Risk Processing Framework for Big Data Security in the Enterprise

*Miza Shazwani Kamarulzaman[1] , Nur Azaliah Abu Bakar[2], Hafiza Abas[3]

[1,2,3]Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia
Jalan Sultan Yahya Petra, 54100 Kuala Lumpur Malaysia
[1]miza1994@graduate.utm.my, [2]azaliah@utm.my, [3]hafiza.kl@utm.my

Article history

Received:
23 Sept 2019

Received in revised form:
2 Nov 2019

Accepted:
20 Dec 2019

Published online:
30 Dec 2019

*Corresponding author
miza1994@graduate.utm.my

Abstract

Big Data is about the growing challenge facing organizations in dealing with large and rapidly growing data or information sources that also present a complex range of analyzes and problems. Big Data creates critical security information and privacy issues, whileThe research process for large data to reveal hidden patterns and privacy referred as Big Data Analytics. Big Data is useful for organizations to gain deeper insights and benefit from the competition. Therefore, the implementation of large data must be analyzed and implemented as securely as possible. This paper gives an overview of Big Data Analytics privacy and security issues existing risk assessment model and overview of risk processing framework in Big Data.

Keywords: *Big Data, risk assessment, IT security, Big Data Value, Information Technology Enterprise*

1. Introduction

Big Data in an enterprise provides organizations with complete customer profiles that enable customer experiences on every point in contact throughout company's entire journey [1]. Big data in a business has been removed so that companies can get a unique view of consumers that contains numerous concise, measured and industry-specific indicators to produce a detailed record of the actions of each customer. Data space is a standout amongst the most encouraging ICT areas with generous desires both in favour of market developing and configuration move in the territory of information stockpiling management and investigation [2].

This research emphasized on identify risk in big data and describes more risk mitigation strategies for system quality assurance and maintenance. This study also provides risk assessment process. The research engines were explored to answer questions such as lists of risks, example of threats in big data, and how to mitigate the risks. This research also shows the existing security framework to mitigate any attacks that possible happen to the big data in the Enterprise.

* *Corresponding author. miza1994@graduate.utm.my*

2. Risk of Big Data

2.1. Data Privacy

The greatest challenge for Big Data from a security perspective to protect the privacy of the user. Big data often contain enormous amounts of personal identifiable information (PII) and the privacy of users is a major concern [1]. User authentication and access must be strictly supervised since data in organizations usually in a shared environment maintaining the Integrity of the Specifications. Big data analytics facilitate violations of privacy. It is a fact that the consequences of data protection for end-users are not yet fully understood [3].

2.2. Regulatory Compliance

Another possible risk in big data is related to compliance with regulations, especially data protection laws. Particularly, if data can be stored or processed in some jurisdictions, these laws are stricter than others. Organizations must consider carefully the legal consequences of their data storage and processing in order to comply with the regulations they face [4]

2.3. Collect and Process Sensitive Information

One of the security issues related to Big Data is collect and process sensitive information about customers and employees, property, trade and financial information. As companies seek to gain value from such information, they seeking to aggregate data from wider range of stores and applications in order to create a wider context for increasing data value [3]

2.4. Threats and Fraud

Another security advantage is that big data sections can be exploited for security events[1]. The industrial information age with open and interconnected technology brings benefits such as convenience and efficiency while posing potential threats to security, such as the spread of viruses, Trojans and other threats in the industrial system [5].

2.5. Privilege Escalation

For the privilege escalation issue, it might have over-special accounts that increase the danger of insider attackers in Big Data. As an example, the administrators should not have full access to Hadoop bunches and every one of their information. Rather, similarly, as with the slightest benefit get to, administrator access should be restricted to the explicit activities and directions required to carry out the command. This implies authorizing a narrow set arrangement of access and benefit rights than the local root account permits [6]

2.6. Lack of Visibility

The lack of visibility in the Big Data Analytics tool which is the Hadoop cluster create challenges for IT companies. Without recording the sessions, it is almost

impossible to identify, mitigate and remedy potential security problems. Therefore, it is difficult for IT organizations to prove compliance with regulatory and standards requirements without auditing capabilities. Since Big Data implementations comply with certain standards, IT organizations must implement auditing capabilities [7].

