Trade Surveillance with Big Data

The rise of real-time, high-frequency trading has regulatory compliance teams working hard to keep up with widening pools of structured and unstructured data. Today, new tools and techniques can improve trade surveillance and help spot abuse and irregularities before they do harm.

Executive Summary

Financial markets are undergoing a dramatic transformation. Starting a decade ago, real-time computing (compliments of the Internet) emerged – enabling what is now commonplace: high-frequency trading (HFT).

Today, data is a trading firm’s most valuable commodity. With the processing capabilities provided by the latest generation of computers, the more data you have, the more opportunities you can spot. But managing exceptionally large volumes of data has its own share of problems, which tend to surface when it comes to trade surveillance.

Regulatory compliance teams, for example, are challenged by the rise of algorithmic trading, where split-second execution decisions are made by computers, as well as the explosion of trading venues and the exponential growth of structured and unstructured data to be scanned. Existing surveillance techniques are effective for tracking “click trading.” Beyond that, capital markets firms require a radically new and more modern approach.

This white paper highlights key issues faced by regulators and compliance teams, and examines new big data solutions that can help to manage them.

New Challenges Facing Trading Firms and Financial Regulators

In a world where high-frequency trading accounts for more than 50% of trades measured by volume, the amount of data an organization receives and generates has grown exponentially. As a result, making sense of all of the data that now flows into, through and out of firms has become a monumental task – made even more complicated when one considers that capital market firms need to manage the data deluge in real time.

Now let’s look at some of the obstacles facing the surveillance community today:

• Capturing and recalling each event in the lifecycle of a trade. Trading firms need to maintain and recall end-to-end lifecycles of individual trades. This helps internal compliance teams perform deep-dive analyses and dig into any issues that may arise. It also enables the firm to respond accurately to regulatory investigations when the need arises. In some instances, this becomes mandatory, as detailed in the following examples:
  - Per Dodd-Frank regulations, a trading firm must be able to provide step-by-step details of the complete lifecycle of any swap or related transactions – including information on related deals – in five days.
Financial Industry Regulatory Authority (FINRA) Rule 5270, effective September 3, 2013, requires that no FINRA member broker-dealer shall execute an order to buy or sell a security or a “related financial instrument” when that member has material, non-public market information concerning an imminent block transaction in that security, and that information has not been made public or has not become stale or obsolete. In the case of an inquiry, firms must quickly recall and reconstruct the entire lifecycle of the block trade. This can become very complicated; a single block order sent via an algorithmic engine can be broken into thousands of smaller orders, and may get routed to multiple execution venues over the course of hours, days or even weeks.

Moreover, complete and rapid recalls of the details of a trade require:

- Trade-related data from trade systems and ancillary systems.
- Electronic and telephonic communications data from instant messaging applications, e-mails, phone call logs and transcripts.
- Data from microblogging sites such as Twitter.

Firms must also be able to merge structured and unstructured data. This way, they can recreate every detail of the trade, and also gain insight into market conditions and any other information that may have influenced the trader.

- Curbing market manipulation in HFT. It is the responsibility of the market surveillance team to guarantee that market participants receive a fair play. Quote stuffing, spoofing/layering and momentum ignition are among the market-manipulation techniques that regulators on both sides of the Atlantic seek to detect and eliminate.

In September, 2013, the Commodity Futures Trading Commission (CFTC) announced plans to build a regulatory framework around high-speed and algorithmic futures trading. The CFTC subsequently released a 137-page document that requested public input on various proposed ways to control the associated technology risks while enabling more trades to be made faster and with less human interaction.

Detecting such market manipulation techniques requires real-time surveillance. But given the number of trades executed by the human traders and their robotic partners at venues spread across continents, connecting the dots in real time can be daunting; a trader working at multiple venues would deploy a high-frequency algorithmic system to process trades at exceptional velocities – making those trades impossible to track without equally fast technology.

- Assembling the bigger picture for efficient and effective “cross-market and cross-asset surveillance.” At a large firm, traders have access to multiple trading venues; they can view liquidity across all markets. They have a consolidated order book and can access its full depth across different asset classes and venues, including exchanges, ECNs and dark pools of liquidity.

Consider the tools that compliance officers have at their disposal for detecting market abuse. These tools can scan traders’ executions, but do not take into consideration the entire order book that was visible to the trader, i.e., trades made across multiple venues worldwide. They do not track a stock’s price movement against the volume traded for the stock across all venues on which the firm can trade. Nor do they compare the trader’s actions with the positions in related future contracts. Furthermore, compliance teams cannot compare the bid placed on an ECN with the best bid/ask. And they cannot detect spoofing if the trader is trading on several venues in different time zones with different currencies.

Finally, they do not provide a complete picture of the market. Compliance officers cannot perform any cross-market analysis with the information accessible to them. They need to have the same view as the trader – a consolidated order book across multiple venues.

