IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 18, Issue 7 .Ver. I (July 2016), PP 64-67
www.iosrjournals.org

Big Data @Work: A New Whistleblower On Financial Crime, Businesses Are Finding The Best Defence Is A Good Offense.

Kaushik Das, Kaustav Dutta, Archisman Saha
Student, Institute of Engineering and Management, Y-12 Engineering Campus, Salt Lake, Sector V, Kolkata, India

***Abstract**: Due to financial fraud and money laundering, banks and other financial organizations are losing a significant amount of their annual revenue. The emergence of E-Commerce along with the change of customer buying patterns and rapid progress of technology are responsible for increasing in fraudulent transactions. To prevent money laundering and fraud, Big data and Analytics have a significant role to play. This emerging technology has ample potential to support behavioural authentication, which can identify phishing activity and stop fraud and recognize the origin of money in suspicious laundering activities-without hampering routine works. Today, data from every possible corner of the world bombard us. How we bring those data into line and make productive application of it is the pivotal issue. This paper points out the application and merits of Big data technology that can be applied to figure out financial crimes.*

***Keywords**: Analytics Big Data, Data, Data Mining, Fraud Detection, Hadoop, Money Laundering.*

I. Introduction

Fraud along with financial crime can no more be considered as a decent cost of running a business. With recent times fraud schemes are becoming more and more practical, the costs are rising high and customer expectations have been exceedingly high. Along with the financial losses, fraud has led to increase in investigative as well as legal costs, lowers consumer confidence and at the same time destroys the brand image.[1] If these threats are not appalling enough, cybercriminals develop even more refined methods which are very hard for security teams of IT organizations to indentify and take necessary actions quickly. However anti-fraud teams make sure that the people who are requesting for secure information has the correct identity and they are scrutinised before their requests are met. The director of research of GOOGLE, Peter Norvig said, "*We don't have better algorithms. We just have more data*." The concept lies in the fact that the more is the data the more are the ways to detect fraud, as IT teams find ways to use data involving in deeper levels of fraud analysis, all the way down to the transaction level .The goal is to find a unique solution that has the capability to work with the existing relational data which is used to run the business at present time, and consolidate large quantities of less structured data to quickly identify hidden patterns as well as discover new insights and even can make predictions before any troublesome activity is likely to happen. Thus combining both these types of data with the use of relational as well as non-relational technologies is one of the keys to what we generally refer to Big Data at Work. In order to meet these challenges, the banking industry is combating fraud in new ways by making use of big data and analytics capabilities. [2][3][4]

II. What is Big Data?

The experts of industry define big data as the processes, tools and procedures which allow an organisation to compose, process and at the same time manage large data sets within acceptable timeframes in addition to the storage required to store such volume of data. However it is not at all sufficient to shorten the definition of big data to data volumes only… the variety, velocity as well as the complexity of the data also plays a significant role. Brian Hopkins who is a leading Business Intelligence (BI) analyst at Forrester Research makes use of the "four Vs" to give his simple definition (Figure 1). If an organisation has high volume or velocity, then the concept of big data may not be appreciable. The two main drivers mainly are volume and velocity; however variety and variability also shifts the curve. Extreme scale is much more economical, which means more people are capable of doing it, leading to many more solutions, etc. Douglas Adams mocked the perception of data analysis as being the sole domain of computer scientists. He depicted that philosophers create a single function computer which can provide a solution to a problem, however waiting a long time for the answer, and finally understanding neither the answer nor even the question. Fortunately, today's reality equip us with BI tools that can bring data analysis more closely to the coal face. New solutions vest users with powers to access the enterprise data directly so as to arrive at efficient business decisions without having to wait for IT resources.[5]

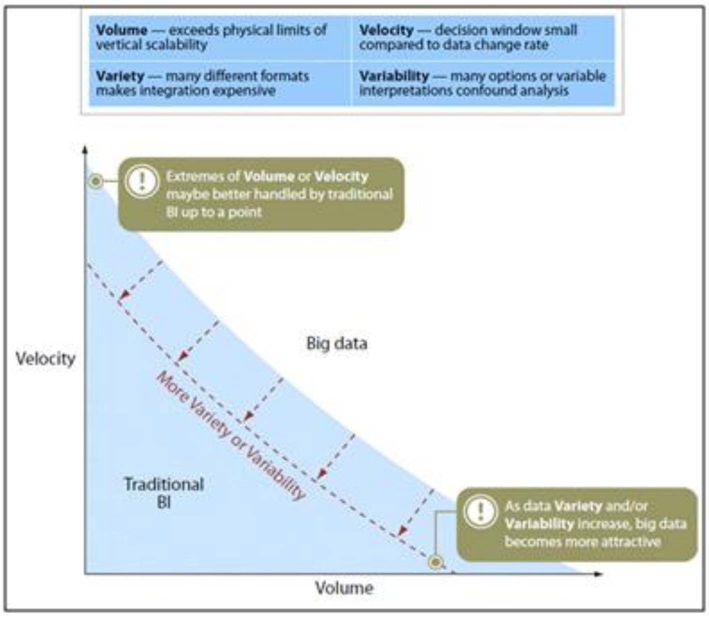
Volume — exceeds physical limits of vertical scalability
Velocity — decision window small compared to data change rate
Variety — many different formats makes integration expensive
Variability — many options or variable interpretations confound analysis
Extremes of Volume or Velocity maybe better handled by traditional BI up to a point
Velocity
Big data
More Variety or Variability
Traditional BI
As data Variety and/or Variability increase, big data becomes more attractive
Volume

