

© [2009] IEEE. Reprinted, with permission, from [Mao Lin Huang, Jie Liang and Quang Vinh Nguyen, A Visualization Approach for Frauds Detection in Financial Market, 2009, 13th International Conference Information Visualisation, 2009]. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Technology, Sydney's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it

A Visualization Approach for Frauds Detection in Financial Market

Mao Lin Huang¹, Jie Liang¹, Quang Vinh Nguyen²

¹ Faculty of Engineering & Information Technology, University of Technology, Sydney, Sydney, Australia

² School of Computing and Mathematics, University of Western, Sydney, Australia
{maolin@it.uts.edu.au, jiliang@it.uts.edu.au, vinh@scm.uws.edu.au}

Abstract

The traditional solutions to the stock market security are not sufficient in identifying attackers and further attack plans from the analysis of existing events. Therefore, it is difficult for analysts to prevent future unexpected events or frauds by only monitoring the real-time trading information. The event-driven fraud detection in financial market could not help analysts to find attack plans and the further intention of attackers. This paper proposed a new framework of visual analytics for stock market security. The proposed solution consists of two stages: 1) Visual Surveillance of Market Performance, and 2) Behavior-Driven Visual Analysis of Trading Networks. In the first stage, we use a 3D treemap to monitor the real-time stock market performance and to identify a particular stock that produced an unusual trading pattern. We then move to the next stage: social network visualization to conduct behavior-driven visual analysis of suspected pattern. Through the visual analysis of social (or trading) network, analysts may finally identify the attackers (the sources of the fraud), and further attack plans.

Keywords--- Information visualization, visual analytics, finance security, stock market surveillance.

1. Introduction

To ensure the security of financial markets, the surveillance of stock trading becomes more and more important, as it saves customers a large amount of money in trading activities every day around the world. Up to date, there have been many researches working on applying computational intelligence to financial modeling and market security [3, 4], including some visualization technologies [5, 6, 7, 8, 9, 10, 11, 12, 13]. To prevent market imperfection, various applications have been introduced into market practices. Failure to employ modern technology and methods used in conducting surveillance can mean significant financial loss to market abuse or loss of market goodwill as investors choose to trade their assets elsewhere.

Traditional financial fraud detection relies on many AI techniques either tries to detect known frauds or indicates potential new frauds by identifying “abnormal”

behavior [4]. The main problem is that these practical detection systems often produce far too many false alarms which are costly if acted upon or ignorant from many undetected frauds which even cause more costly damages. While it may be theoretically possible to remove all security vulnerabilities through formal methods and better engineering practices, it is practically infeasible. Thus, financial frauds detection will remain a crucial security practice for years to come.

Although some existing fraud detection systems help by automatically identify potential frauds (or attacks) from financial trading activity and produce alerts describing the details of these intrusions. They’re often just a starting and useful first step in uncovering security compromises. The complex task of analyzing and determining the accuracy and severity of security events uncovered during monitoring is still with the human analyst.

Current support for frauds detection often focuses on increasing automation. While automation allows for increased scalability, humans provide the ability to handle exceptional events. There has been little research into understanding how fraud detection systems can be designed to assist the cognitive process of the human analyst. Information visualization shows great potential in supporting the monitoring task of frauds detection.

However, studying the available visualization-based systems, we found that most current solutions are more focusing on real-time and interactive visualization of large volume of data with data mining algorithms. They either lack techniques for visualizing appropriate abstractions of a large volume of financial data, or can only deal with low level raw-data.

They haven’t considered the analysis of related social networks, which could be based on a behavior driven visual pattern analysis, as part of the frauds detection process. However, it is very common that a suspected attacker, who intend to make a large number of further frauds or attacks in his plan and these attack plans are hard to be discovered by using the traditional approaches.

