to the global financial crisis remains a significant and unresolved challenge. Visual analytics, which combines the strength of machine information processing with the best of human information processing through the use of interactive visual interfaces offers a promising approach to addressing these challenges. Our system aims to make a novel contribution to the application of visual analytics in the domain of financial systemic risk analysis. It is, however, still a work in progress.

7 References

Amar, R, and Stasko, J. (2004, October) *A knowledge task-based framework for design and evaluation of visualizations*. Paper presented at the 10th IEEE Symposium on InfoVis (INFOVIS '04), Austin, Texas.

Credit Default Swaps (2013). In *Investopedia.com*. Retrieved from, http://www.investopedia.com/terms/c/creditdefaultswap.asp.

Financial Stability Board & the International Monetary Fund (FSB/IMF) *The financial crisis and information gaps*. Retrieved from http://www.imf.org/external/np/g20/pdf/102909.pdf

Isenberg, P., Zuk, T., Collins, C., & Carpendale, S. (2008) *Grounded evaluation of information visualizations*. Paper presented at *AVI Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV)*, Florence, Italy.

Johnson B. and Schneiderman. B. (1991) *Tree-maps: A space-filling approach to the visualization of hierarchical information structures.* Paper presented at the IEEE Conference on Visualization '91, San Diego, California.

Lee, B., Plaisant, C., Parr, C.S., Fekete, J-D. & Henry, N. (2006) *Task Taxonomy for Graph Visualization.* Paper presented at the *AVI Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV)*, Venice, Italy.

Lemieux, V. Ed. (2013) *Financial Analysis and Risk Management: Data Governance, Analytics and Life Cycle Management*. Frankfurt: Springer.

Lemieux, Victoria L. (2012a, May) *Canadian and Global Challenges in Visual Analytics for Financial Systemic Risk Analysis*. Workshop, Banff Centre for the Arts, Banff, Alberta.

Lemieux, Victoria L. (2012b, October) *Panel Discussion on Visual Analytics for Financial Systemic Risk Analysis*. IEEE Visual Analytics Science and Technology, Seattle, WA.

Munzner, T. (2009) A Nested Model for Visualization Design and Evaluation. *IEEE Transactions on Visualization and Computer Graphics 15*(6), 921-928.

Segoviano, M.A. and Goodhart, C. (2009) *Banking stability measures.* (Discussion paper 627). London: Financial Markets Group, London School of Economics and Political Science.

SmartMoney™ Map of the Market. (2013) Retrieved from http://www.smartmoney.com/map-of-the-market/.

Thomas J.J. & Cook, K.A., Eds. (2005) *Illuminating the Path: The research and development agenda for visual analytics*. Washington, D.C: IEEE Computer Society Press.

Yi, J.S, Kang, Y.A, Stasko J.T, & Jacko, J.A. Toward a deeper understanding of the role of interaction in InfoVis. *IEEE Transactions on Visualization and Computer Graphics 13*(6) (2007), 224–231.

8 Table of Figues