Figure 1: Source: Forrester's Brian Hopkins blog.

III. The New Way To Fight Fraud And Financial Crime

Many organizations are vulnerable to these fraud and financial crime because they don't use the modern capabilities to fight these crimes. These capabilities rely more and more heavily on new and improved big data as well as analytic technologies which are now available in the market. With the help of these technologies, banks can easily manage and study huge volume of third-party as well as historical data—far better than they could ever. This ability of analyzing massive data volumes allows banks to built highly accurate predictive models for recognizing and at the same time preventing future fraud also. Organizations also feel the necessity of big data capabilities for examining streaming data in real time. With the help of this technology, banks can also analyze transactions when they occur, detect fraud and can stop it before it is able to cause any serious damage. Therefore by adopting this extensive counter fraud approach, organizations can easily protect themselves as well as their customers while at the same time growing their business. Potential business benefits are substantial, which includes lower addressing fraud cost by means of automation and earlier detection, thus improving operational effectiveness without the increase of staffs and also meeting regulatory compliance obligations. [1]

IV. Why Use Data Analysis for Fraud Detection?

Data analysis software allows fraud examiners to examine the business data of an organisation so as to understand more about how well the internal controls are functioning and also to keep an eye on transactions which indicate fraudulent activity or have huge risk of being fraud. Data analysis may be applied to just anywhere inside an organization specially where e-transactions are recorded and then stored. Data analysis also provides a way to be more ardent in the fight against fraud. Even though whistleblower hotlines allow people to report suspicious fraudulent behaviour but hotlines alone are not enough. Why is it only necessary to be reactive and have to wait for a whistleblower to come forward? Why should people not find out indicators of fraud in the data? By that way, organizations may identify fraudulent activity indicators much earlier and prevent it much before it becomes material and is able to create financial damage. To test for fraud effectively, every relevant transaction should be examined across all applicable business systems. If the business transactions are analyzed at the source level then it helps auditors to provide a better insight and a more complete view of where fraud occurring has maximum probability. It helps to focus and take investigative action to those transactions which are suspicious or which may illustrate control weaknesses that can be oppressed by fraudsters. Follow-on tests should be undertaken so as to increase auditor's understanding of the data and also to identify symptoms of fraud in the data. There is actually a spectrum of analysis that can be used for detecting fraud. It ranges from point-in-time analysis which is conducted in an ad hoc context for one-off fraud investigation or exploration, through to repetitive analysis of business processes where fraudulent activity is more likely to occur. Therefore in areas where the risk of fraud is very high, organizations can use an "always on" or a continuous approach to detect fraud– especially in those fields where preventive controls are not functional or effective. Once an organization gets initiated with data analysis, they actually get interested in digging deeper into the data. Modern organizations have already increased management demands for information and the audit paradigm is undergoing a shift from the traditional cyclical approach to a more continuous as well as risk-based model. Technology therefore provides a wide range of solutions, which varies according to the size and sophistication of the audit. From ad hoc analysis, through to repeatable automated procedures, as well as continuous auditing and monitoring, analytics tries to provide more information on the integrity of financial and business operations with the help of transactional analysis. Technology provides more precise audit reports and also better insight into the framework of internal control, and hence improves the ability to access and manage business risk.[6][7]

V. Four Essentials of an Effective Program

An effective enterprise program comprises of the following phases of counter-fraud measures –[1]

1. Detect: Predicting the fraud before it occurs

Advanced analytics must be implemented to all strategic fraud data to understand if an action is fraudulent before losses actually occur. By looking at small sets of security data, like event logs, reduces the ability of a bank to detect sophisticated crime. An organization's success against internal and external threats largely depends on the pace at which the organization can analyse the data as well as the total amount of data it can analyse.

2. Respond: Applying new fraud insights

The present analytics technology allows banks to take actions in real time and when it matters the most. Organizations, thus, can confidently take legitimate actions while preventing as well as interrupting suspicious actions by responding to criminal activities immediately.

