ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 2, Issue 10, October 2017

Security Issues And Prevention Techniques From Data Analytics

[1] J.Bagyalakshmi, [2] S.Kamalakkannan, [3]P.Kavitha
[1] Research Scholar Vels University Chennai – 600117 Tamil Nadu, India.
[2]Assistant Professor Vels University Chennai – 600117 Tamil Nadu,India.
[3]Assistant Professor St.Joseph College Chennai-600119 Tamil Nadu, India.

Abstract: ***The big data portent is a direct result of the digitization and "datafication" of every activity in personal, public, and commercial life. The major downside is loss of data privacy- Online and offline actions are being stalked, aggregated and explored at dizzying rates. Flipkart E-book can know the details of books you ever bought or viewed by evaluating big data gathered over the years. The NSA (National Security Agency) can track the phone number that you ever dialed. With the arrival of many digital modalities all these data has grown into BIG data and is still on the upswing. Eventually Big Data technologies can exist to provide greater insights in business decision making faster when needed but with the snag of loss of data privacy. In this talk the big data security issues and the extent of financial fraud is discussed.***

Keywords ***:*** **Big Data, security, privacy, security Practices in E-Commerce.**

I. INTRODUCTION

The usage of internet grows day by day. The survey of www.internetlivestats.com says that around 40% of the world population has an internet connection today leading to bulk "data fiction" or digitization and thus arrived the term "big data" .Big data is not just big. It's
also diverse data types and streaming data.[1]. Big data refers to large data sets which traditional data processing application software such as Relational database management systems and visualization-packages are not able to gather, manage and process within a tolerable pass by time. The target users of big data are Scientists, Business people, Medical practitioners, Advertising Media and Governments. Simple application of Big Data- A Facebook can and will analyze big data and tells the birthdays of people. The major benefit that big data provides for its users is thoughtful and faster decision making and to provide greater insights when needed but with the downside of loss of data privacy.

II. BIG DATA SECURITY ISSUES

Security and privacy issues are overstated by velocity, volume and variety of big data. Large scale cloud infrastructures, the various data sources and diversity of formats, flooding nature of data acquisition and high volume inter-cloud migration roots to the security issues. The use of large scale cloud infrastructure with different software platforms, spread over the large networks of computers also increases the outbreak surface of entire system. Therefore traditional security systems, which are personalized to secure small scale static data are inadequate.

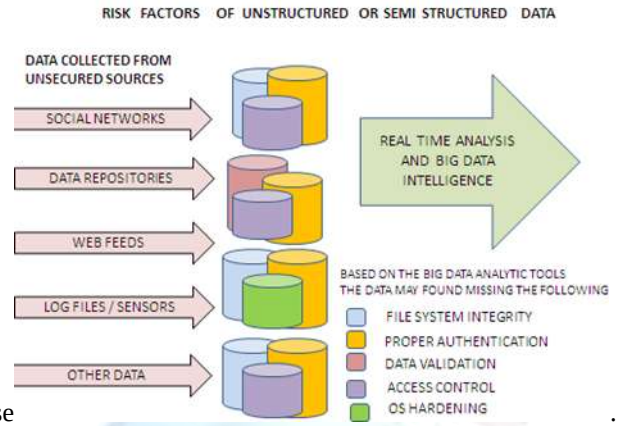
Security of Big Data- a hundred million dollar question:
Being the head of the governmental entity and responsible for protecting U.S. consumers, Ramirez in her talk stated-[2] "The larger the concentration of sensitive personal data, the more attractive a database is to criminals, both inside and outside a firm. The risk of consumer injury increases as the volume and sensitivity of the data grows."

As per the previous year survey report the successful cyber attacks on midsize-to-large companies is more than 60%- according to Cyber Edge[3].The Verizon 2015 Data Breach Investigations Report (DBIR) equated approximately 80,000 security instances, including 2,122 confirmed data breaches.[4]Security breaches affect organizations of all sizes and types. A report says that an individual breach affected 1 million to 10 million records a few years ago and today, in the age of mega-breaches, a single instance of breach can involve 200 million records—or more.
Since the restrictions of standard IT security practices are transparent, the assailants uses software subversion to inject malicious software into applications and operating systems, which is a serious and growing threat whose hostile impact is

intensified by big data[.9]. Now the Hundred Billion Dollar question is what security and privacy technology will be adequate for controlled assured sharing the direct access of Big Data.

Privacy and Security:

With the large volume of personal data such as the buying preference of medical/ healthcare records, site based behavior of customers collected by big data applications and transferred over networks, the question of privacy and security naturally arise .



This leads to an urgent demand on technologies that strive to enforce privacy and security in data transmission. The more of number of resource of huge volume of data requires a new generation of encryption solutions (e.g., homomorphic encryption).

On the other side, big data technologies can also be used to explain the security challenges in networked systems. Data traffic of specified pattern in the network are usually generated by the network attacks and intrusions. By analyzing the big data collected using different sources on different platforms, using network monitoring system, those misconducts can be identified proactively, thus greatly reducing the potential loss.

