



Review Article

Survey on Big Data Analytic and Challenges to Cyber Security

¹Anandakumar Haldorai, ²Umamaheswari Kandaswamy and ³Arulmurugan Ramu

¹Department of Computer Science and Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

²Department of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India

³Department of Computer Science and Engineering, Presidency University, Bangalore, Karnataka, India

Abstract

To evaluate and inspect a high-quality diverse and existing definition of data, technological innovations are enhanced regardless of the evidently tumbling storage funds in developing, collection of data and statistics in computation alongside developed computers processing power. Using these innovative applications known as big data in formulating both internal and external sources of data, covering connections that define data can easily be identified. Aftermath, relevant strategies can be expressed to associate the identification of big data with new technologies and economic expansion. Big data can be framed in two perspectives grouped both negative and positive. The innovation promised much in the future but also points out some major security problems, ethical consideration and personal security matters. When these issues are not put into consideration, they will later form major obstacle to initiate success and consummation of utilizing big data innovations. This paper mainly considers analyzing current usage of big data that can be considered both in personal levels and the community as a whole while also concentrating on seven significant area of usage. These key areas are big data and health-care, big data for business optimization and customer analytics, big data and science, big data as enablers of openness and efficiency in government, big data and finance, big data and emergency of energy distribution systems and big data security systems. Moreover, compelling issues in privacy, ethical concerns and security which have been outlined will also be illustrated.

Key words: Big data, business optimization, technological innovations, security problem, innovative applications

Citation: Anandakumar Haldorai, Umamaheswari Kandaswamy and Arulmurugan Ramu, 2018. Survey on big data analytic and challenges to cyber security. Inform. Technol. J., CC: CC-CC.

Corresponding Author: Anandakumar Haldorai, Department of Computer Science and Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India Tel: +919486137910

Copyright: © 2018 Anandakumar Haldorai *et al.* This is an open access article distributed under the terms of the creative commons attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Analysis of big data is not an extremely new phenomenon in the world because it has existed and discussed by many analytics. Despite the fact that this phenomenon has been viewed to be a so sensitive slogan, big data analytics is a form of over-focused large data in identification of obscured patterns, consumer preference, new associations, market trends or preferences and other information related to business. Many business commercial structures are guaranteed to undertake a skip in the decision-making process in order to maintain a competitive business environment whereby they have to intensively control data to formulate strategies that can be implemented¹. Over the past few decades, major explosions in the magnitude of data and much of its output at personal levels have been observed. These extremely large data is easily duplicated at low prices which is conveniently stored in databases available in the public domain.

According to an estimation done by the IBM published recently, a significant total of 2.5 billion GB of data is released everyday throughout the entire world with its volume evidently growing very single minute. Other aspects to add that correlates with Web 2.0 applications is the factual foundational costs concerning computation storage, the sector of artificial and mining of data that is the process of fundamental innovations breakthroughs and fast penetrating of a wide range of computation prototype such as cloud computation². These aspects all join hands with a collection of sensors which are endowed mobile friendly instruments which help intensively in enhancing big data. Big data does not have a strict description but can be used to explain a benevolent feasibility and wide-range of data in-conjunction with the nature of momentum at which data of various formats, nature and origin are transited and formulated.

The United States National Institute of Standard and Technology and study specialists Gartner in association have enhanced and promoted a practicable definition of large quantity of data involving three key areas of data: Volume (data amount), velocity (data speed) and variety (data types and sources). Big data cannot be related to traditional forms of data warehouse or any other analysis area of commercial intelligence which has been in existence for a few decades³. Large data that is quite unsurpassable is not structured and entails literally raw form of data that is created with an increasing velocity compared to what it was before.

In today's world, analysts are constantly able to wring over extremely difficult patterns of data, identification of correlation and jerk out of data which is valuable from collection of actual time data over the internet by nature. This

is accomplished by using technologies high in performance, statistics correlation algorithms, storage of infrastructure nominal on pricing and extensive techniques in mining of data. Samples of big data sources are varied and rich in nature ranging from big corporate business like government directories and intranets that are available on the web. The immense volume of searches, mobile traces and networks, records of users and all their online social media interactions and cyber corporal schemes such as Intelligent Transport Systems (ITS), clean energy distribution, modern home equipment and smart cars that emerge into many entertainment necessities and domestic appliances that use emotional recognition, face and movement sensors.