This paper proposes a new visualization approach that is specifically designed to solve the fraud detection problems in financial markets. In the new visualization framework, we use: 1) a 3D Treemap to perform visual surveillance of stock market by extracting and displaying the most recognizable analytical features of a large volume

of dynamic financial data that allows users to see the overall view of market operation and identify an abnormal performance of a stock in a single visual workspace, 2) a smart behavior-driven visualization system will display and analysis the social networks, including a stock trading network, of the suspected stock. The system will identify a fraud by conducting an orthogonal analysis (or pattern recognition); matching an unusual pattern to similar others in the pattern database.

We will first discuss the traditional framework and existing visualization solutions to the stock market security problem and then propose a new framework with extra stages for detecting frauds and attacking plans in the stock market.

The effective detection and subsequent analysis of the types of fraudulent activity which occur within the stock exchange markets are varied and it changes with the appearance of new technologies and new forms of the financial activity. The dynamic nature of fraudulent activity as well as the changing nature of normal usage of a service has rendered the detection of fraudulent intent from observed behavioral patterns a research problem of some considerable difficulty.

It is recognized that a common probabilistic theoretical framework be employed in the development of dynamic behavioral models which combine a number of prototypical behavioral aspects to define a model of acceptable behavior (e.g. patterns of mobile phone communication, web browsing, stock trading) from which deductions of the probability of unusual behavior can be made. In addition to these inferential models a means of visualizing the observed behavior and the intentions behind it will significantly aid the pattern recognition abilities of human fraud analysts in the financial market. Employing the common probabilistic modeling framework which defines the “fraud detection” models visualization tools will be developed to provide meaningful visual representations of dynamic activity which has been observed and visualizations of the evolution of the underlying states generating the observed activity. The development of detection & analysis tools from the common theoretical framework will provide enhanced detection and analysis capability in the identification of fraud.

2. Traditional solution

Current surveillance applications monitor the real-time raw trading flow (a large amount of information) and highlight any unusual price or volume movements. They traditionally compare the electronic signal containing all details of trading against a series of parameters. When these parameters are exceeded, the system triggers an “alert”. Each alert is referred to an analyst, who assesses the market conditions that caused it. For example, the analyst might determine that public information, such as a report on the stock written by a stockbroker, a company announcement, a newspaper report or broader market conditions can explain the market activity. Surveillance analysts also look for “patterns of alerts”, which may

indicate that an attempt is being made to interfere with normal market forces, requiring further investigation.

To find similar patterns, as a prerequisite, the behavior of financial market has to be described by a numerical form, as traditional applications reply on numeric figures. Numerous researches are proposed to model the dynamics beneath a single financial series, and they mainly focus on the time dependency of input series.

2.1 Traditional graphical standards

A commonly used graphical representation in stock market is the two-dimensional line plot in which every financial portal or newspaper displays, for example in yahoo money, see Figure 1. The basic idea is to show time on the *X-axis* and price or volume on the *Y-axis*. Price is often represented by using a line series and volume by using a separated bar chart. This standard chart representation is very appropriate and accepted in technical analysis of market evolution in low-dimensional time series. However, based on the nature of financial data, it would become insufficient and inefficient for multi-dimensional analysis. [1]

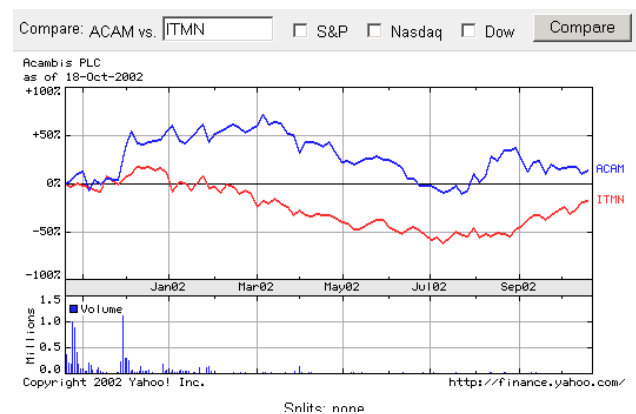


Figure.1: A visualization used in Yahoo web finance that compares two stocks by price and volume.

