Using big data analytics to fight financial crime

Turning volume, velocity, variety and variability of data into insight to protect your business
The amount of digital data available is more than doubling every two years. According to a recent IDC study, ‘Extracting Value from Chaos’ a colossal 1.8 zettabytes is to be created and replicated in 2011 alone. The growing volume of multimedia content has contributed to the exponential growth in the amount of data. In today’s digitised world, consumers create their own enormous trails of data by communicating with each other, browsing, buying, sharing and searching. While 75% of the information in the digital universe is generated by individuals, enterprises are responsible for 80% of information at some point in its digital life. Companies churn out large volumes of transactional data and capture petabytes of information about their customers, suppliers, and operations. Millions of networked systems are in place – embedded in devices such as mobile phones, smart energy meters, automobiles and industrial machinery which can communicate and exchange data via the Internet. **How can financial institutions use this data to get ahead in Financial Crime Management, when traditionally, firms utilise less than 5% of available data to combat illicit financial activities?**
What is Big Data?

Industry experts define big data as the tools, processes and procedures allowing an organisation to create, manipulate, and manage very large data sets within acceptable timeframes along with the storage required to support the volume of data. However it is not sufficient to restrict the definition of big data to data volumes alone… the velocity, variety, and complexity of data all play an important role.

Source: Forrester’s Brian Hopkins blog (http://blogs.forrester.com/brian_hopkins/11-08-29-big_data_brewer_and_a_couple_of_webinars)

Brian Hopkins, a leading Business Intelligence (BI) analyst at Forrester Research uses the “four Vs” to give his simple definition some body, which is illustrated in the chart above - if an organisation just has high volume or velocity, then big data may not be appropriate. As characteristics accumulate, however, big data becomes attractive by way of cost. The two main drivers are volume and velocity, while variety and variability shift the curve. In other words, extreme scale is more economical, and more economical means more people do it, leading to more solutions, etc.

Douglas Adams satirised the perception of data analysis being the sole domain of computer scientists. He depicted philosophers creating a single function computer to provide a solution to a problem, waiting a long time for the answer, and understanding neither the answer nor the question. Fortunately, unlike fiction, today’s reality provides us with BI tools that bring data analysis closer to the coal face. New solutions empower users to access enterprise data directly in order to arrive at efficient business decisions without having to wait for IT resources.
Managing big data in Financial Services

Financial institutions, faced with both growing customer demands for improved and more personalised services along with increased regulatory demands now have to cope with petabytes of data. However, the increasing number and sophistication of electronic channels available today provide an excellent opportunity to exploit data and arrive at intelligence-led business decisions in a timely manner.

Recognising that data is a significant corporate asset, a number of organisations are appointing chief data officers and giving in to the pressure of policing the integrity of the numbers (Economist Special Report “Data, data everywhere”). There is an increased focus on data integrity with business leaders wanting more consistency in information and regulators expressing concerns about the quality of data they receive. The demand for data accuracy is clearly visible in the growing references to integrity of data in new regulatory requirements. For example, Solvency II requires insurers to have "internal processes and procedures in place to ensure the appropriateness, completeness and accuracy of the data". These processes and procedures could (and usually do) involve technology, but should also include data policies, standards, roles and responsibilities to ensure data integrity is appropriately governed. Under the Data Protection Act 1998, organisations are obligated to ensure customer data is “accurate and, where necessary, kept up-to-date”.

Chartis notes that new regulatory requirements are especially focused around anti-money laundering and anti-fraud measures, with more initiatives expected to follow. Such regulatory requirements may present opportunities to leverage storage rationalisation options that make more intelligent use of various storage technologies.

Big data analysis is the next step in the evolution of BI tools. These tools are expected to provide financial crime analysts with the ability to interrogate large data sets quickly leading to early detection and even prevention of financial crimes. This capability will reduce costs by eliminating reliance on the vendors of financial crime management software, and enabling the enterprise to interrogate complex queries on random data sets more quickly and efficiently.