3. Investigate: Turning fraud intelligence into actions

Intelligent investigation of suspicious activity involves performing as well as managing inquiries that are supported by thorough analysis and information accessibility. With the help of these tools, banks have the ability to quickly confirm fraud as a result of which actions like recovery, prosecution and placement can be undertaken.

4. Discover: Leveraging existing historical data

Many banks have a huge treasure of historical data that can be unlocked using the big data analysis. They can search this data for various patterns of fraud as well as financial crimes, and then finally apply the patterns to present activity.

VI. Becoming Less Reactive and More Predictive

There can be many sources and types of errors. Progress has been made in ensuring protection of the networks from cyber attacks; but hacking can be harmful to any business. Hence, multiple layers of protection has been started to be provided to reduce the risk of harmful threats like theft of identity or other important and secret credentials. The increasing focus on prevention of fraud causes the Id management and business rules to flag suspect behavior. Blacklist has become checklist levels of protection in many organizations. This data can be processed by normal data warehouses. But with the sophistication of the threats, improvements of solutions are sought for by companies by taking into consideration more variables to takeout subtle anomalies as quickly as possible. This can be done by big data. Organizations and their security teams must move from analysis of basic queries or business rules into a more advanced state that can sniff out trouble as its happening. Patterns and relationships among apparently unrelated information are identified in the data by predictive analytics. This helps in the analysis and anticipation of behavior and takes actions to prevent them. These real time systems increase the amount of data to be collected, stored and analyzed. Several criteria are to be taken into account to make the approach more sophisticated to reduce fraud and theft. The first key element is the data itself. The historical data, customer transactions, logs, chat transcripts, email, and other data are to be accessed. The second requirement is the infrastructure to handle such huge amount of data. The relational database when made to work with a non-relational platform like Hadoop, can handle such high amount of structured and unstructured data. To identify the patterns among the data, data-mining and statistical analysis tools are used by the IT teams. They operate on the data that are collected. Thus the third requirement is the tools for the analysis of the data. Oracle Advanced Analytics can discover relationships hidden in the data, including unstructured data in Hadoop environments. Oracle data mining helps in uncovering of suspicious submissions of reports and forms, reducing fraud. [2]

VII. Relational and Non-Relational Data: A Powerful Combination

The use of big data has increased rapidly to every organization. Previously, values from data were extracted by careful selection and standardization of the collected data based on some predefined relationships.

But with the rapid increase in the volume, variety and velocity of data, new ways of analysis of data has been developed. The usefulness of a particular data is often unknown now. Formation and testing of hypothesis is quicker in non-relational environment, thus having a better understanding of the data which would not have been possible otherwise. This change can lead to better results. Hadoop complements a relational database. It has ability to store and process huge amount of non-relational data. Relational technologies helps in organizing the data to make it do specific things, whereas non-relational technologies takes the data in its raw form, as it is and tries to identify what can be done with it. Thus the first technology helps in running the business, whereas the second one helps in changing the business. When they work together, they become a very powerful combination.[2][8]

VIII. Conclusion

Several new data sources are added to the chain of information by Big Data. Analysis of the available data helps in prioritizing the information, improving its business value. This helps in reducing its vulnerability to risk of theft and fraud. Only 5% of the available data is utilized in many financial firms for combating illegal activities. The usage of data is so little because the data is considered very expensive for dealing with. Big data helps the firms to utilize more of the remaining 95% of the data available, thus increasing its safety, making them more protected [3]. It is a better approach to data governance. The companies will then be able to take more realistic decisions and protect customer interest. Data-driven decision making using big data enables the companies to have a better understanding of its environment. They can then in turn improve their business strategies accordingly to improve. Big data has a top- down approach. The complexities of the projects that involve the data are of utter importance. Hence it is important to change the business requirements regarding big data. A better understanding of the semantics of unstructured and structured data will be gained by over time business. Big data thus helps to have a better analysis of the patterns of the data. It helps in more accurate and quick analysis by the multitude of sources available.

Reference

[1]. Fighting fraud in banking with big data and analytics, White paper IBM Software.
[2]. Big Data @Work White paper Oracle, Intel, Custom Solution Group.
[3]. Halevi, G., & Moed, H., "The evolution of big data as a research and scientific topic: overview of the literature. Research Trends",
[4]. McAfee, A., and Brynjolfsson, E., Big Data: The Management Revolution, (2012), Harvard business review, 90 (10), 60-68.
[5]. Using big data analytics to fight financial crime,Unisys.
[6]. Fraud Detection Using Data Analytics in the Banking Industry, Discussion Paper, ACL.
[7]. Mulani, N., The Million Dollar Opportunity: Reaping Returns from Analytics. Information Management, (2013).
[8]. Seddon, P. B., Constantinidis, D., and Dod, H. (2012). How Does Business Analytics Contribute to Business Value?