ACCESS CONTROL MODELS AND TECHNOLOGIES FOR BIG DATA PROCESSING AND MANAGEMENT

Sairam Madasu

8347 sandstone crest lane, Indian Land, South Carolina 29707

Email: Rammadasu5@gmail.com

Abstract

The growth of data and digital information created a gap in security. Therefore, the study aims toward analysing the role of access control technology in big data processing. Additionally, access control technology is analysed from the point of view of an additional layer of security. Additionally, rationality and problems of the analysis are discussed in the study along with objectives and questions. Past literature related to the topic was analysed in order to determine the variability of factors. Moreover, an analysis of the past literature was conducted in order to achieve an appropriate path for the analysis of the topic. In order to conduct the empirical analysis primary method of data collection is conducted along with quantitative analysis. Additionally, Data collection is done with the help of a questionnaire. The finding of the study has all the quantitative analysis related to the empirical analysis. In the discussion part of the study an overview of the study is presented. Moreover, in the discussion part, a summative presentation of the entire study is done. In order to conclude the findings in a brief manner concussion is constructed, moreover it was found that there is a relationship between Access management and other factors.

Keywords- *Access management, Big data analysis, Big data management, MAC and RBAC, Factors of access control*

Introduction

Growing cyber threats had created a major gap in the process of big data analysis. According to the opinion of Wang et al. (2020), the magnitude of data security and privacy issues are justified by data volume, variety, and data velocity. Therefore, factors like "Mandatory access control (MAC)" and "role-based access control (RBAC)" pays an important role in big data processing and management.
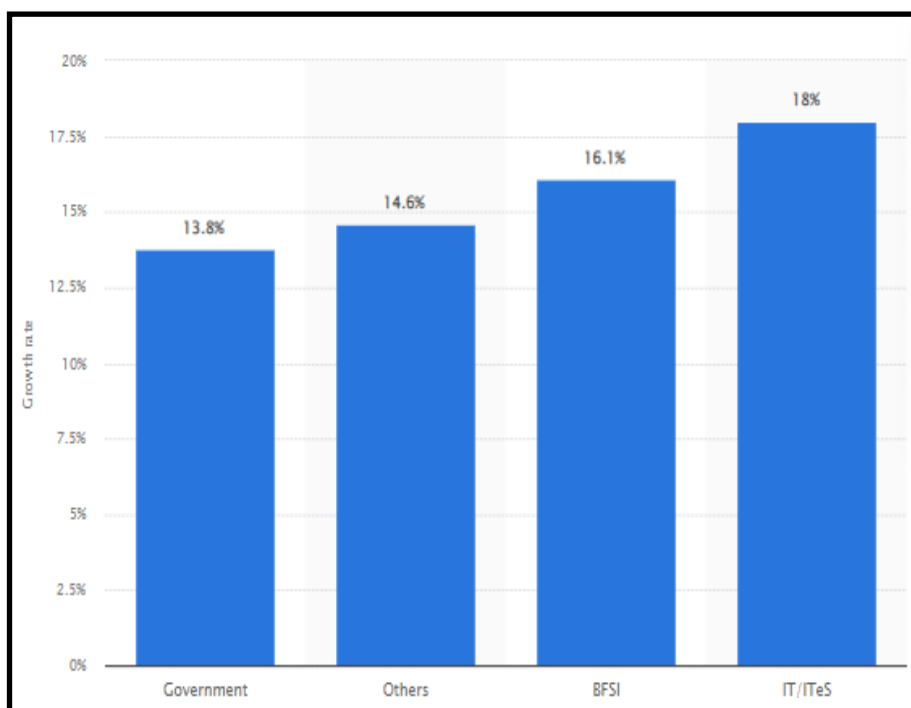
Figure 1: Rate of growth in cyber safety spending in India from 2019 to 2022

(Source: Statista, 2022)

Figure 2 of the empirical Study is related to the expansion rate for cyber safety spending in India from 2019 to 2022 in different sectors. It can be seen that for IT industries the growth rates are the highest at 18% (Statista, 2022). However, the most noticeable fact is that there is a 13.8% growth in the government sector (Statista, 2022). Therefore, it can be said that the demand for cyber security is high in India. Additionally, the high demand justified the rationality of the topic of ate study.

Additionally, there are some problems related to the types of access control which are discussed in the study. As per the suggestions of Qiu et al. (2020), the demand for access control changes according to the demand for big data dispensation and organization systems. Therefore, factors like a dearth of integration with another system and the outdated system is a major issue in access control technology.