2.2 Current work in financial data

Later when stock market behavior descriptors that could extract and characterize the dependencies among stock market quantities from data distribution are introduced, visualization techniques come into the solution. The literature showed that a number of researchers attempted to solve problems with visualization. Several efforts in today’s financial graphics world are briefly discussed below. Some of the innovative solutions coming out of academic research have been commercialized into the software in the market, For example, Technology Road Map (TRM) surveillance system that is used by London stock exchanges and some others are focused on three-dimensional representations.

2.2.1 Financial Viewpoints

An interesting financial analysis tool is created in an information environment named “Financial Viewpoints” [2] by Lisa Strausfeld of the MIT Media Lab. The visualization employs three movable planar spreadsheets, one on each axis to represent 7 mutual funds, with 50 data elements each. Bar charts and video feeds have also been integrated to support interpreting the data. However, this architecture can not overcome the limits of standard line plotting, as it doesn’t have ability to illustrate the relationships beyond three linear dimensions.

2.2.2 Self-Organizing Map

A new visualization tool: Self-Organizing Map [5] for assisting technical analysis was proposed by Kresimir Siyunic from the institute for Media Communication, Fraunhofer Gesellschaft in Germany. An index of assets is to present a single price path or shape for a given period and instead of visualizing charts explicitly. This work uses the principal distribution of chart shapes and clusters similar chart shapes of component asset. It also shows the distribution of patterns within the aggregate and maps them into 2D display. This 2D display method enables the several interaction mechanisms which allows analysts navigate from a large index overview to sectors. However, visualization relationship between the chart shape and the future course of the chart is not realized in this work yet.

Based on various existing applications available in the market, we can summarize the traditional solution to the problem of financial markets security: to detect unusual patterns, and frauds, the traditional market surveillance relies largely on real-time visualization and interactive visualization of large volume of transactions, and in cooperation with the data mining algorithms. However, the traditional solution might only be used to identify attackers from a significant event once and it would be still difficult for analysts to prevent future unexpected events, as monitoring real-time trading information only could not help analysts to find the attack plan and the intentions of attackers. The framework of traditional approach is illustrated in Figure 2.

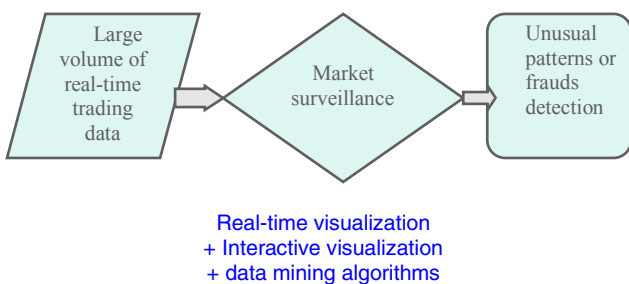


Figure 2: The Pipeline of Traditional Solutions.

3. A New Framework for Stock Market Security

For the complete solution, conducting behaviour analysis of suspected traders should be added into the pipeline. Based on behavioural analysis of social networks

visualization, including trading history network, email network, telephone network, etc. analysts can explore suspected the abnormal behaviours and finally identify attacker (the source of the fraud), and attack plan. To make sure the attacker and plan, analysts should further explore the patterns of information based on a similarity-based clustered visualization, then may identify a group of trading accounts and finally the action can be taken to exclude attacker and his or her accounts from the stock market, see Figure 3. Therefore, the new solution consists of two stages: 1) Visual Surveillance of Market Performance, and 2) Behavior-Driven Visual Analysis of Trading Networks.