This trend is growing at a faster rate and used in many services which affect our daily life activities and our social-economic scenarios. Big data analysts are capable of deriving and applying algorithms and utilize artificial intelligence that uncovers sealed insights from extremely large data that is obtained from various parameters of life. These parameters range from decision-making optimization obtained from a given collection data that can be used in the police force by tactical actions that are proactive and help reduce crime rates and, hospitals to reduce patient risks aspect over a certain illness which are calculate alongside the massively spreading contagious infections.

This also helps in comprehending the nature and means of interaction of human beings and consequences that comes with social and technical situations. Currently, the acceptance of data is considered as a form of currency in rare commodities. The sustain ability of large data reduces to undergo privacy issues at both personal and community levels because these data are linked with other databases. Storage, analytics and decision-making process are automated using computation algorithms have a significant impact on both the individual and community but also pose a significant quantity of threats such as prejudicial results and unfair rates of discrimination. In world today, institutions and business are wholly driven by data. In that regard, the utility of big data and its wide expend of globalization has significantly affected the rate of innovation, productivity and growth of the economy. This factor enables both businesses and the society to reap the benefits of big data extensively.

BIG DATA FOR OPTIMIZATION OF BUSINESSES

Application of big data is a fundamental cornerstone considered as the key in today's innovative business environment. This implies that personalized service provision will be delivered in organizations and could also be applicable to marketers and pioneers⁴. The concealed

patterns can be viewed and isolated by analysts in business impacting an in-depth knowledge from a variety of sources (both external and internal) obtained from assimilation of advanced analysis of large amount of data and today's warehouse data sources. The knowledge of big data in that case acts as a leverage to organizational leaders because they significantly add more knowledge to the process, operational efficiency and achieving a competitive advantage in the end. When a large amount of data is interlinked which are complex and extremely heterogeneous together with data that is available externally enhances short-listing of effective strategies by analysts and pioneers⁵. These strategies are applied in advertisement and marketing campaigns used when identifying the exact customer needs, usage, purchasing trends and other development patterns which can be identified at any level of customer engagement.

Most online interaction and other organizational retailing rely massively on big data innovations to dealing with clients and histories of their customers during purchases, transactions and database inventory to:

- Collect a greater degree of knowledge about clients
- Delivery of personalized services, goods and guidance that is recommendable to potential customers
- Outlining of any changes in client-based necessities

In consideration to the growing amount of data, clients who depend on mobile devices for all their purchases, transactions and utility of electronic cards and managing the use of OSN to register opinions to share personal and intimate views can allow marketing experts of organizations to attain leverage using basic analytics technologies. This aspect is evidently effective in future because its course has already been established. This implies that a healthy competitive commercial situation and continuance of activities are kept stable in the near future keeping in mind technological changes and marketing trends.

BIG DATA AND BIG SCIENCE

The dimension of science can be caused to change with the assistance of highly effective big data. The prospective future of data intensiveness and science is made using developments that were witnessed in the past few decades. The section of increased performance in computation simulation alongside practical analytics and continuous growth in amount of data obtained from various databases such as medical data, internet browsing, security patterns,

video surveillance data, genomic data that output data of global images, wire-free networks, sensors and mobile networks that assist in the growth of large amounts of data and the significant of big science. Nearly new views to mitigate challenges of invention and exploration in the area of science looms so much with future of big data when science is used.

The universal volume of data analyzed allows a significant amount of data output which require simulation, new tools, computation and management of data. This form of approach of data intensive big science allows the lease of promises that integrates a couple of features of life in social and physical sciences. This aspect also covers the field of application from computation of the earth, nature, genomics and environmental research to computed social studies. This is achieved with a completely pure objective and ambition of mankind which will be enhanced by using big data intensive research to face new global challenges like, pandemics, global warming, monitoring global health issues and usage of energy resources.

BIG DATA IN MEDICINE AND HEALTH CARE

The health-care sector as a whole has begun its dependence in all aspects of healthcare divisions such as those monitoring public health-care, delivery of services and research of healthcare analysis of big data. The evidently expanding dependence on IT, company's stakeholders and economics are capable of collecting, developing and distributing various forms of data using systems that constitute samples of biological data of individuals, medical imaging, medical prescription, health analysis and health-care statistics⁶.

