International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Peer Reviewed & Referred Journal |

| Volume 6, Issue 5, September 2019 |

Survey on Security Challenges and Major Advantages of Big Data

Dr. Archana Verma

Assistant Professor, Computer Science & Engineering , Bipin Tripathi Kumaon Institute of Technology, Dwarahat, Uttarakhand, India

ABSTRACT: The Big Data field applies to manage data sets whose size is too large for commonly used software tools to capture, man- age, and analyse that amount of data effectively. Data volume is expected to double every two years. Data from all these sources are very often unstructured, and come from a wide range of sources, including social media, sensors, scientific applications, surveillance, video and image archives, Internet search indexing, medical records, business transactions, and system logs. Currently, big data is gaining more and more traction as devices connected to the Internet of Things (IoT) are increasing rapidly, producing large amounts of data that must be transformed into valuable information.

KEYWORDS: security, big data, advantages, challenges, survey, social networking, analysis, privacy

I.INTRODUCTION

Data generation and collection quickly surpass the bounds in the digital universe of today. The data has been doubling every 2 years since 2011 . It is predicted that the data will increase 300 times, from 130 exabytes in 2005 to 40,000 exabytes in 2020 [1]. As a result of this technological revolution, the big data is becoming increasingly an important issue in the sciences, governments, and enterprises. Big Data is a data set, which is difficult to capture, store, filter, share, analyze and visualize on it with current technologies .Despite such difficulties[2], if you can cope with big data, it provides you with generating revenue, executive efficiency, strategic decisions, better services, defining needs, identifying new trends, and developing new products, all of which is covered in the data science .In addition, data science studies parallel and distributed processing, similarity search, graph analysis, clustering, stream processing, search ranking, association analysis, dimensionality reduction and machine learning algorithms . However, in this complex computation environment, traditional security and privacy mechanisms are insufficient to analyze big data.[3] These challenges in big data consist of computation in distributed and non-relational environments, cryptography algorithms, data provenance, validation and filtering, secure data storage, granular access control, and real time monitoring .[4]

In 2012, 2.5 quintillion bytes of data were generated daily, and 90% of current data worldwide originated in the past two years . During 2012, 2.2 million TB of new data are generated each day. In 2010, the market for Big Data was $3.2 billion, and this value is expected to increase to $16.9 billion in 2015 . As of July 9, 2012, the amount of digital data in the world was 2.7 ZB Facebook alone stores, accesses, and analyzes 30 + PB of user-generated data . In 2008, Google was processing 20,000 TB of data daily[5] . To enhance advertising, Akamai processes and analyzes 75 million events per day . Walmart processes over 1 million customer transactions, thus generating data in excess of 2.5 PB as an estimate.[6]

More than 5 billion people worldwide call, text, tweet, and browse on mobile devices [46]. The amount of e-mail accounts created worldwide is expected to increase from 3.3 billion in 2012 to over 4.3 billion by late 2016 at an average annual rate of 6% over the next four years. In 2012, a total of 89 billion e-mails were sent and received daily, and this value is expected to increase at an average annual rate of 13% over the next four years to exceed 143 billion by the end of 2016 . In 2012, 730 million users (34% of all e-mail users) were e-mailing through mobile devices. Boston.com reported that in 2013, approximately 507 billion e-mails were sent daily. Currently, an e-mail is sent every 3.5×10^{-7} seconds. Thus, the volume of data increases per second as a result of rapid data generation.[7]

Growth rates can be observed based on the daily increase in data. Until the early 1990s, annual growth rate was constant at roughly 40%. After this period, however, the increase was sharp and peaked at 88% in 1998 . Technological progress has since slowed down. In late 2011, 1.8 ZB of data were created as of that year, according to IDC . In 2012, this value increased to 2.8 ZB. Globally, approximately 1.2 ZB (10^{21}) of electronic data are generated per year by various sources. By 2020, enterprise data is expected to total 40 ZB, as per IDC.Based on this estimation, business-to-consumer (B2C) and internet-business-to-business (B2B) transactions will amount to 450 billion per day. Thus, efficient management tools and techniques are required.[8]

