UNIVERSITY OF MORATUWA

Faculty of Engineering

Stock Market Visual Pattern Recognition using Complex Event Processor

Initial Project Proposal

Module CS4202: Research and Development Project

Department of Computer Science and Engineering

Project Members:

100039D Arachchi R.W.M.
100056C Bandara K.A.M.N.
100442G Ranasinghe Y.R.M.
100517R Somatilaka R.D.C.G.

Supervisors:

Dr. Srinath Perera
Director - Research
WSO2

Dr. Daya Chinthana Wimalasuriya
Department of Computer Science and Engineering,
University of Moratuwa.

CONTENTS

INTRODUCTION

Today, in the era of big data, the velocity and volume of data inflow is very large and complex. Many events happening around are captured and produced as event streams with a large business value hidden in them. Though it is so, the technology and methodologies used to convert these data into useful information are comparatively low. To fill this gap in the technology the field of Complex Event Processing is emerging over last two decades.

"Event processing is a method of tracking and analyzing (processing) streams of information (data) about things that happen (events), and deriving a conclusion from them. Complex event processing, or CEP, is event processing that combines data from multiple sources to infer events or patterns that suggest more complicated circumstances. The goal of complex event processing is to identify meaningful events (such as opportunities or threats) and respond to them as quickly as possible."

~Wikipedia~

This concept of CEP has being implemented and used in the industry. There are open source as well as commercial CEP engines available. But currently they are used only in very limited number of business domains and applications. The field of CEP is yet to be extended at large.

This project is an attempt to use CEP in the domain of stock market, to analyze and identify stock market visual patterns. The input event streams for this application will be the transactions happen in the stock market. The charts are drown on last-traded-price versus time. These price values change rapidly, within nanoseconds or micro seconds. The professional specialists in the stock trading industry have identified that there are some chart patterns which play a critical role in technical analysis of stock market. These patterns can occur within any time frame: monthly, weekly, daily or an intra-day and they repeat themselves over and over again. Sooner one identifies such a pattern in a stock market he has a huge advantage in the stock market over others.

Demonstrated below is an example of such a pattern and there are more than 13 such patterns identified. The left hand picture is a model of the 'triple bottom' pattern and the right hand picture shows how it appears in real data, with lots of noise, fluctuations and outliers.

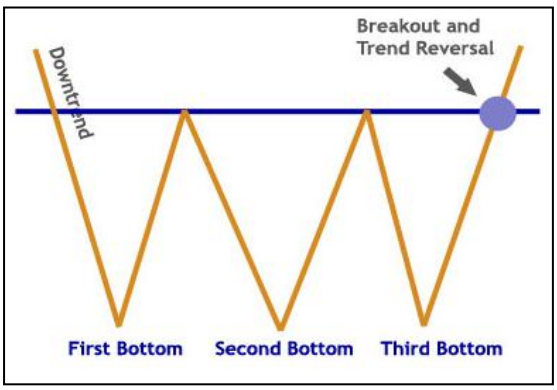

Figure 1:
Source: http://optionalpha.com/13-stock-chart-patterns-that-you-cant-afford-to-forget-10585.html

Figure 2:
Source: http://thestockbandit.com/triple-bottom/

The aim of this research is to identify these patterns using the first class functions provided by a complex processor, directly processing the stock market data stream. Here we are trying to come up with a set of algorithms, which will identify several important stock market patterns in streams of stock market data and implement those algorithms on an available open source CEP platform.

As the CEP platform for implementation we selected the WSO2 CEP. The reasons for selecting the WSO2 CEP are that it is 100% open source middle ware which promises high performance, massive availability and provides a wide array of CEP functionalists. Also our project supervisor (external) has a sound background knowledge and experience with WSO2 CEP.

IMPORTANCE OF THE PROJECT

Stock market is the place where the wealth is created and distributed among investors; the trading which takes place at the stock market affects the large enterprises as well as the economy of a country as a whole. Even a difference of a nanosecond in a transaction can save or loose thousands of millions of profit. Therefore identifying the correct time to buy or sell stock is very important.