While it is important to ensure the integrity of data provided to executive management and regulators, unlocking the insights embedded in the data to better understand customers, competitors and employees represents a significant opportunity to gain competitive advantage. Although regulatory pressures are forcing organisations to improve the integrity of the data, many financial organisations see improved data quality and the use of analytics as an opportunity to fundamentally change the way decisions are made and to use the data for commercial gain.

The benefit of big data analysis goes beyond having large volumes of clean data, in fact it can help minimise fraud. Big data helps build a model; a model built on every incidence of fraud going back many years for every single person, rather than restricting it to a sampling group. Each time there is deviation from the model, it can be re-built to help prevent further fraud.
Financial Crime Management and Data analysis

As financial institutions allow access to services across new channels, the risk of new attack vectors increases; the rise of mobile banking has introduced a new set of devices and banking and mobile app builders have to overcome inherent weaknesses in the design of these devices. As the smartphone market matures and manufacturers begin to take defensive action against “phreakers” and strengthen the security of their devices, new data items will become available which will strengthen fraud detection/prevention rules engines.

Data analysis will be used to determine financial crime management (FCM) solution rules by detecting the correlation between financial crime and attributes of the transaction, or series of transactions.

However, research and experience show that having created a set of rules to detect financial crime one cannot rest on one’s laurels. External fraudsters, for example, are nothing if not persistent, and given the size of the rewards, will spend a lot of time and effort in trying to circumvent an institution’s defences. Money launderers will go to extremes to hide the true source of their funds and internal fraudsters will exploit procedural weaknesses in order to reap their rewards.

There can, therefore be both cost and time constraints when new attack vectors are encountered, the implementer will need to be brought in to further customise the application interface in order to broaden the amount of data available to the FCM solution.

Some FCM software vendors have learnt this lesson and are providing tools where end-user can create their own rules. These client-input rules strengthen the standard rule set provided by the vendor, and are added to the rules engines – though the data that the new rules can be applied against is often restricted to that which the application already knows about.

On occasion a quick rule can be derived without requiring extensive data analysis, for instance a list of specially constructed customer numbers that have “responded” to phishing attacks can be used to identify IP addresses used by fraudsters thus preventing other phished customers from being defrauded. Other situations may require deeper analysis, for instance establishing a correlation between fraudulent cheque deposits and other transactions a fraudster may make between the time of deposit and the availability of funds for withdrawal. This raises a problem, financial crime management (FCM) solutions may not retain enough data to enable data analysis to arrive at meaningful conclusions – which in turn means providing sufficient enterprise data to specialist teams within the financial crime management department without necessarily impacting core production systems.

- The United Nations Office on Drugs and Crime has recently stated that $1.6 trillion dirty dollars are floating around the global economy and that this number is increasing every year.
- The latest annual statistics from the UK’s National Fraud Authority show that more than £38bn has been lost over the last 12 months due to fraud. This amounts to an increase of more than 25%.
  - The public sector (£21.2bn) reported the biggest part of the loss, while the private sector cost the government £12bn, with another £4bn in losses from fraud against individuals.
  - The financial services industry recorded a £3.6bn in fraudulent losses in 2010.

Our experience suggests that the logistics of modifying these databases and solutions can take between 2-4 weeks; for financial institutions this timeline is simply too long. But the real question is then, do you know how long your organisation can afford to wait.
When analysing customer initiated transactions, the entire dataset will be structured – that is, data which has an “enforced composition to the atomic data types”. In the example given above then, as long as the appropriate data sources are available it would be a relatively trivial exercise to establish a relationship between, say, a cheque deposit and repeated balance enquiries on an account followed by an attempted withdrawal of the majority of the value of the cheque deposited.

**In cases of investigating fraud involving employees** it is a chain of transactions that triggers suspicion. A simple example would be of an increase in the number of enquiries made against customer accounts to determine available funds followed by a series of transfers to an account the employee has control of. After suspicion has been confirmed, there may be a large amount of unstructured data (in the form of spreadsheets, word processor documents, and emails) to analyse in order to identify other customers the employee, or anyone they might have colluded with, might have defrauded, adding context to the structured data.