Big Data is large, diversified sets of data sourcing out of multiple channels: social media platforms, websites, electronic check-ins, sensors, product purchase, call logs, the choices are limitless. Big Data has three unique characteristics: volume, velocity, and variety.[9]

- Volume: Big Data contains an undecided and unfiltered volume of information. The data collected is different for different businesses. Therefore, the efforts paid are unique. Nonetheless, filtering valuable data from the voluminous pile is essential. Companies need to process this high-volume information to address their business challenges.
- Velocity: It is the speed at which data is created and collected. Mobile, SaaS solutions, e-commerce transactions, and IoT devices are a few of the primary sources of acquiring real-time data. The velocity at which data is generated at scale requires real-time handling and processing for augmenting Data Analytics.[10]
- Variety: Conventional data types consist of structured data that fit well with relational databases. However, with Semi-structured and Unstructured data in the landscape, the information received requires additional preprocessing to convert it into digestible formats. While Structured data can be quickly dealt with, Semi-structured and Unstructured data need to be converted into predetermined models or formats before turning them into actionable information.[11]

Processing Big Data has become the go-to technique to collect information that can further be used to enhance business operations. However, the process is not straightforward. Considering its diversified nature and content, traditional relational databases are incapable of capturing, managing, or processing Big Data into digestible formats.[12]

II.DISCUSSION

Data Analysts harness different data types primarily to make better and improved business decisions by understanding customer behavior and their purchasing patterns. Data Mining, Machine Learning, and Predictive Analytics are a few of the newly-evolved techniques used to achieve new insights into untapped data source areas for optimizing business processes. Let’s discuss the main benefits that businesses can reap from Big Data:

- Big Data allows companies to improve their products and create tailored marketing by gaining a 360-degree view of their customers’ behavior and motivations.[13]
- It enables businesses or service providers to monitor fraudulent activities in real-time by identifying unusual patterns and behavior with the help of Predictive Analytics.
- It drives supply chain efficiencies by collecting and analyzing data to determine if products are reaching their destination in the desired conditions to attract customers’ interest.
- Predictive analysis allows businesses to scan and analyze social media feeds to understand the sentiment among customers.
- Companies that collect a large amount of data have a better chance to explore the untapped area alongside conducting a more profound and richer analysis to benefit all stakeholders.[14]
- The faster and better a business understands its customer, the greater benefits it reaps. Big Data is used to train Machine Learning models to identify patterns and make informed decisions with minimal or no human intervention

Businesses are adopting Cloud Data Storage to move their data easily to expedite business operations. However, the risks involved are exponential with security issues. Even the slightest mistake in controlling the access of data can allow anyone to get a host of sensitive data. As a result, big tech companies embrace both on-premise and Cloud Data Storage to obtain security as well as flexibility.[15]

While mission-critical information can be stored in on-premise databases, less sensitive data is kept in the cloud for ease of use. However, to implement security policies in on-premise databases, companies require cybersecurity experts. Although it increases the cost of managing data in on-premise databases, companies must not take security risks for granted by storing every data in the cloud.Fake Data generation poses a severe threat to businesses as it consumes time that otherwise could be spent to identify or solve other pressing issues. There is more scope for leveraging inaccurate information on a very large scale, as assessing individual data points can be a daunting task for companies.[16]