Researches which have done on stock market data analysis, have pointed out some time series patterns on stock marketing trading price changes. Simply put, chart patterns are just a series of price action that occurs in a stocks trading. These can happen on any time frame really; monthly, weekly, daily and intra-day. The great thing about chart patterns is that they repeat themselves over and over again. They are a result of human nature and trading psychology and play a critical role in technical analysis of stock market.

Just as volume, support and resistance levels, RSI, and Fibonacci Retracements can help your technical analysis trading, chart patterns can help identify trend reversals and continuations and breakouts. So if you can learn to recognize these patterns early, you will gain a real competitive advantage in the markets. That is why the academic researchers and investors pay close attention on how to identify these patterns.

The technological platform we use is the complex event processor which is highly efficient and salable. The newly emerging field of complex event processing comes with a promise of efficiency over existence systems. The advantages of processing data/events in real time is there.

So applying the CEP technology in Stock Market technical analysis and coming up with algorithms which identifies chart pattern real time and alert the users consuming less resource and spending less time will have a high business value.

Not only that but also the approach which used to implement the algorithm would be helpful for applications of other domain which requires real time data processing and analysis.

We are developing this on open source platform and the output will also be open source so that it will serve a wider user community as well.

OBJECTIVES OF THE PROJECT

The main objective of this project is to design some algorithms to identify some of the already identified visual patterns in stock market and implement them using a CEP in real time. For that we should identify similar researches currently happening in the field as well.

SCOPE

The expected outcome will be set of algorithms implemented using CEP query language (Siddhi SQL) and some java back end extensions.

First there will be a research on existing solutions and approaches/technologies relevant to the topic. We have to identify strengths and support available in a CEP for a task like this.

We have to identify or come up with a proper noise filtering, outlier identifying and data smoothing methodology as well.

Then identification of each stock pattern has to be done. One pattern will be represented as set of states and a transition between those states. Then algorithms should be developed to identify the patterns.

At the implementation phase these algorithms will be implemented using query language and some extended Java functions.

As the final outcome, the CEP should be able to identify the implemented patterns if they are available in a provided input stream of real data.

We can summarize the scope of our project as follows

- Input to the system – Stock market data of various frequencies
- Approach – Filter the data stream, smooth and identify visual various patterns emerge in data
- Domain – Stock Market Analysis
- Output - A module to identify visual patterns of stock market data, in a real time event stream.
- Back end – Java based WSO2 complex event processor

DELIVERABLE

The following will be the major deliverable of our project:

- Set of algorithms to be used to identify visual patterns resides in a stream of events.
- Module for Visual Pattern Recognition in Stock Market Data , implemented on a CEP
- Comprehensive literature review on the work related to the topic
- Project report
- Project Website
- Research paper

ASSUMPTIONS/LIMITATIONS

- The concept and definition of complex event processing limits the number of processing and computational operations we can do over the real time data.
- The strengths of the filtering mechanisms available are limited.
- Unavailability of most of the research details of the field due to the high market value and competition is another difficulty faced. Most of the products and researches done in technical analysis of stock market are vendor specific and not revealed to the community.

PROJECT PLAN - DETAILED PROJECT PLAN FOR THE FIRST PHASE OF THE PROJECT

1. Research Component
   1.1. Gather research papers related to the topic, explore possible techniques that can be utilized.
   1.2. Analyze the approaches/techniques used in existing researches
   1.3. Getting familiarize with the WSO2 CEP and Siddhi engine. Identify strengths and weaknesses.
   1.4. Based on the above findings select a suitable approach, and do further research.
2. Development of Algorithms
   2.1. Identify state machines for the basic patterns
   2.2. Develop the algorithms
   2.3. Test the algorithms and calculate precision of them
3. Implement the proof-of-concept module
   3.1. Implement the algorithms
   3.2. Test the algorithms
   3.3. Calculate precision
   3.4. Compare and analyze the outcome
   3.5. Documentation