Considering the large volumes of data that result from multiple venues worldwide and the accelerated speed of trading, it is simply not possible for a human being to detect all possible instances of market abuse. Consequently, future surveillance tools must be programmed to detect any suspicious activities, store very large volumes of data and analyze that data in real time.
A Regulatory and Industry Roadmap for Overcoming the Hurdles

The explosive growth of data over the last few years is taxing the IT infrastructure of many capital markets firms. Fortunately, there are emerging technologies that can help these companies better manage and leverage ever-bigger data pools. These tools can help trading firms end data triage and retain useful historical information. By building a big-data architecture, IT organizations can keep both structured and unstructured data in the same repository, and process substantial bits and bytes within acceptable timeframes. This can help them uncover previously inaccessible “pearls” in today’s expanding ocean of data.

Big data analytics involves collecting, classifying and analyzing huge volumes of data to derive useful information, which becomes the platform for making logical business decisions.

For today’s capital markets firms, big data sets can reach multiple petabytes (one petabyte is a quadrillion bits of data). Not surprisingly, relational database techniques have proven to be inadequate for processing large quantities of data, and hence cannot be applied to big data sets.

Big data analytics involves collecting, classifying and analyzing huge volumes of data to derive useful information, which becomes the platform for making logical business decisions.

Figure 1 depicts a representative big-data architecture appropriate for modern day trade surveillance.

**Figure 1**

**A Big Data Analytics Reference Architecture**
There are some key guidelines that firms should bear in mind while formulating big-data strategies for surveillance and compliance.

**New types of data sources should be included to gain a complete picture.** Real-time news and data from social media can be helpful in evaluating the circumstances under which a trade was executed.

- **Fuse data across the timeline.** Historical, new and real-time data must come together to provide a 360-degree view of market activity, sentiments and trading behavior.
- **Put activity in context.** Once validated against historical data, a suspicious activity can turn out to be an aberration or a pattern. The NoSQL database aims to supersede and replace legacy RDBMS systems to deal with the rising data tide.
- **Real-time analytics is a must-have.** Market data must be analyzed at the speed of automated trading. CEP technologies can provide the real-time analytics needed for modern surveillance systems.
- **Data quality is essential.** Last but not least, the quality of the data being captured is jobs 1, 2 and 3. IT organizations must make sure to allocate the required resources to ensure that data is captured at the highest quality possible for proper and meaningful analysis.

**Looking Ahead**

Firms need to equip their regulatory compliance teams with tools and technologies that can keep pace with robot traders. They also need to manage the torrents of data coming at them in various formats – from various venues and media – in order to discover and apply actionable intelligence.

Using big data solutions, capital markets firms can look forward to the following:

- **A complete view of historical activities,** including highly granular details down to very small time intervals.
- **Quick recall, review and analysis of large volumes of historical data** from algorithmic trading programs. This provides a better view of the trading patterns needed to uncover anomalies.
- **Deploy CEP technologies** to uncover practices of market abuse that are hard to detect with existing technologies.
- **Be seen as a proactive player** in the eyes of regulators, the market and customers.

Large global banks in the U.S. are already raising the bar by using big data. Implementing big data and CEP together allows them to capture data from active feeds and interpret the signals in real time. This sophisticated technology fuses intelligence feeds, such as those from Bloomberg, exchanges, ECNs and block trade services. The list does not end there. The ability to bring together streams from market news, information from social media such as Twitter, exchange data, market quotes and a firm’s own executions in real time is extremely valuable.

By tapping into these new tools and data streams, surveillance teams can effectively uncover more – and more complex – patterns of abuse, and move beyond the specific alerts that today only spot insider trading.

**Footnotes**


2. Data throughput rates that can range to millions of events or messages per second.
About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world's leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 171,400 employees as of December 31, 2013, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.

References

- http://www.bigdata.com/blog/

About the Authors

Pritesh Bhushan is a Consulting Manager in Cognizant's Business Consulting practice, focused on the capital markets domain. Pritesh has worked on several projects for large investment banks, providing technology consulting in the areas of compliance and trade surveillance. He has over 11 years of experience in technology consulting to capital markets firms. Pritesh has an MBA from Indian Institute of Management, Bangalore and a B.Tech from Indian Institute of Technology, Kanpur. He can be reached at Pritesh.Bhushan@cognizant.com.

Mrugesh Kulkarni is a Senior Consulting Manager in Cognizant Business Consulting, focused on the capital markets domain. With over 12 years of experience in capital markets, consulting and technology, Mrugesh has worked in areas such as business/IT strategy, project roadmaps, portfolio analysis, and requirements management for large investment banking firms, and has supported multiple sales and business development initiatives. He has an MBA from the Indian Institute of Management, Lucknow, and a BE from Mumbai University. Mrugesh is a member of the Global Association of Risk Professionals and has completed the FRM certification. He can be reached at Mrugesh.Kulkarni@cognizant.com.

About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world's leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 171,400 employees as of December 31, 2013, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.