False flags for fake Data can also drive unnecessary actions that can potentially lower production or other critical processes required for running businesses. One way to avoid this is to ensure that companies should be critical of the data they are working on for enhancing business processes. An ideal approach is to validate the data sources by periodic assessments and evaluate Machine Learning models with diverse test datasets to find anomalies.Data Privacy is a big challenge in this digital world. It aims to safeguard personal or sensitive information from cyberattacks, breaches, and intentional or unintentional data loss. Businesses must follow stricter Data Privacy principles with the help of access management services in the cloud, including very rigid privacy compliance, to strengthen Data Protection. It is best to follow a few rules alongside implementing one or more Data Security technologies. The general rules are knowing your data, having more grip over your data stores and backup, safeguarding your network against unauthorized access, conducting regular risk assessments, and training the users regularly about Data Privacy and Data Security.[17]
A security breach can have crushing consequences on businesses, including the vulnerability of critical business information to a completely compromised database. Deploying highly secured databases is vital to ensure data security at all levels. A superior Database Management System comes with various access controls. While it is advisable to follow rigid and rigorous physical security practices, it is even more essential to follow extensive software-based security measures to safeguard data storage. A few methods to effectively achieve this goal are—practicing data encryption, data segmenting and partitioning, securing on-the-move, and implementing a trusted server. Besides, a few security tools can integrate with databases to automatically monitor data sharing and notify businesses when data has been compromised. Controlling which data users can view or edit enables companies to ensure not only data integrity but also preserves its privacy. But managing access control is not straightforward, especially in larger companies that have thousands of employees. However, a shift from on-premise solutions to cloud-based services has simplified the process of working with Identity Access Management (IAM). IAM does the job of controlling data flow via identification, authentication, and authorization. Following relevant ISO standards is a good starting place to ensure organizations meet the best IAM practices. Today, there are several Machine Learning solutions like chatbots that are trained on a colossal amount of data. The advantages of such solutions are that they keep on improving as users interact. However, this leads to Data Poisoning, a technique to attack Machine Learning models' training data. It can be considered as an integrity attack as the tampered training data can affect the model's ability to provide correct predictions. The results can be catastrophic, ranging from logic corruption to Data Manipulation and Data Injection. The best way to beat the evasion is through outlier detection, wherein the injected elements in the training pool can get separated from the existing data distribution[18]

Advance data culture has allowed every employee to hold a certain level of critical business information. While it boosts data democratization, the risk of an employee leaking sensitive information, intentionally or unintentionally, is high. Employee Theft is prevalent not only in big tech companies but also in startups. To avoid Employee Theft, companies have to implement legal policies along with securing the network with a virtual private network. In addition, companies can use a Desktop as a Service (DaaS) to eliminate the functionalities of data stored in local drives. Based on the enlisted concerns, it is apparent as to why enterprises are seeing Big Data Security as a major concern. However, the good news is that with the right information, resources, skilled manpower, detailed coping strategy, and commitment towards data integrity and privacy, many of such challenges can be easily addressed. The absence of threats to Big Data will lead businesses to achieve their ultimate goal of harnessing data for better customer experience and enhanced customer retention.[19]

Extracting complex data from a diverse set of data sources can be a challenging task and this is where Hevo saves the day! Hevo offers a faster way to move data from Databases or SaaS applications into your Data Warehouse to be visualized in a BI tool.[20]

III.RESULTS

Historically, before the term big data was introduced to represent the vast amounts of data in the digitized world of today, the three commonly used dimensions characterising massive databases and data warehouses were the 3Vs, namely, Volume, Velocity, Variety However, with changes in technologies, big data has been characterised by additional dimensions that are more semantically applicable to big data per se. One of the foremost additional dimensions recognised is Veracity, which refers to the credibility or quality of data. Subsequently, Validity, Volatility and Value have formed the most relevant dimensions of big data . Further, certain dimensions such as Variability and Visualisation of big data have evolved, and more and more technical challenges are being identified in big data In particular, Valence and Vulnerability dimensions have been associated with the privacy and security challenges characterising big data of today. Due to the growing dimensions of big data beyond the traditional 3Vs, in this section, we describe the properties and characteristics of all the eleven dimensions to prepare ourselves for the security challenges surrounding the big data initiatives of the future.[21]

Figure 1 shows a summary of the 11Vs as dimensions identified in a typical big data system. These 11Vs are described below.
Dimension 1: Volume

The first known property of big data is Volume, which is attributed to the sheer size of data being collected With the developments of business websites getting integrated with social engineering and mobile applications (Apps), majority of today's data have been collected only recently, within the past couple of years There is an exponential growth in data generation and storage per day. Large multimedia files amounting to at least 300 hours of videos are being uploaded to YouTube and other social sites every minute Several social media sharing services are being leveraged by businesses in addition to their traditional transactional data, resulting in several trillions of data storage. The data volume of a single data source could be growing from petabytes to exabytes and zettabytes. A recent report by International Data Corporation (IDC) predicts that data would grow worldwide by 61% reaching 175 zettabytes in 2025. The volume of big data impacts the security and privacy in at least two major aspects as listed below:

i) data is stored in multiple locations (servers, nodes, clusters, etc.) in a distributed manner where conventional database systems and software tools are unable to continuously monitor and enforce standardised security protocols;
ii) any failure of a cluster or node can affect data transactions and performance within the tolerance time limits and is prone to security vulnerabilities.[22]