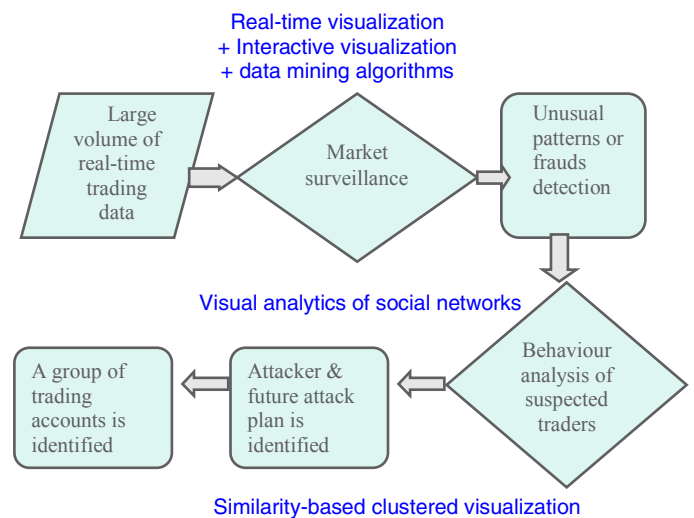


Figure 3: A proposed framework for stock market security.

3.1 Stage 1. Visual Surveillance of Market Performance

In stage 1, visual surveillance should have the following functions, including 1) real-time raw data collection, 2) feature extraction and abstraction, 3) overall view of the performance and 4) detailed view of the individual’s performance. An individual could be a particular business category, or a group of similar stocks, or a particular stock.

On extraction of features from raw data, the visual analytics process can be conducted effectively and efficiently; exploring new learning rules to discover patterns and outline, defining essential features and attributes, those to be important in identifying unusual patterns that may be related to a fraud, from the large volume of raw financial datasets; and discovering how a specific event is related to other information available in the massive dataset. We believe that more can be characterized and abstracted from the stock trading data to assist the visualization process to perceive suspicious frauds or identify actual security breaks. However, current work mostly relies on simple relationships between elements of raw data and hence we believe more can be revealed from better abstractions.

The proposed on-line visualization system can identify a particular stock that produces an unusual

performance through a 3D visual surveillance of recognizable features extracted from the real-time financial data and then passes it to the next stage “trading network visualization” for further analysis on the social behaviour of this suspected stock.

We use a Three-Dimensional EncCon Tree (treemaps) [14] to perform the stage one that monitor the market performance. It provides an overall view of the real time stock market performance, see Figure 4.

The proposed 3D treemaps was upgraded from a 2D visualization called “SmartMoney” [15], a well done implementation of 2D treemaps shows 500+ popularly held stocks, organized by industry groups, size-coded by market capitalization, and color-coded to show rise or fall.

SmartMoney is also known as a “cluster treemap” that uses the squarified treemap algorithm to generate square-like rectangles. It avoids the thin rectangles in the slice-and-dice algorithm. The SmartMoney website also offers a for-fee premium service, SmartMoney Select, with advanced treemap features such as dynamic query filters and other data search services.

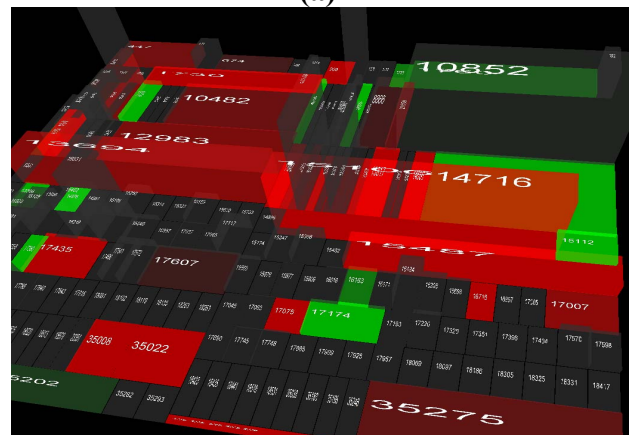
To uncover an additional dimension of complex financial data, we upgraded the SmartMoney to 3D by combining a commonly used real-time price chat with a 2D treemaps for visual surveillance in a 3 dimensional visual space. The proposed visualization shows not only the performance of real-time stock market through the visual properties, e.g. size of cubes, colors and labels, but also the current price of stocks indicated by the height of cubes and changes of prices according to the time. The view of an interested individual company can be converted into plot charts view as an overview of price change.