3. Existing Risk Assessment Process

In this section, the authors explained the existing security risk assessment in the Big Data Analysis. According to ISO 31000:2009, the risk is defined as "uncertainty effect on defined goals". It means that objectives must be defined or known before risks can be defined [3]

These objectives are usually defined taking into account the institutional context of the respective organization. As examples are the advents of Big Data Analysis security problem. The traditional issues related to the protection of data integrity, availability completely outsourced infrastructure and deployment plan on security and privacy [7]. Below are several lists security risks model existed in Big Data to the enterprise security:-

3.1 Security Risk Assessment Process

Figure 3.1 shows four (4) risk mitigation strategies are available to security which are preventing, reducing, spreading, transferring and accepting. Most security programs are designed to reduce risk [4].

Figure 3.1 Security Risk Assessment

Threat Assessment: This assessment is a logical process to determine the probability of events affecting your assets and to validate levels of security. We use various sources of data to evaluate conceptual threats [8].

Vulnerability Assessment: Analysis of safety weaknesses and adversarial exploitation opportunities in one or more of the categories mentioned above. The

aim of a vulnerability assessment is to identify and block the possibility of asset attacks [3]

Security Risk Modelling: The objective of a Security Risk Modelling is to develop a model that incorporates the variables to identify risks to people and inform security decisions at each site. The purpose of a security risk model to maximize protection by concentrating on the factors that have an effect on security risk. [9]

Liability Analysis: The objective of liability analysis is the authors' evaluations include detailed crime analysis, including predictable property and area crime, vulnerability identification, risk mitigation strategies, and cost-effective security solutions. The written report will be provided with findings and recommendations for reasonable security measures [10].

3.2 Multi-tier Risk Prevention

There are five layers of multi-tier risk mitigation to avoid fraud during transactions shows as in Figure 3.2 below [10] :

Figure 3.2 Multi-Layer Risk Prevention Framework

Account Check: The account check is the first layer which includes buyer account data and seller account information [11]

Device Check: Device check is to check if there a large number of transactions from the same device and any transaction involve in bad equipment [10].

Activity Check: Activity check is tot check the pattern of historical records, buyer and seller behaviour, linking accounts, devices, and scenarios.

Risk Strategy: Risk strategy is the fourth layer and assesses actions on the basis of previous analyses. Some are sent to self - determination because of obvious fraud [10].

Manual Review: Suspicious cases are manually examined without strong evidence, where more information is disclosed and telephone calls to remind buyers can be made [10]

3.3. Risk Assessment Process

Figure 3.3 below shows the risk assessment process in the IT industry that has been proposed by researchers.

Figure 3.3 Risk Assesment process

Risk Assessment: The objective of risk identification to determine what could cause potential loss and to understand the process of loss could occur. In the security risk identification activity, the web application assets to be managed and a list of asset-related business processing, threat sources, potential vulnerabilities, the likelihood of threats, existing controls and the effects of asset impact must be simultaneously considered [12]

Risk Estimation: Risk estimation uses a scale, descriptive or numerical, to describe the extent of asset loss and the likelihood of impact. The likelihood of impact and the extent of the loss of assets are combined to create risk levels and reveal the major risks [12].

Risk Evaluation: During the risk assessment phase, estimated risk levels should be compared with the criteria for risk assessment and criteria for accepting risk. For the proposed controls, a cost-benefit analysis could be carried out to show that the cost of implementing the controls can be justified by reducing the risk level. In addition to the estimated risks, other factors can be taken into account are contractual, legal and regulatory requirements. The output of the activity is a list of risks prioritized according to risk assessment criteria [12]

3.4. Existing Security Framework

Below is brief description of some components of the existing security framework model in Big Data Analytics for answer the second research question of this study:

3.4.1. SIEM for Network Analysis Visibility (NAV)

The organization should decide who can view the data or give them physical access in the case of an administrator. They should supervise continuously and changed as

employees change their job in the organization so that they do not accumulate excessive rights and privileges that abused in order to ensure access controls are effective [11]. This is done using Security Information and Event Management (SIEM) technologies that collect log information from a wide diversity of applications on the network. SIEM systems collect aggregate and filter alarms from many intrusion detection sensors and other sources and provide security analysts with actionable information. SIEM systems extensively use external data sources [11]. One of the vendors known as AlienVault that expanding its capabilities called Network Analysis Visibility (NAV) which helps to make the SIEM tools more practicable. NAV provides the ability to capture network traffic and look for potential threats and insider attacks [13]

3.4.2. AMSEC

The authors proposed AMSEC (Attack Modeling and Security Evaluation Component) as one of the key SIEM subsystems and the techniques used to perform the main AMSEC functionality as showed in figure below :-

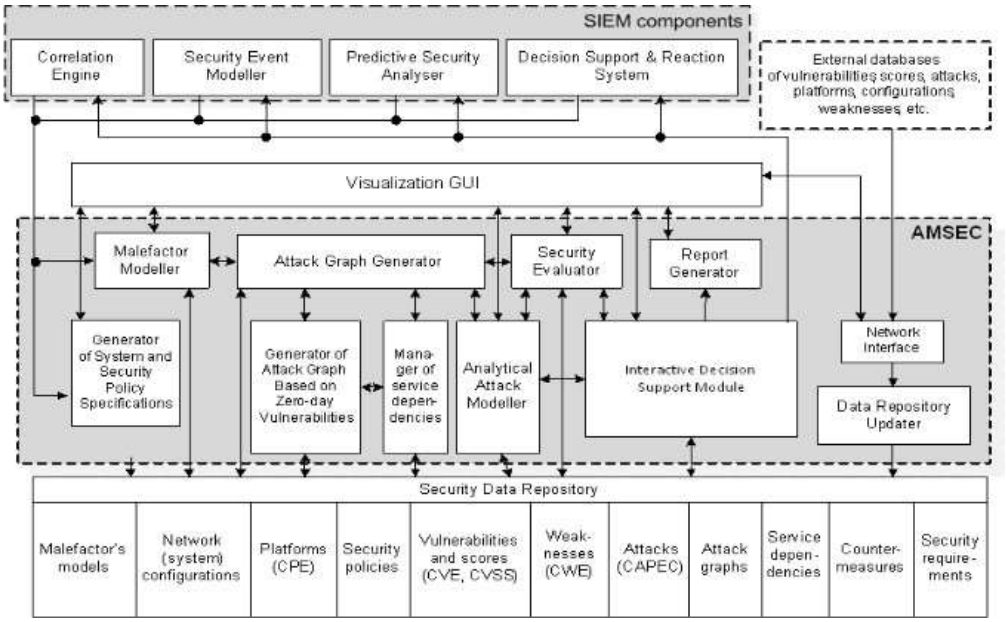

Figure 3.4 Attack Modeling and Security Evaluation Framework

Network interface: Helps to support interaction with external environment [11].

Interactive Decision Support Module: The interactive decision support module enables the user to decide on counter measurement solutions [11].

System and security policy specification generator: Change to internal representation the network configuration and security policy information received from the data collection and correlation components or the user [11].

Data repository updater: National Vulnerability Database (NVD) downloads and translates open vulnerability databases, attacks, settings, weaknesses, platforms, countermeasures into AMSEC [11].

Report generator: The report generator shows vulnerabilities that has been detected by AMSEC, appoints infirm locations known as the majority of attacks, generates recommendations to increase safety levels [11].

Malefactor Modeler: It is responsible for the modeling of the malefactor and used in the design and operation of AMSEC.

Attack Graph Generator: It is used for building graph attacks. The Topological Vulnerability Analysis (TVA) is used to generate an attack graph [11].

4. Security Risk Assessment Process for Big Data

This section shows the overview of risk assessment process for Big Data by researchers to answer the last question for the objective of this sudy to give overview of risk assessment process related to big data in Figure 4.1.