Extent of financial fraud

A list of top trends in fraud across the world has been reported by Experian [5], a reputed global information services company, based on the survey conducted in 2014 on financial fraud.

The trends that are listed gives an idea of how the fraudulent attacks have been evolving. Here are some of the salient findings from the report.

- The rate of detected frauds has increased from 24 to 35 frauds per 10000 applications. So, data and identity are becoming increasingly more vulnerable.
- Of all types of fraud, identity theft has increased at the fastest rate. From 30% in 2013, incidents of identity theft has risen to 47% in 2014.
- Third party involvement was the most notable finding in frauds related to cards, loans and savings products.
- Mortgage application frauds accounted for the highest among all detected frauds. In August 2013, 82 of every 10000 mortgage applications were found fraudulent.
- Current account application fraud increased by 60% from August 2013 to August 2014.
- Third-party involvements have been on the rise in the case of all types of frauds. Since August 2014, third-party involvements accounted for 79% of all fraud discoveries.

ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 2, Issue 10, October 2017

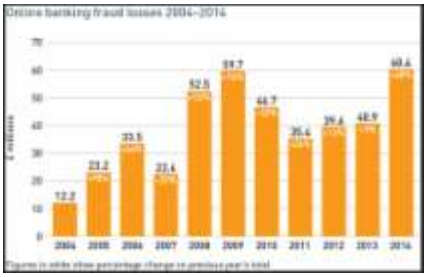


Fig. 1: online banking fraud losses between 2004 and 2014 [7]

According to AFPs (**A**ssociation for **F**inancial **P**rofessional) payments fraud and control survey, following image shows the statistics for payment frauds in 2015.

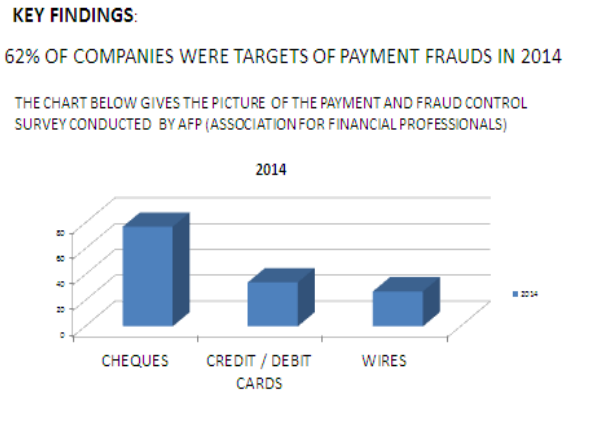


Fig. 2: Payment fraud statistics in 2014[8]

III. DEPLOYING BIG DATA FOR FRAUD DETECTION

A. The role of Big data Analytics in preventing fraud

Financial fraud methods are becoming more erudite and the techniques to contest such attacks also need to advance[7]. Financial institutions have been investing on secured technologies that prevent and combat fraudulent attacks. But the extent and nature of financial fraud continues to change. Eliminating financial frauds is not a practical goal, but there is a need to evolve the ways fraudulent attacks are to be handled. Clearly, the methods need to change and make big data, a novel and potentially potent weapon.

B. Role for big data analytics in helping insurance companies and E-commerce companies to find ways to detect fraud:

(i) INSURANCE COMPANIES

Insurance companies may need to stop fraud early. By designing predictive models based on both historical and real-time data on salary of an individual, health claims, counsel costs, demographics, weather data, voice recordings and call center notes, companies are able to identify suspected fraudulent claims at the earlier stages.

For example, a personal injury claim may be a fake medical claims or a staged accident. Companies might have experienced an increase in sophisticated crime rings to commit auto insurance or medical fraud. These rings may have

similar methods of process that are enacted in different zones of the country or using different pseudonyms for the claimants.

Big data analysis can easily find match for patterns in historical claims and identify similarities or raise questions on a new claim before the process gets into its full swing.
An actuary helps the companies in identifying fraudulent claims at the first notice of loss. A team of experts at insurance companies with the skill of identifying risk and fraud along with actuarial analyst and business managers view big data analytics as having all potential to deliver a huge benefit in anticipating and decreasing the attempted fraud and decreases.

Consider the following example. An insurance company may need to improve its way to make real-time decisions more quickly and accurately, when deciding how to process a new claim. The company's cost expenditure including legal process payments related to fraudulent claims has been rising consistently. The company has extensive policies to help the actuary in evaluating the legitimacy of claims, but the actuary often did not have the data at the right time to make an informed decision.

The company gears a big data analytics platform to provide the integration and analysis of data from multiple sources. The platform incorporates extensive use of social media data and streaming data to provide a real-time view. The Call center agents are able to have a much deeper perception into possible patterns of behavior and relationships between other claimants and service providers when a call first comes in.