Aim

Eur. Chem. Bull. **2023**,*12(Special issue 8), 6886-6902*

6887

The following empirical study primarily aims toward analysing the role of Access control technology for big data processing in order to ensure an additional layer of security.

Research objective

- To analyse the role of access control technology as an extra layer of security big data dispensation and organization

- To analyse the implementation of access control according to the type of big data dispensation and administration system

- To discuss the importance of "Mandatory access control (MAC)" and "Role-based access control (RBAC)" in a big data processing and management system

- To understand the factors that has a direct impact on access control of a big data dispensation and administration system

Research question

- How does access control technology provides an extra layer of security in big data processing and supervision system?

- What are the implications of access control according to the type of big data processing and supervision system?

- What is the importance of MAC RBAC in a big data processing and management system?

- What are the factors that have a direct impact on access control of big data administration system?

Literature review

Importance of Access control for a big data processing and organization system

In order to provide a reliable understanding of access control for big data processing and administration system reliable past literature was observed. It was found that, according to the opinion of Saini et al. (2020), Access control is essential as it limits access to information and information processing systems. Moreover, the access is based on the user and authentication.
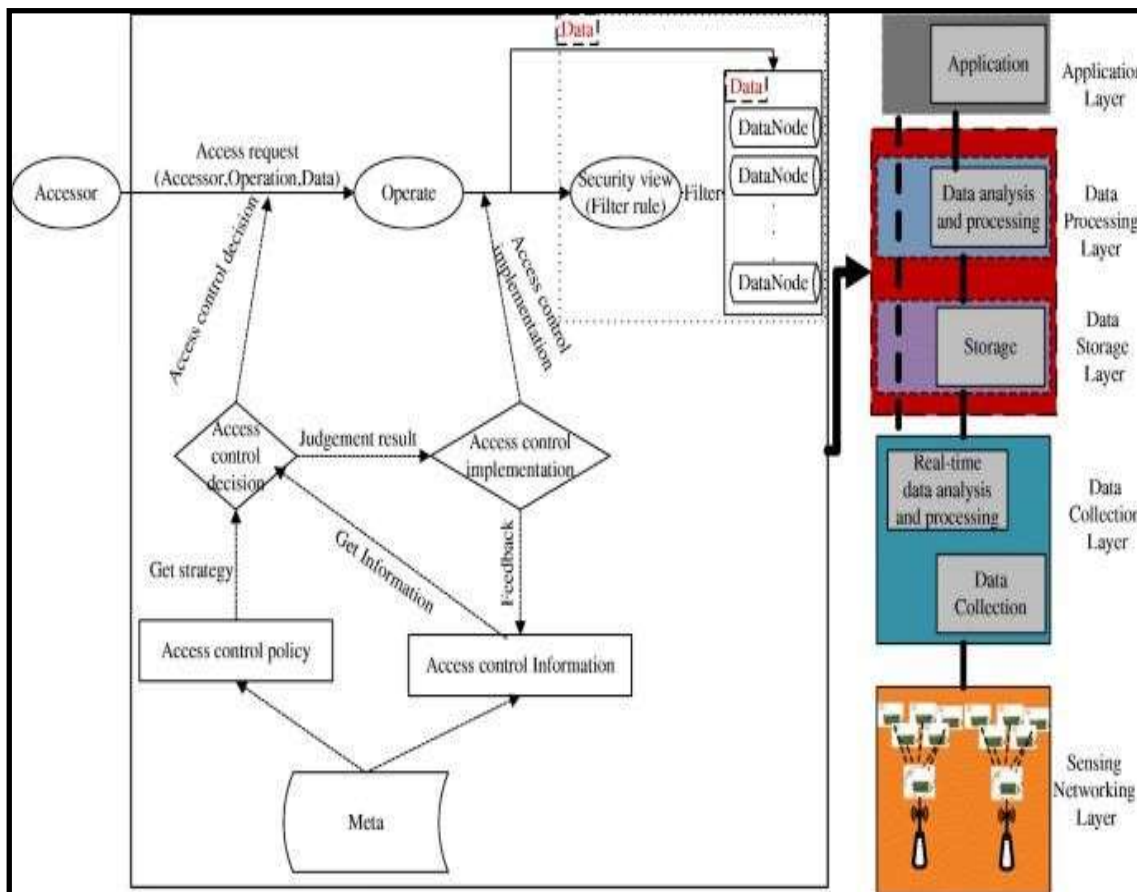
Figure 2: Access control for a big data processing and organization system