We utilized the “treemaps” view developed by Ben Shneiderman and first time being used in stock market by Jungmeister and Turo in 1992 [6]. The map of the market also has the capability to show the changing of stock prices the relative returns and the capitalization of companies in one compact 3D view. The treemap is partitioned at the top of its hierarchy by industry sectors, to categorize a large number of companies. The visualization updates every 15 minutes. Colored rectangular tiles are used to represent stocks and ordered tree-maps are used to lay out the structure.

The proposed 3D visualization uses visual cues, such as size, colors, positions, etc, to represent the properties of major stocks concurrently. Two major properties of the stock are displayed by default: they are market capitalization (indicated by the size of a cell) and the current price of stocks (indicated by cell attributes). The color code of red indicates the decrease of the price. The dark red indicates the low decrease of price, and the light red indicates the high decrease of price. Oppositely, green means the price is increasing. Similarly, the darker green means the low decrease. The brighter color is the higher increase is. Black indicates no change.



(a)



(b)

Figure 4: Two displays of a 3D EncCon Tree that combines a real-time Price Chat and 2D Treemap for Visual Surveillance of Stock Market.

3.2 Stage 2: Behaviour-driven Visual Analysis of Trading Networks

The second stage is to conduct behavior-driven visual analysis of trading networks. We use node-link diagrams to reveal social relational structures among traders. We aim to identify “suspected trading patterns” in the trading network by matching the similar patterns in the historical pattern database with the suspected one in the historical trading times slots, and finally to identify fraud patterns, suspected traders (attackers), and attack plans.

In the visualization of stock trading network, a node (circle) represents a particular trader and an edge (link) represents the trading relationship. The area of a circle (node) represents the trading value (buy & sell) or the capitalization of a stock share. The width of an edge represents the weight or exchange amount. The arrows of an edge indicate the exchange directions from one trader to another, see Figure 5.

This visualization dynamically tracks the trading patterns of a particular trader and highlights any unusual pattern occurred at a time. For example, in Figure 6, a suspected behavior of domination of the stock price has been identified.

Once a suspected trading pattern is found, an

orthogonal analysis is followed. By orthogonal, the analysis is done independently and expanded to: analyzing in the time domain (temporal analysis); analyzing in the spatial domain (spatial analysis); analyzing in the frequency domain (spectral analysis); tracking where the sources of frauds (attack) are in the analytical process and how they got there; gathering other sources of information without leaving the visualization console, and incorporating the results into the view. Most current frauds detection systems analyze financial data in either the time domain (time series analysis), or the spatial domain (space analysis) to correlate coordinated frauds, and hardly any explores the analysis in the frequency domain. However, many interesting events can hardly be seen in the time domain but in the frequency domain, these events can be readily revealed. Furthermore, spatial correlation (such as space diversity) in antenna design can provide more reliable identification of the desired signal by “receiving” signals from different locations. For these reasons, we propose the “orthogonal analysis” in the second stage.

These analyses are performed only after the visualization stage when the human analyst is able to point out areas (circles) of interest. These may be some unusual financial events or suspected trading activities across the visualization screen and they need confirmation. Data associated with these unexpected visual patterns will be isolated and analyzed to find temporal correlation, spatial correlation, and dominant frequencies that may link seemingly isolated events.

In comparison with the existing solutions that emphasize either on event-driven analysis or a behavior-driven analysis, the proposed solution seems to be a more completed solution that combines the visual surveillance of real-time performance and the behavior analysis stages that could effectively assist users to detect not only the isolated frauds and attacks, but also the intensions or future plans of suspected frauds (or attacks), see Figure 7.