**From an Anti Money Laundering (AML) perspective**, big data provides vendors with opportunities to link unstructured data from publicly available sources, such as social media, with high risk entities. AML data vendors already employ smart analytics to map relationships between Politically Exposed Persons (PEPs) and family/associates. Will the next step for vendors be to sort and aggregate unstructured data on those entities in a way that provides banks with a holistic view of their relationship?

As the amount of publicly available data multiplies, how can banks benefit from mapping this information to high risk entities? Will adding due diligence on account opening by using social media be plausible? What can you learn about your customers using the massive amounts of data available in the world?
“Big data” answer

Traditional data analysis tools require data sets to be available locally, necessitating sufficient RAM to manage the data being analysed. As data sets become larger, this approach increasingly requires more and more data to be stored away from centrally managed database servers increasing the risk of data security failures, data loss and data atrophy, thus reducing the quality of findings based on that data. Furthermore, the types of data used for financial crime management often include sensitive customer data, which needs to be protected, as well as maintained, strengthening the argument that these data sets should be centrally managed.

In-database analytics tools enable business logic, developed and supported by the business to be applied directly to large data sets with minimal involvement from IT departments.

Massively Parallel Processing (MPP) databases are not a recent innovation, the concept was discussed in a 1992 paper by David J. De Witt and Jim Gray where they concluded that “the concept of a database machine built of exotic hardware is inappropriate for modern technology” but they also concluded that the ability to build servers using commodity components, reducing the total cost of ownership, would encourage parallel database systems. As well as providing security, MPP databases enable the central management required to protect and maintain data. Depending on how they are implemented, the also provide faster, more efficient access to large sets of data either by parallelising queries and/or tasks.

Although SQL implementations differ between product vendors, it is a standardised language – making the introduction of new data to an MPP database a DBA exercise. This eliminates dependencies on product vendors having to make the data available to their FCM solution. This simple DBA task can be completed if required in minutes as opposed to weeks, and the data can be made available in seconds offering a massive reduction in the potential financial exposure of an organisation under attack.

MPP databases allow the financial crime management team to exploit large amounts of business data quickly and more efficiently than standard databases. Although IT involvement is required to implement new data sources to ensure the optimal parallelisation of queries and/or tasks, the combination of in-database analytics and MPP empowers the FCM team to respond quickly and provide effective additional business logic in order to combat new attack vectors, further reducing the impact of those new attack vectors without having to rely on third party FCM product vendors.

Map/Reduce, a technology which supports the distributed processing of large data sets across clusters of computers, and designed to scale up from a single machine to several thousand is being used by companies like Google and Yahoo to aid searches. Amazon, as a part of their cloud computing offering, provides data mining capabilities to EC2 customers. Map/Reduce works by assigning sub queries to worker nodes and then reducing the multiple results into one which could allow investigators to discover relationships between unstructured data and anomalous transactions in structured data.

Forrester’s definition of big data is the 4 Vs: volume, velocity, variety, and variability.

- Volume: data available, exceeds physical limits of vertical scalability
- Velocity: decision window small when compared to the decision change rate
- Variety: many different format of data prevalent, making integration expensive
- Variability: many options or variable interpretations confound analysis
The future of Financial Crime Management with big data

Gartner has, over the past year, issued three reports reviewing the capabilities of FCM solution vendors. While all the vendors rank high on capability, the majority have issues with providing adequate levels of customer support. As financial institutions begin to offer services over new channels, and fraudsters continue to exploit existing channels it becomes increasingly important that FCM teams have both a detailed understanding of the capabilities of their existing FCM solutions and the ability to respond quickly to new attacks.

Big data analytics are available. If financial institutions are to arm themselves against the cyber terrorists equipping financial crime management departments with state of the art technologies is imperative.

Conclusion

Big data adds several new high-volume disparate data sources to the chain of information. Several new and enhanced data management and data analysis approaches help in the management of big data. Creating analytics from the available data can help in prioritising information in a big way and also provide its business value. This will in turn help in combating illicit financial activities and mitigate any possible risks.