(Source: Saini et al. 2020)

Figure 2 of the analysis is related to the access control for a big data processing and organization system. It can be seen that there is an overlapping layer of security in the system; therefore contradicting Saini et al. (2020), Yu, Cai & Wu (2020) stated that the limitation of access is based on the type of system, however, the additional access control can be changed according to the demand. Therefore, the limitation of information is subjective according to the use of data. Hence from the above discussion, it is understood that access controller is an essential factor for data management and the limitation of data depends on the type of access control

Factors Impacting Access Control for data organization

During the analysis, it was estimated that there are different factors which directly impact the process of access control for data management. According to the opinion of Al et al. (2019), purpose and the responsibility of the audience directly Impact access control of a data management system. Moreover, the impact of data management is directly linked with the objective of management.

*Eur. Chem. Bull. **2023**,12(Special issue 8), 6886-6902*

6889
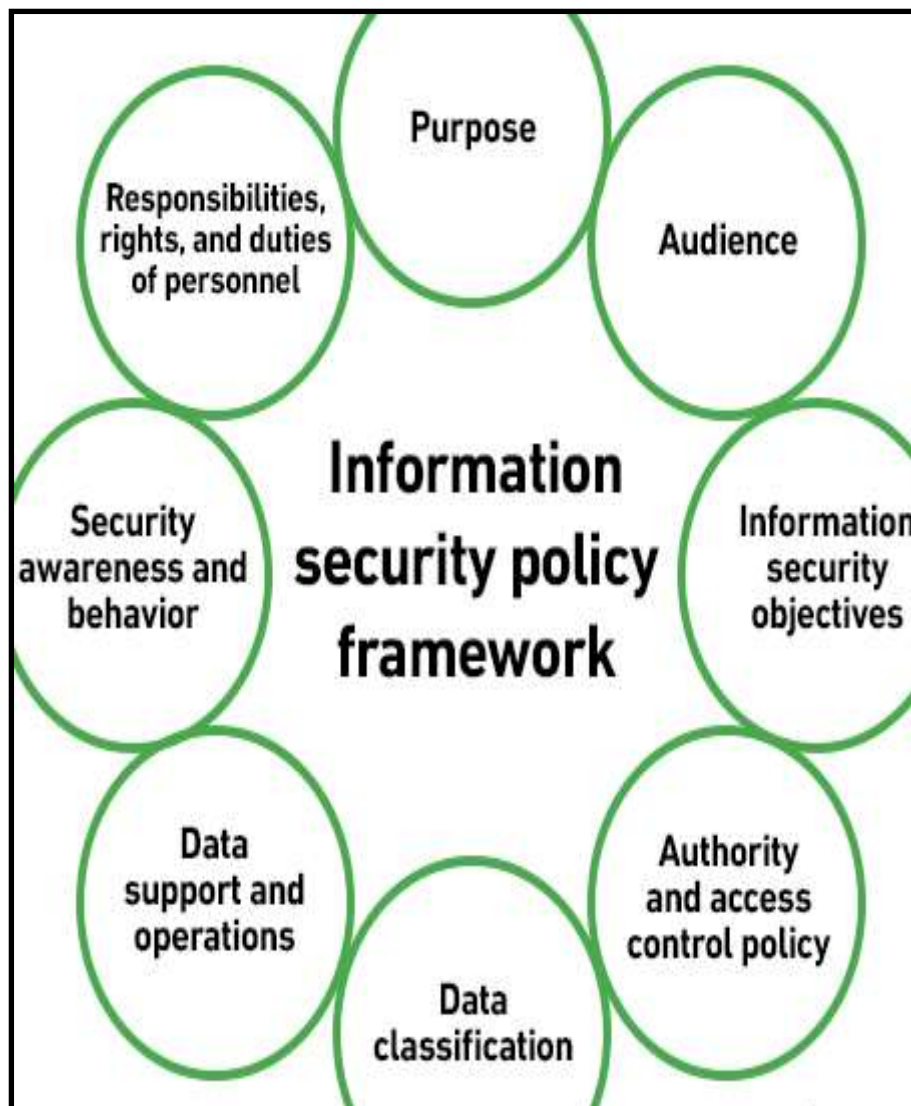
Figure 3: Factors Affecting Access Control

(Source: Samaraweera & Chang, 2019)

Figure 3 of the analysis is related to all of the factors impacting data security and access control. There are various definitive factors that impact the access control of a organization system, Therefore, arguing Al et al. (2019), Samaraweera & Chang (2019) stated that data type, operation type and authorised people are directly related to the access control. Thus it can be said that there are factors that are based on data management that impacts and compromise access control of a big data management system.