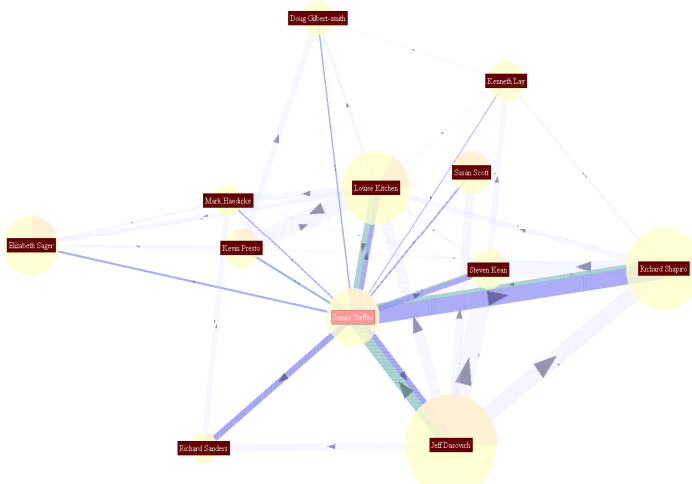


Figure 5: A display of trading pattern in which a circle (node) represents a particular trader and an edge represents the trading relationship.

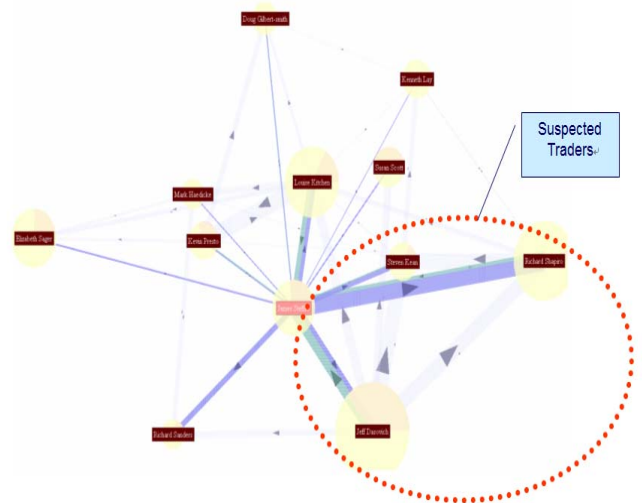


Figure 6: A suspected behavior of domination of the stock price has been identified.

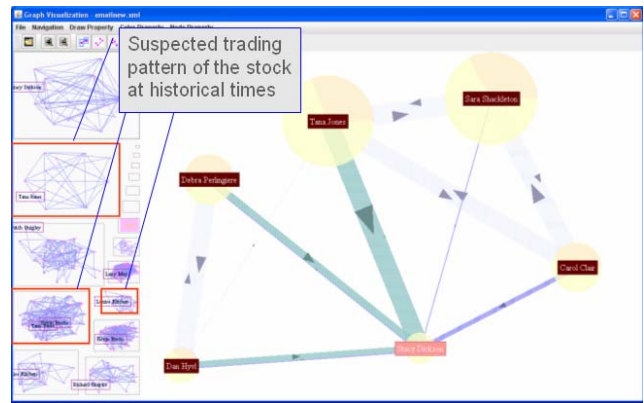


Figure 7: Identify frauds by matching suspected patterns with similar patterns in the historical times.

4. Conclusion and recommendation

A fair trading system is crucial to an exchange market and plays a critical role in determining the overall efficiency of the market. A smart surveillance system is crucial to guarantee such a fair system by detecting market imperfections and human frauds to prevent individual investors' fair trading that will potentially increase investors' confidence in the marketplace.

However, a good market surveillance platform has significantly higher performance requirements than other financial IT systems, both in sheer volume of transactions, as well as the extremely high transaction rates. Integrating with such high performance systems is an area which can involve significant technology advances.

For technology, we believe that most visualization techniques should have higher ability for tracking changing data and real time analysis. However, a single visualization may only be able to achieve one phase of a complete analytical target, especially in the area of complex financial data analysis process. Therefore, a

higher degree of the integration of several visualizations, which take advantages of human vision, as well as the artificial intelligent algorithms, that take advantages of logical reasoning together as a complete visual analytic procedure to assist human for identifying attackers and their further plans of attack is essentially needed. Such an advanced integration should be considered as an optimized method for eliminate the source of a variety of attacks and frauds in the financial market.