Dimension 2: Velocity
The second dimension, Velocity deals with the speed at which new data gets generated and flows into organisations, and the increasing pace at which it needs to be processed in real-time. The impact is on big data analytics where the rate of data creation needs to be matched with the real-time processing speed and capability of computing systems. Even though the volume of data storage could be enhanced, it is more important to consider the velocity at which new data is generated. Even if data is available, unless it can be processed in real-time, business opportunities could be lost. For instance, if weather predictions get delayed due to a slower processing speed that could not match with the velocity of data received, then it affects the right decisions required to be made at the right time. The velocity of big data impacts the security and privacy as faster cryptographic algorithms are required to keep up with the pace for real-time transaction processing. In addition, security audits are required to keep track of historical data by passing through privacy policies that match with the high rate of data accumulation.[23]

Dimension 3: Variety
Big data exhibits heterogeneity with three types of forms such as structured, semi-structured, unstructured, which can be associated with its Variety dimension¨ Majority of the data are unstructured that include files representing audio, video image, and sensor signals, as well as logs coming from social media, satellites, networks, and other machines. The Variety dimension does not refer to such different data representations alone, but also refers to the means and modes in which the same information is conveyed. While most common variety indicates the structural variety of data representation, it is also important to identify media variety or the different medium in which the same data is represented, and semantic variety indicating different meanings based on the different contexts of data. With structured data, a standard query using structured query language (SQL) could be employed to convey the associated semantic meaning, while unstructured data does not involve latent meaning. Recent adoption of email, XML, and other mark-up languages have led to a variety of semi-structured data. The variety of big data impacts the security and privacy with a need to have appropriate data classification and access controls for different data sources, types and formats.[24]

Dimension 4: Veracity
Apart from the 3Vs characterising big data, with more and more uncertainty pertaining to data streaming and data availability, the credibility or quality of data has been considered to describe Veracity as the fourth important dimension of

big data. For big data to be operational with a meaningful analysis, it is important to have the right and accurate data that can be processed in the right amount and at the right time Any data that is redundant incomplete or having errors cannot lead to good results when used in data analysis. When the first 3Vs increase with big data, the veracity reduces leading to less confidence or trust in the data. By improving the veracity of big data, the business risks associated with decision-making could be controlled. This has an impact on the security and privacy policies with respect to enforcing high quality data through appropriate application of data ownership and periodic access review methods.[25]

Dimension 5: Validity

Another dimension related to Veracity is identified as the fifth dimension called Validity, which refers to the applicability of data with a specific context or intended use of the data. Hence, Validity establishes the correctness of the data for a specific use or view of the data in order to reap the benefits of big data analytics in contextual situations Since many organisations spend much time in cleaning the data before any data analysis can be performed, good data governance practices are required to maintain its validity as a continuous quality check process. This requires proper management of third-party vendors and partners enforcing protection of the entire data supply chain.[26]

Dimension 6: Volatility

While Veracity and Validity dimensions characterise the quality assurance of big data, another dimension related to temporal aspects of the data is called Volatility, which determines how long the data is valid for it to be maintained in the data storage. This dimension ensures the currency of the datasets relevant for conducting real-time analysis ’ Due to the cloud storage limitations and expenses associated with maintaining big data, robust policies for backup, and archiving are required to determine how long the data is to be held valid. In order to improve the performance of big data analytics, historic and irrelevant data should be archived regularly. The volatility of big data impacts the security and privacy policies and procedures for data retention, destruction and periodic re-assessment of security solutions.[27]

Dimension 7: Value

The seventh dimension of big data is identified for understanding the benefits of big data associated with different stakeholders of an organisation that add Value to their business. This Value dimension refers to various factors that answers questions such as: which business decisions could leverage on big data insights, when is it most appropriate to make decisions, and who benefits from it directly [34]. In a nutshell, the Value dimension refers to measuring the usefulness of big data in making decisions for improving business performance. It helps organisations to embark on the right big data strategy so that big data analytics could help them to gain more data insights that are required for solving their complex business problems. Since analytics lead to action in businesses, the value of big data is an important dimension and appropriate access controls and approvals over analytical assessments are required. Also, determining appropriate security checkpoints during the development of such data insights are essential.[28]