Methodology

Data collection

In order to collect data for the study ***primary method*** was considered. As per the suggestions of Johansson (2007), primary sources of data aid the process of collecting reliable and relatable data. In addition, the collected data is collected in real-time. Therefore, the results of the analysis are able to indict recent changes in factors. In order to collect data for the study ***65 participants*** having IT background was collected. The selection process of the participants was done through ***a random sampling method*** (Lather, 2013). For collating data, a questionnaire with ***3 demographic questions*** and ***10 variable-based questions*** was created. In order to protect the integrity of the study all the possible manipulative factors were eliminated. Thus, the participants are solely liable for the answers.

Data analysis

In order to analyse the collated data ***quantitative methods*** of data analysis were considered. As per the recommendations by Pandey & Pandey (2021), the method of quantitative analysis provides a distinctive nature to the study. Moreover, with the quantitative method of analysis, it is possible to establish relations between different data points. Therefore, quantitative methods of analysis were done using ***IBM SPOSS software***. Moreover, the relationship between the different data points is established according to their reliability factors (Mishra & Alok, 2022). In order to analyse the data hypothesis based on the variables is created. Additionally, statistical analysis using ***ANOVA, correlation factors, regression statistics and descriptive statistics*** are considered in the analysis.

Findings

Demographics

Gender
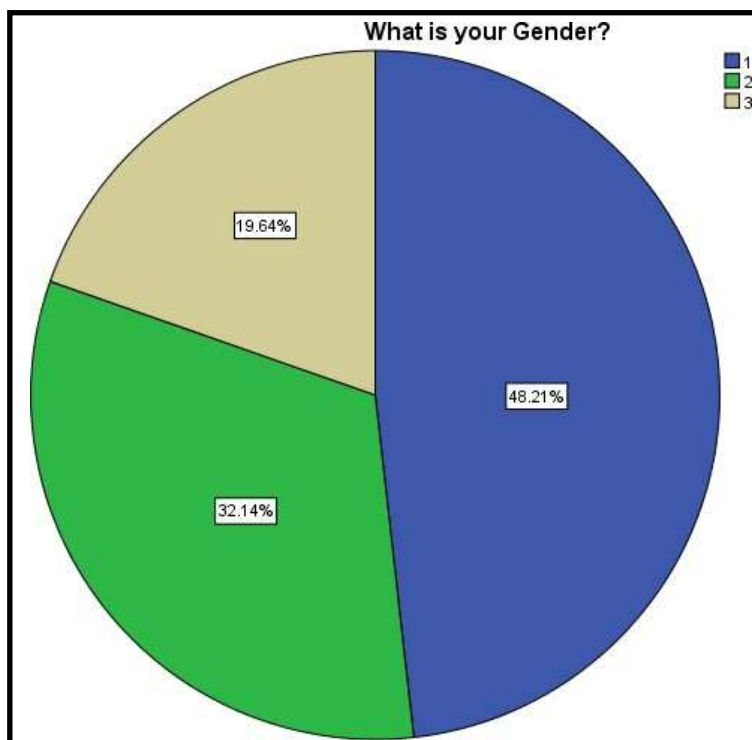
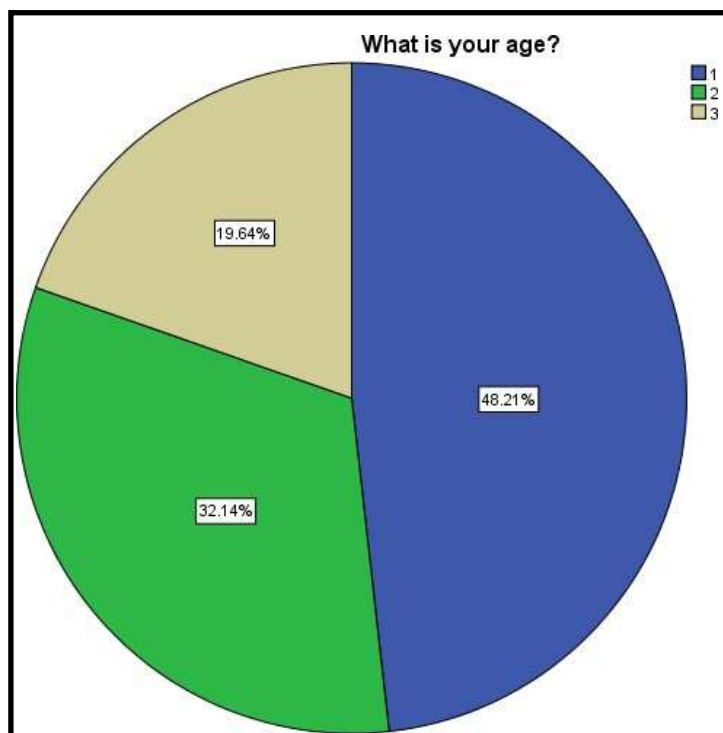**Figure 4: Pie chart associated with the gender of participants**