In this paper, we proposed a new framework of visual analytics for stock market security. Our integration consists of two visualizations: 1) Visual Surveillance of Market Performance, and 2) Behavior-Driven Visual Analysis of Trading Networks. In the first stage, we use a 3D treemap to monitor the real-time stock market performance and to identify a particular stock that produced an unusual trading pattern. We then move to the next stage: behavior analysis of suspected traders. Through the visual pattern analysis, we may finally identify attackers (the sources of the fraud), and their attack plans.

Reference

- [1] Tue D., H., Bazinet, A., Berthier, R., and Shneiderman, B. 2008. "NASDAQ Velocity and Forces: An Interactive Visualization of Activity and Change". *Journal of Universal Computer Science*, 14:1391-1410.
- [2] Strausfeld, L. 1995, "Embodying Virtual Space to Enhance the Understanding of Information", Master's Theses, Massachusetts Institute of Technology, Program in Media Art and Sciences.
- [3] Fan, M, Stallaert, J. & Whinston, A.B. 2000, "The Internet and the Future of Financial Markets", *Communications of the ACM*, vol. 43, no.11, pp.83-88.
- [4] Lai, P.S. & Fu, H.C. 2007, "A Polygon Description Based Similarity Measurement of Stock Market Behaviour", *Proc. IEEE Congress on Evolutionary Computation*, pp806-812.
- [5] Simunic, Kresimir, 2003, "Visualization of Stock Market Charts", *Poster Proc. of 11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*.
- [6] Jungmeister, G.A. and Turo, D. 1992, "Adapting Treemaps to Stock Portfolio Visualization", *Technical Report UMPCSD CS-TR-2996*, University of Maryland, College Park, Maryland 20742, USA.
- [7] Merino, C.S., Sips, M., Keim, D. A., Panse, C. and Spence R., 2006, "Task-at-Hand Interface for Change Detection in Stock Market Data", *Proceedings of the working Conference on Advanced Visual Interfaces*, pp. 420 – 427.
- [8] Nesbitt, K.V. & Barrass, S. 2004, "Find Trading Patterns in Stock Market Data", *IEEE Computer Graphics and Applications*, Vol.24, no.5, pp. 45-55.
- [9] Parrish, Edward, 2000, "StockVis: An Internet-Based System for Visualizing Stock Market Data", *Master's Thesis*, UC Santa Cruz, Department of Computer Science.
- [10] Roberts, P. 2004, "Information Visualization for Stock Market Ticks: Toward a New Trading Interface", *Masters Thesis*, MIT.
- [11] Shreck,T., Tekusova,T. and Kohlhammer, J., 2008, "Trajectory-Based Visual Analysis of large Financial Time Series Data", *SIGKDD Explorations*, Vol. 9, no. 2, pp30-37.
- [12] Teoh, S.T., Ma, K.L., Wu, S. F. and Jankun-kelly, T.J, 2004, "Detecting Flaws and Intruders with Visual Data Analysis", *IEEE Computer Graphics and Applications*, Sept.-Oct. 2004, pp. 27 – 35.
- [13] Ziegler, H., Nietzsche, T. and Keim, D.A. 2007, "Visual Exploration and Discovery of Atypical Behavior in Financial Time Series Data using Two-Dimensional Colormaps", *Proc. 11th International Conference on Information Visualization*, pp. 308 – 315.
- [14] Huang, M.L. and Nguyen, Q.V. 2007, "Three-Dimensional EncCon Tree", In *Proc. of IEEE Int'l Conference on Computer Graphics, Imaging and Vision (CGIV07)*, pages 429 – 433, Bangkok Thailand.
- [15] SmartMoney Software: <http://www.smartmoney.com/>
<http://www.smartmoney.com/map-of-the-market/>