In healthcare setting, maintaining records in all listed data above has been analyzed using age. However, what is new is:

- Capability of interlinking internal and external health facts when considered as a whole to formulate patterns of geographic, health fitness and behaviors over large number health data at an individual level. This allows formation of patterns that help understand healthcare setting in a certain geographical location and its citizens
- Assistance rendered in understanding research, examination of new medical standards and meticulous remarks of obscured health tendencies in a certain society which will however permit better medication goods and services to them

FINANCIAL SERVICES AND LARGE DATA

Actually most of large associated data technology exhibits significant involvement in financial organizations which has led to the rapid growth of the IT sector. This type of technology is mainly utilized in critically investing large quantities of financial information based on economic and personal related such as the market's shares and some stocks. Using this it brings a clear implication that it aids financiers understand the various risks and challenges encountered while using these systems⁷ during financial markets and the various possibilities in seeking fresh investments. In the past years various financial institutions have been using various forms of credit forms that are used in capitalizing the technological companies.

Moreover, such companies are ought to developing the social networking so that they could be easily accessed by many people. Such institutions always aid in identifying several counterfeit behaviors through pointing out some of the complicated designs. On the other hand, such large data monetary accomplishment strategies always aid in drafting complex decisions in other diversified markets places. The basic data will be critically evaluated, feasible and the most profitable trends that would be very difficult in tracing. These will ease the access of significant skills and knowledge so that one can easily established the transition that would emanate in the stock markets accounting the radical conclusion which would create an outstanding effect on some financial organizations.

Large data in rising energy distribution systems: The revolution of the smart grid systems in the energy systems gains momentum in its handling. In order to gather the newer data the techniques and the analytical devices of data-driven use the safe field devices, smart meters and IT components with most modern energy configuration. The realistic in hand with real life completion along with scrutinizing large data aids the force sector to reach to the finale of developing modified ways of generating power, its contribution, allocation and transmission.

Large data in the government sector: The immense revolution of the administration to exhibit large data in the earlier periods had noticed a huge importance all through. An enormous magnitude of public particulars like the physical condition data, population census, meteorological statistics, crime details have been published in the internet for the community to see. This is normally carried out to maintain the

open end liability and clearness of data so as to upgrade the national value in the community. The nation thinks that by permitting this simple admittance and the use will let up a redoubtable financial expansion and also mend recent innovations by the non-government parties and also commercial agency⁸. It's evident that the provision of data by the government shows an affluent source of massive data mine. This data can be well utilized by both public organizations and private ones to higher levels as been predicted by the British government.

The ever-present amount of mismanagement of data by the government data encourages the private sector and some agencies to improvise modern mechanism in data processing and also analysis and therefore aid in government activities efficiency and also cost minimization. From European public sector around 100 billion dollar per year is saved and this is achieved through disabling the usage of large data in its value in the operation. This is encouraged in other nations also. For instance in the private sectors the real estate scenario vary extremely with its awareness of accessible property, where it's found and its value in addition to its criminal data. Real estate is well prepared and organized to educate their customer about the proper place they can invest at.

Using big data in identifying cyber crimes: Most evidently cyber-crime has recorded an outstanding lead in both the private and public sectors. In order to fight such crimes one needs to establish the potential tools and equipment that would be able in tracing any criminal activity. This takes into account various criminal's activities and behaviors that would negatively affect many individuals. Professionals who are involved in solving cyber security related problems always invest much effort in making sure that they safe guard their institutions from some malicious cyber criminals who may engage in some illegal actions⁹. As a result of the accompanying increasing rates of cyber felonies most of the private cooperation always tend to use some of the large data diagnostic devices and high technological innovations for upholding their privacy.

Thus the varying tendency brought about by the unwanted attacks by such cyber criminals would be openly recognized. Moreover, these will create ease in identifying some general patterns based on movement, inspiration and intentions with the constant aid which helps one know more about some security information. Moreover these kinds of data are shared among different companies also in the other countries and information concerning the cyber-crime is shared in order to detect and minimize cyber-crime rate. It also

aids to recognize similar networking activities in major states and country jurisdiction. This collection of large data and its study of security and information assure the yearning of the law menders and its agencies of enforcing the law for recent information from various sources like financial data, traveler's data, internet usage and many other ways such as satellite images and surveillance videos. When all the data are carefully collected and put as one for the purpose of analysis, security experts are proficient to identify the criminal behavioral patterns, which they are and more other information approaching.

