Thodly Dugue
Intro to computer security
Research paper
Dr. Haydar Teymourlouei
Monday, November 21th, 2016

How big data could help prevent cyber attacks

Outline

- Introduction
- Big data
- How big data works
    - Benefits of using big data
- How big data help to prevent cyber attacks
    - JIT response
    - IBM solutions
    - Hadoop
        - Propose solutions/algorithm
    - Cyber ontology
        - Attack event
        - Log file /event file
        - Port scanner
    - Cyber warfare
- Conclusion
- Acknowledgment

Introduction

We know that nowadays every business is dealing with large number of data, whether it is a hospital, a school, a grocery store a bank or the government, they are all collecting information. We can say with the evolution of technologies those data as increase enormously. The use of big data analytics helps you to structure those data; because most of the time those information are personal or sensitive information and you need to make sure that those data are safe from insider or outsider intrusion. This research project is about how to manage big data analytics. How can we protect those data against cyber-attacks and what are the vulnerabilities. What are the characteristic of a big data analytics? What are the requirements to use them? How big data analytics can help you to prevent fraud? What are the methods used to prevent cyber-attacks.

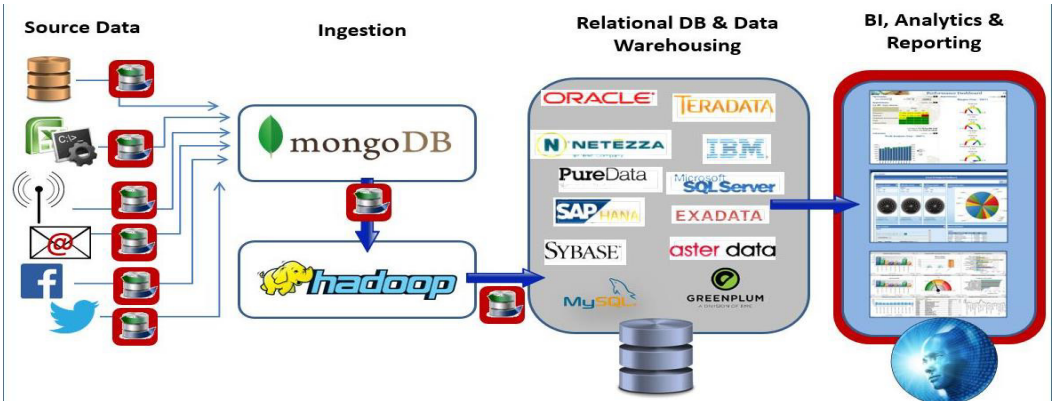
Big data

Big data is used to describe large amount of data or complex data. It is characterized by four "V's", which are volume, velocity, variety, and veracity. "Volume" is the large number of data collected. "Velocity", when talking about velocity, we see the process of analyzing the data collected and also it means that those data are collected at a continual interval. It catches and organizes them. "Variety" is about how many types of information are being collected; structured and unstructured data. Those data are collected from everywhere, from email, mail, telephone, social media, internet, data center or in person. "Veracity" is questioning if the data are accurate, authentic or not.

How big data works

Big data helps to analyze large number of information and also help to take better decisions. Using big data give you special tools that make it more accessible to analyze those large numbers of information. The best known companies for the management of big data are IBM, Teradata and Oracle. The most common tools are Analytic platform and Hadoop which is an open source technologies. Big data collect the information from the data source; like email, social media,

internet and etc... Then come the ingestion process where all the information are put together to create the appropriate data base. And in the relational database they will do the management process where the information will be analyze and coded if needed and the result will lead to the final part Analytics in reporting. In the last part you will have a better understanding of your big data and you will also have a clue about the security level.



Source: http://www.querysurge.com/solutions/testing-big-data

- Benefits of big data:

1. Predict workload
2. Reduce wasted resources
3. Reduce cost
4. Make process more efficient
5. Detect safety issues
6. Detect behavior before it affects your organization.
7. Security managements

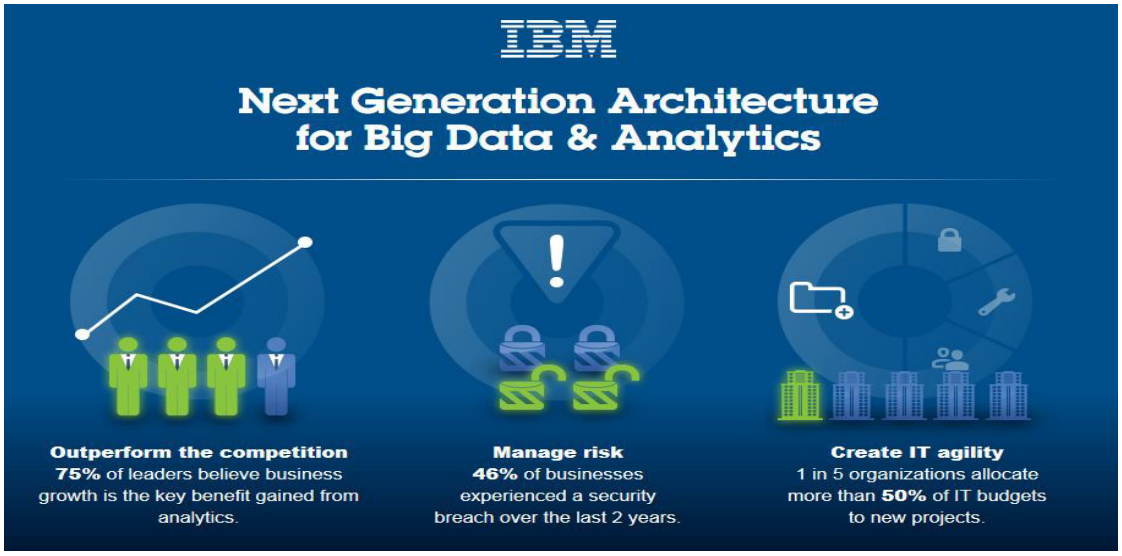
How big data helps to prevent cyber-attacks:

With big data you can have the control of the security inside and outside of your network system, it prevents data loss. For that they use the JIT response which stand for "just in time" response. They say when you put big data and JIT response together they give you an acceptable level of security. The actions taken by big data are: incoming message block, outgoing message block, SSL inspection and lock user's account. To prevent fraud big data also have a "big data log analytics". When big data log analytics is combined with JIT analysis it collects information on the machines that are working in the same system that have an open connection with locations outside the local network. Big data log analytics also predict future attacks and gives you information about previous attack that might have occurred on your system.

IBM solutions

An IBM report shows us that 46% of businesses are experiencing security breach; which shows that the need to protect our information is very high. Also let us not forget that there are

people that are just waiting for the right moment to steal important information from those companies to gain money or just to test their skills. When dealing with personal information, you will always find people that are willing to do everything to be able to get through those files.



IBM came with a solution to that problem the IBM security intelligence with big data. For them we should be in position to protect our data from threats and fraud. Their solutions help to detect risk and intrusion; it analyzes structure and unstructured data. On the IBM official website about big data analytics they tell you what they use and how it works: "IBM Security Intelligence with Big Data combines the real-time security visibility of the IBM QRadar Security Intelligence Platform with the custom analytics of the IBM Big Data Platform. QRadar performs real-time correlation, anomaly detection and reporting for immediate threat detection, and also sends enriched security data to IBM big data products, such as IBM InfoSphere BigInsights. Information is subsequently fed back to QRadar, providing a facility for closed-loop, continuous learning"

Hadoop

Hadoop is a Java-based programming framework that helps to store, organize, process, and analyze data. CounterTack which is a big data endpoint detection; helps to analyze system-level information collected from any kind of data to detect intrusion or malicious behavior. CounterTack use Cloudera which use Hadoop algorithm to protect against security threats. Hadoop allows us to get through data and to manipulate them. It gives you a complete access to the information like this you can know what is going on in your file. In the CounterTack's website it said that: "Endpoint data is collected and analyzed to provide security teams with the information needed to identify and mitigate threats."

Cyber ontology

The ontology system collects data and put them in an ontological database. In the ontological database the data will be coded or encrypted. When talking about cyber ontology we also see the words attack event ontologies. To have a better understanding of cyber ontology they usually use a model called "The Diamond Model". In this model there are four keywords; victim, infrastructure, capability and actor. The actor is the one that need our attention the most. The

actor is the attacker. Terry Janssen, cyber strategist and Nancy Grady, data science had made a research paper about "Big data for cyber-attack management" In that research they have found nine ways to prevent from data loss using cyber ontology:

1) Access control: That give you the necessary information about who accessed the system, when was that and where.
2) Backup and restore: in case of loss you will have most of your information safe somewhere else.
3) Secure coding techniques: encrypted data or information using specific software.
4) Inventory of all devices and software.
5) Risk assessment
6) Configuration management data
7) Continuous monitoring events: using log file or log events.
8) JIT response (Just in time response)
9) Training

Cyber warfare

"Cyber warfare involves the actions by a nation-state or international organization to attack and attempt to damage another nation's computers or information networks." Source: rand and cyber-warfare. Big data has an important role when we have to face cyber warfare. Cyber warfare include digital intrusion and data loss; by using tools that give you a better understanding of large amount of data and that helps you to manage them easily big data helps us from this type of security threats

Conclusion

Now, we have a better understanding on how big data helps us to protect our data from malicious threats. In general we can say that big data helps us to protect our data by identifying malicious activities in computer environments, by improving the analysis of personal information and by improving the maintenance and the storage of data. Big data allows you to manipulate those data and to be in control of everything. We also show that Hadoop algorithms are used in all platform or software that deals with big data. Other methods are used to protect against intruders attacks; like cyber ontology, where the data are encrypted and especially the JIT response or just in time response that gives us another level of security. Big data predict cyber-attacks and also gives you information about previous suspicious behavior that you had on your system or just tells you about the anomalies on your network system.

Acknowledgments

- http://www-03.ibm.com/security/solution/intelligence-big-data/
- http://www.querysurge.com/solutions/testing-big-data
- https://www.actiac.org/system/files/Big%20Data%20and%20Cyber%20Attack%20Management%20-%20SAIC.pdf