Dimension 8: Variability

All the above seven Vs could be affected by the eighth dimension of big data, namely Variability, which refers to inconsistencies in which variable data sources could load data into the data storage in variable speeds, formats or types ’It can also refer to outliers or anomaly detection that can benefit the organisation. When such information about the variability of big data is captured and associated with the data in the data storage, it can be utilised to make meaningful insights from big data. The IT security operations should cater to the big data variability aspects within the various audit log collections and monitoring methods.[29]

Dimension 9: Visualisation

With different ways of data representation such as dashboards, heat maps, cone trees, and k-means clustering, an important dimension of big data that has evolved more recently is its Visualisation for improving data insights ’ Visualisation makes data easy to understand, and simple charts and graphs were traditionally used to communicate the data in a commonly acceptable graphical representation. Today, many sophisticated visualisation tools are being integrated with models for data analytics to make more meaningful graphical interpretations of big data in order to facilitate effective decision making. Hence, we consider Visualisation to be a popular requirement forming the ninth dimension of big data. Privacy and protection policies pertaining to the outputs from various visualisation tools should be established in addition to assigning access controls and privileges based on user roles and responsibilities.[30]

Dimension 10: Valence

While big data could be massive, if the connections between the data items are not established, we would have pockets or islands of disparate data whose interrelationships may not be fully understood nor utilised. Any direct connections could be established when data gets collected as they get streamed. However, discovering indirect connections between data items is more difficult and they add value to the organisation. These interconnections, similar to the bonding between atoms in a molecule, result in the tenth dimension, namely Valence of big data ’It is a measure indicating how dense the data is, and a measure of Valence is determined as the ratio between the actually connected data items and the number of connections that could possibly be established within the data collection. Due to the large heterogeneous access points to data, scalability of hardware, network and systems is essential to maintain the appropriate valence that supports the service level agreement (SLA) of the big data system. Security and privacy management procedures should maintain the level of performance for both current and future growth of big data systems.[26]

Dimension 11: Vulnerability

The last but the most important dimension is the Vulnerability property of big data, which relates to the security, privacy, and technology risks arising due to various rich personalised data collected through products and services using Internet applications, social networks, and IoT devices. The vulnerabilities due to security and privacy gaps and the absence of standards are associated to big data technologies, processes and management’ Recently reported breaches of security and privacy of big data has drawn much attention to look into the Vulnerability dimension of big data Security and privacy policies and procedures for incident management with regard to big data are driven by the abovementioned dimensions that require continuous monitoring. Periodic vulnerability checks and penetration tests are to be developed catering to the unique features of big data. The vulnerabilities of sensitive data leakage must be identified and appropriate measures to review the confidentiality, integrity and availability of big data systems and data are required.[28]