ACTUAL TIME PROBLEMS AND CHALLENGES OF LARGE DATA

It's evident that one's personalities, societies, companies, organizations, public and private sectors get an immense deal of optimistic opportunities with the consumption of large data techniques and its different tools however, there is uncertainty on significant issues of confidentiality and morals. The large data diagnostic devices or mechanism and configuration will have recognizable result which may be pessimistic to the privacy aspect both lawful and ethically¹⁰. This will be the exact stumbling block to the potentiality of large data theory which is to be predicted at this point.

Challenges to security and privacy in large data: One of the challenges is the loss of civil rights led by the consumption of large data diagnostic tools on different areas like social, economic, financial and transactional as there is total deficiency of personal autonomy and privacy. In result they mislay the authority to uphold the power and scrutinize their private data and avoid any exploitation and/or mistreatment by cyber criminals, data analyst though the usage, enlargement and innovation relays on large data maintenance and utilization.

The following cause a few virtual threats to notion of large data's privacy and security.

Improved potential for huge scale theft or violate of susceptible data: The possibility of breach in data is increased due to the increased amount of data in stored, accessed and shared online through the internet by third parties. It results to a long list of questions concerning the personal data's access, storage and usage which is of supreme significant. There are two dissimilar components in unlawful access such as:

- Primary adversary likes to enter into data base of raw data with an intention to concession the way of interpretation

or analysis. It is performed by processing wrong data into the database or takes over huge volume of recognition or financial data which are highly sensitive. The other difficult act is to act upon previous analyzed databases and steal the actions and intelligence of the lawful analyst of the large data¹¹. To carry out this violate of privacy in data, both the software and the hardware of these data platform are scrutinized and their flaws are utilized. Hence the enormous data infrastructures such as cloud platform and the data centers are available to assault are greater protected from security breach and malware attack. All the raw data and inferred knowledge which are really critically are stored in these places

- A person may endure from recognition disclosure, private data like credit, debit card details, financial transactions while a company's breach at last result to severe destruction to its brand image, faithfulness of the partners and its consumers, loss of share market value, intelligence data and fines that are lawful which are in compliance with some solitude policies

Lack of self-control over individual data: There has been a constant appreciation of the large data systems and an accumulation of data which has brought significant challenge to sharing one's individual data. The main backbones of this subject include:

- In the aspect of large data relation, the IT sector always progresses significant amount of data among various individual through acquiring and keeping them so that to attain a conclusion. This always happens when it is not on another person's hands
- The main notion of large data entails the up keeping of much data which comprises of basic security protocols which requires most institutions and personnel that are involved at different wars
- It is essential for the reduction of data and its principle requirements which exert vast opposition to attain consecutively based on the large data trends principles
- There is always contradiction related to data possession extracted using large evaluation which is similar to other social media and e-commerce sites

The mentioned problems they bring about concerns on ones right to regulate and manage their individual data both completely and clearly but in addition whether exposure at proper point can be maintained. Moreover it's hard to know whether monitoring and managing of private data can be successful all through¹². Collectively, there is a critical

issue on the choice to access data of one as per the regulation of European Data Protection Directive which is intimately connected to the personal privacy and management of their personal data.

Elongated period of accessibility of insightful datasets: The government, private organizations and individual researchers are able to accumulate massive data from platforms due to the fluctuating prices brought about due to the developing data requirements. There is rare establishment of one's identity in the database, which explains much of the rest of the behavioral activities, existence and one's attitudes will always be acknowledged. There is high possibility based on the data that always changes one's personal lifestyle¹³. There has been significant transition on the social digitization which explains various debates rose based on one's personal technological life. Much of the data can be easily be manipulated and duplicated making ease of access by most people on the web which reduces loss of data abruptly. Large data which consist of enormous amount of IQ always stores many records that cannot be deleted permanently which relieves people who incur ridiculous mistakes by many people¹⁴. All this would not be put into consideration by any individual based on a particular discipline.