What is your Gender?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	27	48.2	48.2	48.2
	2	18	32.1	32.1	80.4
	3	11	19.6	19.6	100.0
	Total	56	100.0	100.0	

Table 1: Table associated with the gender of participants

Figure 4 and Table 1 are associated with the gender of the participants. It can be seen that 48.2% of the identified themselves as male and 32.1% of the participants were female. However, there were 19.6% of participants identified with another gender group.

Age

Figure 5: Pie chart related to Age

What is your age?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	27	48.2	48.2	48.2
	2	18	32.1	32.1	80.4
	3	11	19.6	19.6	100.0
	Total	56	100.0	100.0	

Table 2: Table of Age

Table 2 and Figure 5 of the study are related to the age of the participant where 48.2% were Below 20 and 32.1% were between 20 to 35 years of age. In addition, 19.6% of the pie chart represented ages between 35 to 60 years of age. Therefore, it can be said most of the participants were young professionals,

Income
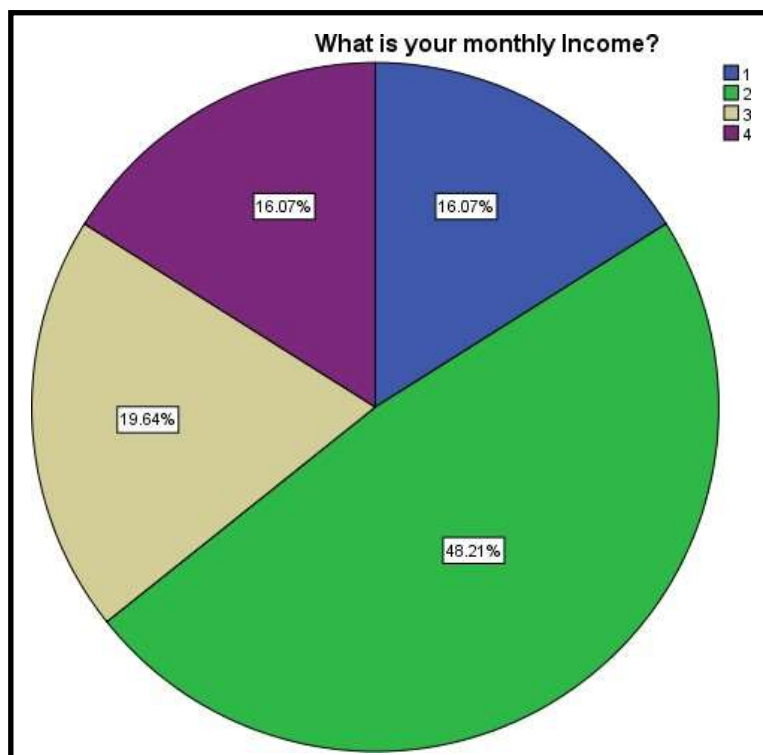
Figure 6: Pie chart related to income

What is your monthly Income?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	9	16.1	16.1	16.1
	2	27	48.2	48.2	64.3
	3	11	19.6	19.6	83.9
	4	9	16.1	16.1	100.0
	Total	56	100.0	100.0	

Table 3: Table of income

Figure 6 and Table 3 of the empirical analysis are related to the income range of the participants. It can be seen that earning below RS 18000 represents 16.1% of the pie chart. However, 48.2% and 19.6% of the pay chart represent participants earning respectively between RS 18000 to 30000 and RS 30000 to 50000. Additionally, there were 16.1% of participants earned above RS 50000. Thus, there was a noticeable diversity in the income range of participants