Big data categorized as high volumes of data that need to be collected with high speed variable.Big Data characteristics as follows [14] :

(a) Decentralized computing nodes
(b) Extremely robust system
(c) Fault tolerant
(d) Generated real-time data
(e) Parallel data generation
(f) Efficient programming paradigms
(g) Non-relational data storage
(h) Variable data sources

The key information in the big data is simplified as below [14] :-

i. To distribute and retrieve data
ii. To data collect, annotated and clean data
iii. Integrating, describing and aggregating data
iv. To create data
v. To analyze data

Below explained more about risk categories and risk management strategies in big data infrastructure [15]:

A. Risk Categories

There are various of risks in big data as the classification of the risk as below [16] :

1) Insecure Computation
   Retrieve sensitive data leads to data corruption in Big Data as Services.
2) Granular Access Control
   Avoid end user to access data.
3) Insecure Data Storage
   Unauthorize user access to data storage when data is distributed at each nodes.

B. Risk Management Strategies :
ISO 31000 standards applied risk management in big data. In addition, ISO 31000:2018 offers additional direction than ISO 31000:2009 and emphasizing more on top management engagement and risk management in the company [8]. Risk management includes in following steps [14]:

1) Define risk from technical and management aspect
2) Plan to implement risk management in big data
3) Managing risk profile
4) Analyze the risk
5) Estimate likelihood of risk
6) Monitor risk acceptability
7) Evaluate and treat the risk