Reliability of data and the derivation of issues: Most of the applications which are using the large data service always comprise of enormous data that are extremely perceptive which the surrounding environment. As a result it requires clear contemplation of the data subjects such as history and genealogy more so bearing in mind some of prospectus relating to the quality of the data, the origin and honesty of the data. All these are clearly related to some potential difficulties related to the analytic of the huge data sources. Such data forecaster always find very difficult in analyzing some possible means of acquiring a steady integrity also and the quality of the data¹⁴. This always results to different difficult scenarios in developing a business, so that one can be able establish a rigid colossal amount of data and also proceeding with any data related activity that would need data driving. The activity section of how the analytic is handled depends on the data quality, its honesty and the integrity fraction. There should be maximum concentration based on data independence and the optimization of calibrated and those that are strategic in nature¹⁵. This would create constraints in any institution and the uncovering of some vices which are investigated by the law enforcement agencies.

Un-preferred data relationship and inferences: The whole amass of the interlinked significant sum of data that most individuals use it in accessing of a variety of sources, always aid in classifying most of the hidden designs that always adjoin to the series of large data related perils¹⁶. The large data is effectiveness is stipulated to increase some security protocols on all the database connection. In the large data sites, always the collection of data among some varied scenery. This will entail a backbreaking activity that constitutes different challenges of re recognition¹⁷.

In adequacy of proper management and transparency: The main two lawful requirements that are desirable by many people based on both private and the public institutions always aid in accumulate, process and distribute personal identifiable data which tends to be of electric approval and the some various notion of notice. Around the world, such techniques always comprise of various fundamental position which are in maintained based on privacy and data safety¹⁸. Though this is an ordinary process since many states and countries may constitute diverse controls which explain differently the idea. The accomplishment of such policies is always a perilous task in the aspect of large data surroundings since there is always transition from different people in the event of giving out options of checking out which makes people bring out consent decisions¹⁹.

Despite of much detailed instruction explained by many individuals some of the practices handled by many people always tend to be very coherent and there is always reduced concentration which is paid to such processes during the setting up of some digital services. Due time, the effects of such actions will explain differently and the circumstance will be motivated by the processing of the large data²⁰. Therefore at these points of focus most of the data that would have been accumulated will always lack a concretes inference and also null prediction will be done and most people will not be able to give a clear explanation on that. Additionally, most forms of data and info accumulated using the aid of the large data technology consists of many reasons and contexts which are used by many people in connections to their different data which have the opportunity of having significant amount of alter²¹. In relation to such trend, it is always difficult for many users to get access to most of their common wants which raises concern when it is in relation to some common analysis which would significant benefit some of the private corporations since they always comprise and always reveal much based on a different point of idea.

Truly, most people are always not familiar to some of the procedures involved in the large data operations technique, their infrastructure and the operation of some to their algorithms. These are always achieved without the recognition of the data subject in context. Moreover they are always not familiar to some of the management procedures of the consent management that always aids in decision making processes²². Therefore there is always the need of openness in instilling knowledge in the subjects which should be handled by some huge data systems. Such restriction related to the large data compilation of the consent mainly brings out change in the strategies which always supports different specifics and clarification in some of the consent management, cleansing the consent process or rather the notice management for example, aiding alternatives to current all or nothing approach that poorly satisfies every users needs and the execution of some processes also the revocation of other data determined business forms of the evolvement of the large data technology.

Huge accountability: The responsibility of some huge data related algorithms always exerts significant challenges which come along with their openness eligibility. The huge data algorithms always aid in attaining maximum output within the society wholly which have a disapproving influence among some individuals related to most increasing profiles, surveillance handled and other types of prejudiced techniques. In order to respond to such problems and results such issues need to be addressed with essentially makes it able to change the trail from the current approach to algorithmic responsibility.

Moral issues and social challenges: Besides, other forms of technological questions in context and other lawful those have been analyzed in more critically. This brings that subject of the large data to be one of the largest ethical and moral implications which creates a huge obstacle that should be analyzed. Below are some of the few listings that should be considered.

Information asymmetry and the issuing of power: Data asymmetry entails always an environmental business issue which involves one party possessing huge amount of data and also information involves being in a position to utilize the data that you have at hand. This can be exhibited through the retail recent transactions whereby the buyer comprises minimum skills that would stand aside in the business sector. Such skills are always that a retailer acquires are always more

than that of the buyer revolving around evaluation of the quality and the accompanying rates of items on sale which is determined by the market values over a long period of time. On the other hand the buyer always has to depend on restricted knowledge and skills and also will abide the retailer's wishes. The irregularity in different data allows data to be selected by withholding power over a wide range of ordinary with the assistance of huge data. Most people in the generate public are always blocked in having access to some knowledge. This brings about a situation where there is possession of authority among people of the state and business establishment and other desirable customers. Therefore when such amount of extraordinary data is accumulated and the ease of maneuvering data is permitted by many organizations, it would result in organizations attaining their mean desires. Additionally most governments and delegates would get rid of some of the unnecessary power, which creates some inconceivable amount of effect which will be linked democratically to a country.