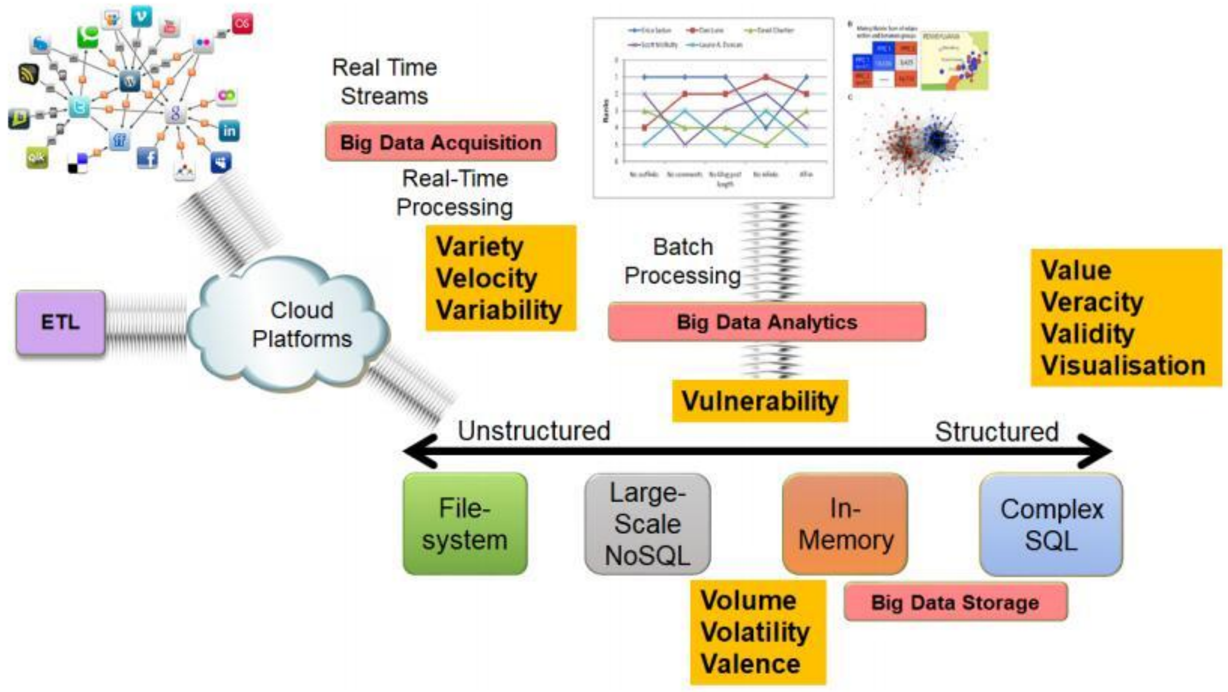

Figure 1. Dimensions of a typical big data system.

CONCLUSIONS

The dynamic integration of different technologies via Intranets, cloud infrastructures, the Internet, social, and IoT networks has resulted in highly complex and heterogeneous big data systems. This paper attempted to satisfy the need to underpin the evolving new privacy and security issues using a holistic approach of applying first principles thinking for understanding the entire big data system. Firstly, we established the premise of big data environment in order to understand its complexities for addressing these challenges. From the evolution of various dimensions of big data, we identified eleven key dimensions or 11Vs namely, Volume, Velocity, Variety, Veracity, Validity, Volatility, Value, Variability, Visualisation, Valence and Vulnerability, which constitute the main characteristics that have a direct or indirect impact on the escalating privacy and security issues in big data. Secondly, we considered the big data system life cycle with three main phases, namely big data acquisition, big data storage and big data analytics. Further, the 11Vs of big data were mapped to each phase based on the privacy and security perspectives. Finally, we provided four practical strategies spanning every phase of big data life cycle by using the application of certain contemporary technologies to address the privacy and security challenges of big data. Four existing technologies using popular techniques of data provenance, data encryption and access control, data mining and blockchain were discussed with suitable adaptation in order to address the security challenges encountered throughout the big data life cycle. [30]

REFERENCES

[1] M. Chen, S. Mao and Y. Liu, Big Data: A Survey, Mobile Netw. Appl., 19 (2014), 171-209. doi: 10.1007/s11036-013-0489-0

[2] W. Tian and Y. Zhao, Big data technologies and cloud computing, Optimized Cloud Resource Management and Scheduling Theory and Practice, (2015), 17-49.

[3] C. L. McNeely, J. Hahm, The big (data) bang: policy, prospects, and challenges, Review of Policy Research, 31 (2014), 304-310. doi: 10.1111/ropr.12082

[4] A. Gandomi, M. Haider, Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management, 35 (2015), 137-144.

[5] D. Laney, 3D Data Management: Controlling Data Volume Velocity and Variety, META Group research note, 6 (2001), 1.

[6] J. Frizzo-Barker, P. A. Chow-White, M. Mozafari, et al. An empirical study of the rise of big data in business scholarship, International Journal of Information Management, 36 (2016), 403-413. doi: 10.1016/j.ijinfomgt.2016.01.006

[7] T. Huang, L. Lan, X. Fang, et al. Promises and challenges of big data computing in health sciences, Big Data Res., 2 (2015), 2-11.

[8] B. Nelson, T. Olovsson, Security and privacy for big data: A systematic literature review. In: 2016 IEEE International Conference on Big Data (Big Data), (2016), 3693-3702

[9] M. Li-chuan, P. Qing-qi, L. Hao, et al. Survey of Security Issues in Big Data, Radio Communications Technology, 41 (2015), 1-7.