Descriptive analysis

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
DV	56	1	3	2.04	.602	.362	-.012	.319	-.102	.628
IV1	56	2.00	4.00	2.8393	.68162	.465	.212	.319	-.795	.628
IV2	56	2.00	5.00	4.0000	.99087	.982	-1.047	.319	.237	.628
IV3	56	2.00	3.00	2.3214	.47125	.222	.786	.319	-1.435	.628
IV4	56	4.00	7.00	5.1250	1.07977	1.166	.552	.319	-.955	.628
Valid N (listwise)	56									

Table 4: Table of Descriptive Analysis

Table 4 of the study indicates the descriptive statistics of the variables related to the topic. As per the opinion of Rawat, Doku & Garuba (2019), the analysis of descriptive statistics aid in understanding the fundamental traits of the different variables. Descriptive statistics are also used to describe the variables' probable reliability (Stergiou, Psannis & Gupta, 2020). As can be seen, the dependent variable's mean value is 2.04, while its standard deviation is 0.602. Additionally, the mean value of other self-governing variables is higher than the standard deviation value. Thus, it can be concluded that the majority of the components are centred on the means of each variable and most participants agreed with the statements.

Hypothesis 1: There is a direct relationship between access control and the security of data

Model Summary[b]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.413[a]	.171	.155	.553	.171	11.109	1	54	.002	1.529

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.400	1	3.400	11.109	.002[b]
	Residual	16.528	54	.306		
	Total	19.929	55			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	1.000	.319		3.131	.003			
	IV1	.365	.109	.413	3.333	.002	.413	.413	.413

Table 5: Regression table for Hypothesis 1

Eur. Chem. Bull. 2023,12(Special issue 8), 6886-6902

6895

Table 5 of the study is related to regression data for Hypothesis 1 and the significance value is 0.02. worth value is lower than 0.05 indicates the reliability of hypothesis related with study (Chen, Lv & Song, 2019). Therefore, the significance value of 0.02 indicates that hypothesis 1 supports the argument with appropriate pieces of evidence.

Hypothesis 2: Implementation Access control of data depends on the type of Big Data Processing and Management system

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.549[a]	.301	.288	.508	.301	23.262	1	54	.000	1.564

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.000	1	6.000	23.262	.000[b]
	Residual	13.929	54	.258		
	Total	19.929	55			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	.702	.285		2.467	.017			
	IV2	.333	.069	.549	4.823	.000	.549	.549	.549

Table 6: Regression table for Hypothesis 2

Table 2 of the quantitative analysis is related to the second hypothesis and provides a worth value of 0.000. Therefore, the value required to be compared with the level of worth, which lies, includes 0.1, 0.05, and (Ouaddah, et al, 2017). Comparing with the significance level it can be said that the null hypothesis can be rejected in all of the scenarios.

Hypothesis 3: Authorization of the user base related with is related to MAC and RBAC in a Data Management system

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.041[a]	.002	-.017	.607	.002	.092	1	54	.763	1.764

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.034	1	.034	.092	.763[b]
	Residual	19.895	54	.368		
	Total	19.929	55			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	2.158	.411		5.247	.000			
	IV3	-.053	.174	-.041	-.303	.763	-.041	-.041	-.041

Table 7: Regression table for Hypothesis 3

Table 7 of the study has the regression data for the 3rd hypothesis where the connotation value is 0.763. According to Ravidas et al. (2019), Vallu more other than 0.05 represents no impact on the dependent variable. Therefore, a value of 0.763 indicates that the 3rd hypothesis is not supported with appropriate evidence.

Hypothesis 4: There are factors of Access control that directly impact Access control of a system

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.566[a]	.321	.308	.501	.321	25.515	1	54	.000	1.908

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.395	1	6.395	25.515	.000[b]
	Residual	13.534	54	.251		
	Total	19.929	55			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	3.654	.327		11.164	.000			
	IV4	-.316	.063	-.566	-5.051	.000	-.566	-.566	-.566

Table 8: Regression table for Hypothesis 4

Table 8 of the analysis is related to the regression statistics of hypothesis 4 of the study. It can be seen that the importance value for the 4th hypothesis is 0.000 which is lower

than 0.05. Therefore, a lower significance value indicates the reliability of a hypothesis (Han et al. 2023). Therefore, it can be said that 4th hypothesis is supported with sufficient evidence.