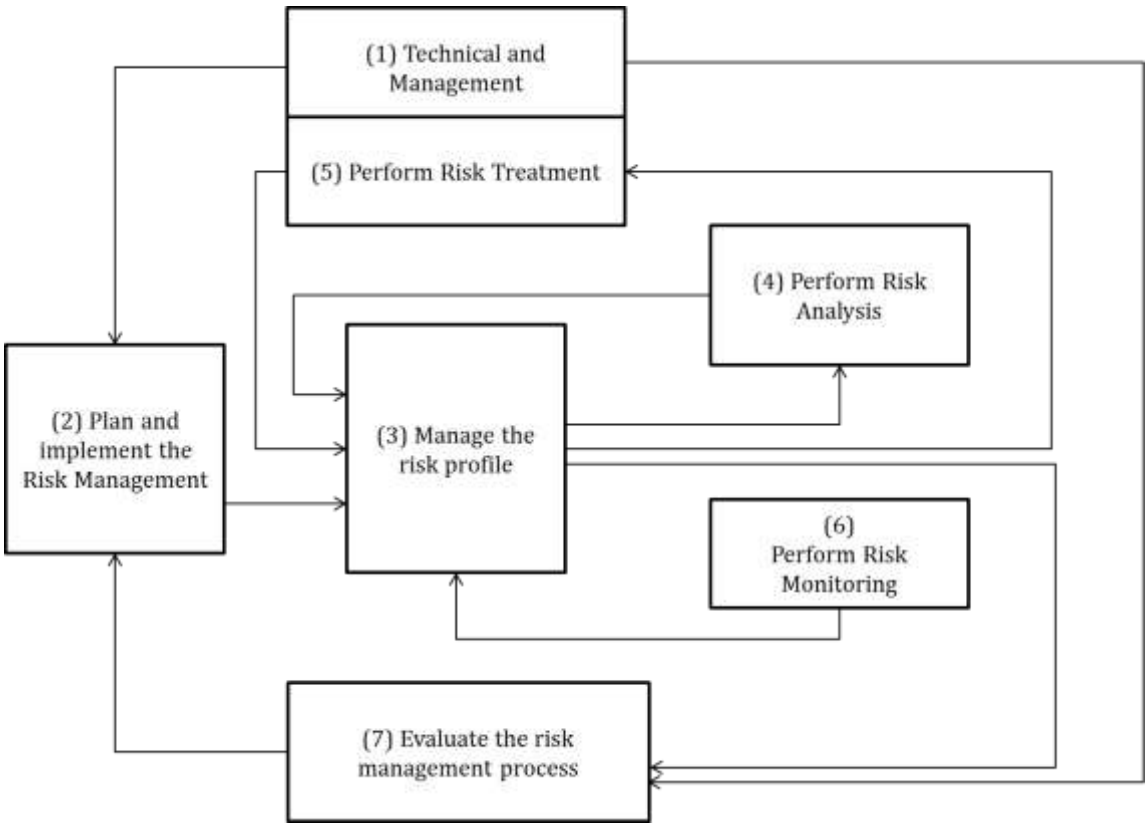

Figure 4.1 Risk Process Framework [14]

5. Conclusion

The purpose of the risk assessment process framework is to give overview on availability of processing framework for risk in the big data that can be applied in the enterprise. It helps all the IT and non-IT recruitments in the enterprise and organization on a deeper understanding of risk assessment in Big Data Security. One of the challenges arises when systems try to handle the Big Data concept. In this paper, it

gives overview on big data security issues, security framework and the risk assessment processing framework to overcome the potential risk in the big data.

References

[1] Fujitsu, "Risk Management Our Approach to Risk Management," pp. 1–4, 2019.

[2] B. A. Berhad, "BUMI ARMADA BERHAD RISK MANAGEMENT COMMITTEE," no. November, 2018.

[3] Christophe Veltsos, "Lessons From the ISO/IEC 27005:2018 Security Risk Management Guidelines," *IBM*, 2018. .

[4] F. Bélanger and L. Carter, "Trust and risk in e-government adoption," *J. Strateg. Inf. Syst.*, vol. 17, no. 2, pp. 165–176, 2008.

[5] J. Lanz, "Conducting Information Technology Risk Assessments," *CPA J.*, vol. 85, no. 5, pp. 6–9, 2015.

[6] T. Ye, "A study of software development project risk management," *Proc. - 2008 Int. Semin. Futur. Inf. Technol. Manag. Eng. FITME 2008*, pp. 309–312, 2008.

[7] P. Willumsen, J. Oehmen, V. Stingl, and J. Geraldi, "Value creation through project risk management," *Int. J. Proj. Manag.*, 2019.

[8] L. Agreement and A. Standards, "AS/NZS ISO 31000:2009 Risk management - Principles and guidelines," 2018.

[9] Technical-Committee, "INTERNATIONAL STANDARD ISO / IEC Information technology — Security techniques — Information security risk management," *Iso/Iec 27005:2018*, vol. 2018–07, no. Third Edition, 2018.

[10] P. Massingham, "Knowledge risk management : a framework," no. June 2010, 2014.

[11] M. Y. Aguria, S. Q. Wang, and M. Dulaimi, "Risk management framework for construction projects in developing countries," no. May 2014, 2004.

[12] NIST, "Guide for Conducting Risk Assessments SP800-30rev1," *NIST Spec. Publ. 800-30 Revis. 1*, no. September, p. 95, 2012.

[13] S. Alhawari, L. Karadsheh, A. Nehari, and E. Mansour, "A Theoretical Framework for Knowledge Management Process : Towards Improving Knowledge Performance International Journal of Information Management Knowledge-Based Risk Management framework for Information Technology project," *Int. J. Inf. Manage.*, vol. 32, no. 1, pp. 50–65, 2009.

[14] V. Malik, "Cloud , Big Data & IoT : Risk Management," *2019 Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput.*, pp. 258–262, 2019.

[15] X. He, M. He, and Z. Han, "A Survey of Network Topology of Data Center," *Proc. - 4th IEEE Int. Conf. Big Data Secur. Cloud, BigDataSecurity 2018, 4th IEEE Int. Conf. High Perform. Smart Comput. HPSC 2018 3rd IEEE Int. Conf. Intell. Data Secur.*, pp. 39–41, 2018.

[16] M. Paryasto, A. Alamsyah, and B. Rahardjo, "Big-Data Security Management Issues," *2014 2nd Int. Conf. Inf. Commun. Technol.*, pp. 59–63, 2014.