[10] F. Deng-Guo, Z. Min, L. Hao, Big Data Security and Privacy Protection, Chinese Journal of Computers, 37 (2014), 246-258.

[11] N. B. Kshetri, The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns, Big Data & Society, 1 (2014), 1-20.

[12] X. Jin, B. Wah, X. Cheng, et al. Significance and challenges of big data research, Big Data Research, 2 (2015), 59-64. doi: 10.1016/j.bdr.2015.01.006

[13] W. Xindong, Z. Xingquan, W. Gong-Qing, et al. Data Mining with Big Data, IEEE T. Knowl. Data En., 26 (2014), 97-107. doi: 10.1109/TKDE.2013.109

[14] V. Chang and G. Wills, A model to compare cloud and non-cloud storage of Big Data, Future Gener.Comp. Sy., 57 (2016), 56-76. doi: 10.1016/j.future.2015.10.003

[15] Z. Goli-Malekabadi, M. Sargolzaei-Javan, M. K. Akbari, An effective model for store and retrieve big health data in cloud computing, Comput. Meth. Prog. Bio., 132 (2016), 75-82. doi: 10.1016/j.cmpb.2016.04.016

[16] N. Kumar, A. V. Vasilakos, and J. Rodrigues, A multi-tenant cloud-based DC nano grid for self-sustained smart buildings in smart cities, IEEE Commun. Mag., 55 (2017), 14-21.

[17] S. Subashini and V. Kavitha, A survey on security issues in service delivery models of cloud computing, J. Netw. Comput. Appl., 34 (2011), 1-11. doi: 10.1016/j.jnca.2010.07.006

[18] H. Cheng, W. Wang, and C. Rong, Privacy protection beyond encryption for cloud big data. In: Proceedings of the 2nd International Conference on Information Technology and Electronic Commerce, (2014), 188-191, IEEE.

[19] P. Jing, A new model of data protection on cloud storage, Journal of Networks, 9 (2014), 666-671.

[20] C. Liu, C. Yang, X. Zhang, et al. External integrity verification for outsourced big data in cloud and IoT: a big picture, Future Gener. Comp. Sy., 49 (2015), 58-67. doi: 10.1016/j.future.2014.08.007

[21] H. Kun, L. Di, L. Minghui, Research on Security Connotation and Response Strategies for Big Data, Telecommunications Science, 30 (2014), 112-117.

[22] T. Matzner, Why privacy is not enough privacy in the context of ubiquitous computing and big data, Journal of Information, Communication and Ethics in Society, 12 (2014), 93-106. doi: 10.1108/JICES-08-2013-0030

[23] D. Thilakanathan, Y. Zhao, S. Chen, et al. Protecting and Analysing Health Care Data on Cloud. In: Proceedings of the 2nd International Conference on Advanced Cloud and Big Data, (2014), 143-149, IEEE.

[24] I. de la Torre-Díez, B. Garcia-Zapirain, M. Lopez-Coronado, et al. Proposing telecardiology services on cloud for different medical institutions: a model of reference, Telemedicine and e-Health, 23 (2017), 654-661. doi: 10.1089/tmj.2016.0234

[25] G. Lafuente. The big data security challenge, Network Security, 2015 (2015), 12-14.

[26] R. Lu, H. Zhu, X. Liu, et al. Toward efficient and privacy-preserving computing in big data era, Network IEEE, 28 (2014), 46-50.

[27] J. W. Crampton, Collect it all: national security, Big Data and governance, GeoJournal, 80 (2015), 519-531. doi: 10.1007/s10708-014-9598-y

[28] D. Lyon, Surveillance, snowden, and big data: Capacities, consequences, critique, Big Data & Society, 1 (2014), 1-13.

[29] X. Hu, M. Yuan, J. Yao, et al. Differential Privacy in Telco Big Data Platform, Proceedings of the VLDB Endowment, 8 (2015), 1692-1703. doi: 10.14778/2824032.2824067

[30] M. Benjamin, S. B. Michelle and T. B. Nadya, Eigenspace Analysis for Threat Detection in Social Networks. In: 14th International Conference on Information Fusion, (2011), 1-7, IEEE.