Discussion

The empirical study is related to access control for programs and models for big data analysis. In order to conduct a knowledge-based analysis primary sources data were considered and quantitative analysis was conducted. According to the opinion of Samaraweera & Chang (2019), Access control of a data management service depends on the types of data storage and authentication of a user. In addition, there are different factors that impact the effectiveness of the access console program (Sollins, 2019). Therefore, a conclusive analysis is done regarding the factors. In addition, an understanding of various factors is drawn based on the quantitative analysis of the study. Furthermore, for a better understanding of the background if the study is drawn in the introduction part followed by the literature review where past literature analysis is conducted. Similarly, a survey is done for establishing the relationship between the factors.

Conclusion

Thus, an empirical analysis is conducted in order to understand the role of Access control technology for big data processing additionally; access control is discussed based on ensuring an additional layer of security for data management. Therefore, the factors of the study are discussed accordingly. It was found that most of the participants agreed with the statements of the questionnaire. On the other hand, a relationship between data security and access control is established with quantitative data. In addition, it was found that the implication and creation of access control programs must be based on the demand of the big data management system. Additionally, the findings are based on the findings of probes of access control therefore, a complete analysis is presented.

References

Al Breiki, H., Al Qassem, L., Salah, K., Rehman, M. H. U., & Sevtinovic, D. (2019, November). Decentralized access control for IoT data using blockchain and trusted oracles. In 2019 IEEE International Conference on Industrial Internet (ICII) (pp. 248-257). IEEE.Retrieved on 27th June 2023, from: https://www.researchgate.net/profile/Muhammad-Habib-Ur-Rehman/publication/335258940_Decentralized_Access_Control_for_IoT_Data_Using_Blockchain_and_Trusted_Oracles/links/5dc84890458515143500699a/Decentralized-Access-Control-for-IoT-Data-Using-Blockchain-and-Trusted-Oracles.pdf

Chen, J., Lv, Z., & Song, H. (2019). Design of personnel big data management system based on blockchain. Future generation computer systems, 101, 1122-1129.Retrieved on 27th June 2023, from: https://journal.ump.edu.my/ijsecs/article/download/4397/821

Han, Z., Wang, K., Zhou, H. C., Cheng, P., & Shi, W. (2023). Preparation and quantitative analysis of multicenter luminescence materials for sensing function. Nature Protocols, 18(5), 1621-1640.Retrieved on 27th June 2023, from https://www.nature.com/articles/s41596-023-00810-1

Johansson, R. (2007). On case study methodology. Open house international, 32(3), 48-54.Retrieved on 27th June 2023, from: http://d30037385.purehost.com/HTMLobj-3839/Case_Study_Methodology-_Rolf_Johansson_ver_2.pdf

Lather, P. (2013). Methodology-21: What do we do in the afterward?. International journal of qualitative studies in education, 26(6), 634-645.Retrieved on 27th June 2023, from: https://www.tandfonline.com/doi/abs/10.1080/09518398.2013.788753

Mishra, S. B., & Alok, S. (2022). Handbook of research methodology.Retrieved on 27th June 2023, from: http://74.208.36.141:8080/jspui/bitstream/123456789/1319/1/BookResearchMethodology.pdf

Ouaddah, A., Mousannif, H., Abou Elkalam, A., & Ouahman, A. A. (2017). Access control in the Internet of Things: Big challenges and new opportunities. Computer Networks, 112, 237-262.Retrieved on 27th June 2023, from: https://www.researchgate.net/profile/Aafaf-Ouaddah/publication/309893261_Access_control_in_The_Internet_of_Things_Big_challenges_and_new_opportunities/links/5a5f21c7458515c03ee1c7af/Access-control-in-The-Internet-of-Things-Big-challenges-and-new-opportunities.pdf

Pandey, P., & Pandey, M. M. (2021). Research methodology tools and techniques. Bridge Center.Retrieved on 27th June 2023, from: http://dspace.vnbrims.org:13000/jspui/bitstream/123456789/4666/1/RESEARCH%20 METHODOLOGY%20TOOLS%20AND%20TECHNIQUES.pdf

Qiu, J., Tian, Z., Du, C., Zuo, Q., Su, S., & Fang, B. (2020). A survey on access control in the age of internet of things. IEEE Internet of Things Journal, 7(6), 4682-4696.Retrieved on 27th June 2023, from: https://ieeexplore.ieee.org/abstract/document/8968396/

Ravidas, S., Lekidis, A., Paci, F., & Zannone, N. (2019). Access control in Internet-of-Things: A survey. Journal of Network and Computer Applications, 144, 79-101.Retrieved on 27th June 2023, from: https://www.academia.edu/download/82730094/Fardapaper-Access-control-in-Internet-of-Things-A-survey.pdf

Rawat, D. B., Doku, R., & Garuba, M. (2019). Cybersecurity in big data era: From securing big data to data-driven security. IEEE Transactions on Services Computing, 14(6), 2055-2072.Retrieved on 27th June 2023, from: Rawat, D. B., Doku, R., & Garuba, M. (2019). Cybersecurity in big data era: From securing big data to data-driven security. IEEE Transactions on Services Computing, 14(6), 2055-2072.