Most financial services firms utilise less than 5% of the available data to combat illicit financial activities. Why so little? That is because the rest data is simply considered too expensive to deal with. Big data is new because it lets firms affordably dip into that other 95%. If two companies use data with the same effectiveness but one can handle 15% of available data and one is stuck at 5%, who do you think will be better protected? Big data is not a traditional BI tool. It requires new processes, an and a fresh approach to data governance. Only then will organisations be able to make pragmatic decisions while safeguarding their reputation and protecting customer interest.

In this paper we have discussed how big data tools can help address one business problem, Financial Crime Management. Data-driven decision making, using big data allows the enterprise to gain a better understanding of the “physical, societal, financial, and business environment” which will, in turn deliver strategic value to the organisation. Big data requires a top-down approach, board level buy-in. It is important that technology departments and organisations understand not only the value that big data can bring, but also the complexity of projects involving that data.

We have seen a number of approaches to big data projects ranging from strategic exercises to those which create silo data marts designed to address one business department’s problem. Our own product portfolio includes databases that could be considered big data marts.

It is clear however, that technology should respond to the changing nature of business requirements regarding big data. Over time business will gain a better understanding of the semantics of unstructured data and the relationships between unstructured and structured data. Not only should technology support the access to and the availability of big data, but specifically in the area of Financial Crime Management, it can derive value from better analysis of patterns and more accurate and timely intelligence due to the multitude of sources available through big data.
Disruptive IT Trends

Smart computing, often termed intelligent analytics, is all about utilising the incredible processing power of today’s computing environments to optimise business decisions in real time and convert the ever-growing flood of data into both meaningful information and actionable intelligence. The strategic significance of smart computing in today’s globally-connected business environment is compelling: in less than fifteen years, the number of global web users has exploded by more than a hundred-fold, from 16 million in 1995 to more than 1.7 billion today (Source: http://www.cabinetoffice.gov.uk/content/cyber-security). By 2015, there will be more interconnected devices on the planet than humans.

All these interconnected users, devices and sensors are contributing to data volumes that simply cannot be interpreted and acted upon using traditional techniques. The issue is broader than simply a storage issue around big data. What’s required is intelligent data analytics that can interpret data streams in real-time and make appropriate and automated business decisions.

In addition to Smart Computing, Unisys believes there are five additional disruptive trends that are creating new opportunities for IT to deliver value to the business – enabling a step-change in productivity via new styles of interaction, automation, and service delivery.

- Cloud Computing
- CyberSecurity
- IT Appliances
- Consumerisation of IT/Mobility
- Social Computing

For more information on those trends, visit www.DisruptiveITtrends.com
Unisys Credentials

For more than a decade, Unisys has had an anti-fraud and Financial Crime practice that focuses on a holistic and pragmatic approach to CSM. We regularly conduct independent worldwide surveys on financial crime, consumer rust and security issues in financial institutions. These surveys help us provide benchmark and standards for security across geographies and give us sound knowledge of the best technologies and regulatory controls in various regions. In addition, our expertise spans across industries in both the private and public sectors, giving us a comprehensive understanding of enterprise risk and enabling us to offer complete solutions and services for fraud detection and prevention.

Over 1,400 Security experts, 160+ Certified Information System Security Professionals (CISSP)

ISO 27001 certified Security Operation Centres

Enterprise wide coverage via our - Secure Network, Secure Infrastructure, Secure Data and Secure User services

40+ years experience securing people, places, assets & information

Fully managed, complete 24/7 monitoring & management

Unisys recently transformed the infrastructure of a tier-1 UK bank to provide a single view of customer deposits for the Financial Services Compensation Scheme. We migrated over 400M data records, 4 banking licenses, over 35M customers and 63M accounts and orchestrated post-integration flow of 69 different input files.

For more information visit www.unisys.co.uk

©2011 Unisys Corporation. All rights reserved. Specifications are subject to change without notice. Unisys and the Unisys logo are registered trademarks of Unisys Corporation. All other brands and products referenced herein are acknowledged to be trademarks or registered trademarks of their respective holders. 12/11