For example,

An agent may receive an alert about a new claim that indicates the claimant was a previous witness on a similar claim six months ago. After identifying other unusual patterns of behavior this information may be presented to the claimant and the claim process may be halted before it really gets going.

In other cases, social media data may point out that the conditions described in a claim did not take place on the day in query. For example, a claimant may quoted that his car was totaled in a flood, but credentials from social media showed that the car had actually been parked in another city on the day the flood occurred.

Nowadays Insurance fraud is incurs a huge cost for companies that executives quickly move to incorporate big data analytics and other advanced technology to solve the problem of insurance fraud. Insurance companies not only takes the impact of these high costs, but also have a negative impact on customers who are charged higher rates to account for the losses.

By using big data analytics to look for patterns of fraudulent behavior in huge amounts of unstructured and structured claims-related data, companies are able to identify fraud in real time. The return on investment for these companies can be huge. They are able to analyze complex information and accident circumstances in minutes as compared to days or months before implementing a big data platform.

(ii) E-COMMERCE COMPANIES

The use of Big Data can contest fraud in E-commerce companies in three main ways.

Analysis should be done with all the data. In the past, retailers used a sample or a subset of their data for analyzing the fraud. It took much time and money to use the complete data set. With Big Data, new data sources can also be introduced apart from analyzing the fraud. Analyzing the full data set leads to several benefits: (a)Based on the defined fraud rules a review of all transaction can be done (b) New fraud patterns can be identified and added to the growing list of fraud rules; (c) False positives should be minimized to avoid revenue loss and turning away customers.

For example, the delivery of product can be analyzed by big data by analyzing the data streams from social networks to minimize return fraud. Big data can conduct image analyses by partnering with third parties like eBay and Craigslist to

access their listings. All this data can be gathered and analyzed using Big Data tools like Hadoop. Real-time fraud detection. Online transactions are combined with data from other sources, like existing data warehouses, to detect fraud in real-time. This can prevent credit card fraud where the transaction is curtained against a set of pre-defined fraud rules as part of the credit card authorization. This includes combining spot data with data from customer's social feed, the geo data gathered through customer's smart phone apps, purchase history, and web logs. The purpose of doing this is to decline the fraudulent authorization done in real-time. As a part of the real-time authorization process new fraud patterns can be automatically added to the set of fraud rules and this can be done by enabling big data solutions to analyze the historical transactions happened in the previous weeks, months or years. The recent reports of Visa identified $2 billion annual incremental fraud opportunities by using vigorous fraud management system that identifies more than 500 different patterns of fraud transactions. The alternative method of preventing fraud on online purchases/ real- time transactions is by processing the streaming data from sensor attached to high-value items to transmit their location. This helps in minimizing the return fraud, as the retailer now knows exactly when the item was delivered to the customer. Use of visual analytics: Even though the data is derived from different data sources, the capability of visual analytics- an another Big Data tools offer the capability to visually analyze data and derive insights. Retailers can use these tools to identify the customers, products and the areas that have a higher fraud rate based on historical analysis.

This tool helps the companies in identifing the areas where time and money should be invested to minimize fraud. Visualization also reduces manual efforts to reviewing every order. The graphical reports can depict the probability of fraud for each order transaction and connect to email or SMS alerts for escalation, as needed.

Case study of Alibaba the Chinese Ecommerce Company

Alibaba, the Chinese ecommerce company, has been effectively using big data to tackle fraud. At Alibaba, any attempt of potential fraudster has to pass through 5 stages of verification which is a tough scheme. These five stages are (1) Account Check, (2) Device Check, (3) Activity Check, (4) Risk Strategy and (5) Manual Review.
Each stage involves the use of huge volume of data related to the seller activities.
For example, in the first layer of verification, several questions may be asked such as the details of last order and so on to track the suspicious activity . The second layer inspects the devices such as the IP and device processes, number of devices the seller is possibly going to use and so on.

IV. CONCLUSION

Thus Big data analytics provides Insurance and finance companies the opportunity to prevent fraud to a large extent and helps in security. However, the usage of big data for this purpose is till at its early stages and a lot needs to be done in this regard.

References:
[1] "Big Data Analytics" – by By Philip Russom, fourth QUART ER 2011TDWI best practices Report

[2] "The Privacy Challenges of Big Data: A View from the Lifeguard's Chair," speech by FTC Chairwoman Edith Ramirez, 2013
[3] "2014 Cyberthreat Defense Report," CyberEdge Group, 2014
[4] "2015 Data Breach Investigations Report," Verizon, 2015
[5] KDnuggets.com- Reports on " How to combat financial fraud by using big data?" (16:n11) by By Kaushik Pal, TechAlpine.
[6]Source: www.theukcardsassociation.org.uk/plastic_fraud_figures/
[7] Source: http://www.afponline.org/fraud/
[8] "The Truly Personal Computer," The Economist, 2016.
[9] WenguangChai . Analyzes and Solves the Top Enterprise Network Data Security Issues with the Web Data Mining Technology. 2009 First International Workshop on Database Technology and Applications, 2009.