Saini, A., Zhu, Q., Singh, N., Xiang, Y., Gao, L., & Zhang, Y. (2020). A smart-contract-based access control framework for cloud smart healthcare system. IEEE Internet of Things Journal, 8(7), 5914-5925.Retrieved on 27th June 2023, from: https://www.researchgate.net/profile/Akanksha-Saini-3/publication/346312690_A_Smart_Contract_Based_Access_Control_Framework_fo r_Cloud_Smart_Healthcare_System/links/60dea85b92851ca9449f0130/A-Smart-Contract-Based-Access-Control-Framework-for-Cloud-Smart-Healthcare-System.pdf

Samaraweera, G. D., & Chang, J. M. (2019). Security and privacy implications on database systems in Big Data era: A survey. IEEE Transactions on Knowledge and Data Engineering, 33(1), 239-258.Retrieved on 27th June 2023, from: https://www.computer.org/csdl/api/v1/periodical/trans/tk/5555/01/08765784/1bLyn2 RvkbK/download-article/pdf?casa_token=dmrFbtH-SHEAAAAA:2IIvBDe4WdsbZ4dGSNRXUpCfhjS7BYZUehs7eJxKndoAmUp8Wrn EFVdMH3WK941C6pPDuKRG

Eur. Chem. Bull. **2023**,*12(Special issue 8), 6886-6902*

6900

Sollins, K. R. (2019). IoT big data security and privacy versus innovation. IEEE Internet of Things Journal, 6(2), 1628-1635.Retrieved on 27th June 2023, from: https://ieeexplore.ieee.org/ielaam/6488907/8709863/8643026-aam.pdf

Statista, 2022, the *Growth rate of cyber security expenditure in India between 2019 and 2022, by sector*Retrieved on 27th June 2023, from: https://www.statista.com/statistics/1099920/india-cyber-security-expenditure-growth-rate-by-sector/

Stergiou, C. L., Psannis, K. E., & Gupta, B. B. (2020). IoT-based big data secure management in the fog over a 6G wireless network. IEEE Internet of Things Journal, 8(7), 5164-5171.Retrieved on 27th June 2023, from: https://ruomo.lib.uom.gr/bitstream/7000/808/1/IoT-based%20Big%20Data%20secure%20management%20in%20the%20Fog%20over%20a%206G%20Wireless%20Network.pdf

Wang, J., Yang, Y., Wang, T., Sherratt, R. S., & Zhang, J. (2020). Big data service architecture: a survey. Journal of Internet Technology, 21(2), 393-405.Retrieved on 27th June 2023, from: https://jit.ndhu.edu.tw/article/viewFile/2261/2274

Yu, B., Cai, Y., & Wu, D. (2020). Joint access control and resource allocation for short-packet-based mMTC in status update systems. IEEE Journal on Selected Areas in Communications, 39(3), 851-865.Retrieved on 27th June 2023, from: Yu, B., Cai, Y., & Wu, D. (2020). Joint access control and resource allocation for short-packet-based mMTC in status update systems. IEEE Journal on Selected Areas in Communications, 39(3), 851-865.

Appendices

Appendix 1: Survey questionnaire

Survey link:

https://docs.google.com/forms/d/114mLEceQ8pzM_P02ouKELavLEcezwSo0Xt31jpdV4mE/
edit

1. What is your Gender?

2. What is your age?

3. What is your monthly Income?

4. Access control for Big Data Processing and Management ensures an additional layer of security

5. Implementing the Access control model is related to the type of Big Data Processing and Management system

6. Implementing the Access control model is indirectly related to a personalised experience

7. Mandatory access control (MAC) is essential for the authorization of the user

8. Mandatory access control provides security to the sensitivity of the information

9. Role-based access control (RBAC) is suitable for multiple uses in Big Data Processing and Management systems

10. Role-based access control (RBAC) restricts access according to competency and responsibility

11. Access control for data management and accountability of users

12. Access control for data management is essential for blocking unauthorised access

13. Access control is related to Domain-specific Big Data approaches