Huge surveillance: The big data technology has a specialized field referred as surveillance which holds a concrete background. It is essential to know that there is constant growth in various systems related to the surveillance when one is online and also offline. Such systems are always not enacted by the business sectors and individuals who tend to take control of intelligence and law organizations. The large data algorithms are always used service providers when these systems are always online. On the other hand they amass and inspect significant amount of data based on the customers desires and tastes which tends to boost peoples services and aims in aiding the commercialized advertisements. With the constant and nonstop surveillance among the customers aids in investing huge amount of skills based in the customer's need, tastes and the daily expenses and always affects one's way of life over a long period of time. Currently, there is always nothing that tends to be very secretive any longer due to the large amount of data that categorically results to being outrageous. When using some of individual's files it aids most organizations who always tend to look for sensitive data based on different individuals thus helping in maintaining a label while in the internet reality.

Categorical social control: Conservatively, it is always postulated that some of the minor section of a given population always requires maximum innovation and need much attention. Although in the aspect of large data, such data is limited to various individuals since it comprises only

minimum details such as privacy, fairness and some ethics which in turn becomes very questionable. With the use of large data, there will be minimum biasness mainly focused at specific people and this is always achieved in an easy manner with minimum forecasting accuracy. With optimistic nature of the large data always permits the compilation and the study of huge amount of data which always nourishes the effectiveness of some operations. Thus the making of decisions always provides some of the important needs of some people and such type of technology always tends to isolate the similar kind but on the wrong hands. In some periods there is always some discrimination of some law enforcement, consumer scoring and some weird behavior could be handled in a negative manner.

Thus higher replica in the mathematic models always help in comprehending some of the distinctions and similarities of some groups which always aid in classifying them in some cadres. In such scenario it is always used in the opposite manner since most disabled individuals may encounter the challenges of discrimination but are always stored in isolated areas since they do not have the sufficient means of rescuing themselves. In the department of healthcare society and the insurance corporation they always utilize the big data at a high accuracy level. They always determine one's data the extent to which they have been affected by the sickness and most probably some to the patients historical levels will clearly explain the main problem facing an individual. On the other hand, it also permits in distinguishing some of persons based on one's gender biased course.

Additionally when aimed at the real estate department, obviously the data based on the property and different crime statistic which have been exposed by official organs always aid at creating an area which initiates the isolation of most deprived personnel and their homeland will depreciate in value. Such methods could be utilized by most complicated and well off hotels airline corporations so that they can accurately analyze and gather data based on the extensiveness of the rooms, sufficiency of the seats and also the supply and demand of the services. Such result always initiates a clear variation in the fluctuation in the prices and its availability in improving their profits.

Therefore the environment recommends that urge is to use up in respect to a particular services or computational knowledge. They also give out the credit worthiness that is based on the large data models which consists of rational facts and data such as one's sex, religion, income, race, location and the accompanying routine of buying a product. On this aspect, the customer will always be isolated for good through

categorizing some who are handled in a special way. With the law enforcement organizations, the use of the large data investigation always set a position and subdivides the society in minor sections and with the insufficient skills and knowledge a wide range of correlations is always achieved. Therefore the ethnic groups are always shaped in order persuade the dark nature of the preventing strategy. There will also be evaluation of civil rights and the derivation of some pessimistic labels which are easy to attain.

CONCLUSION AND FUTURE RECOMMENDATIONS

This study discovered the big data analytic that can be beneficial for Cyber Security. The impending figure of the large data analytics has been thought to be a worldwide business happening which has come up and assures maximum security to many individuals. This type to technology has been established through the innovation of the large data technology which has shown to be applicable in today's business world. On the other hand essential decision making processes can be effectively be achieved through various analytical techniques and tactical information administered by some relatively associated organizations. Being one of the upcoming data evaluator, an individual should be heavily be trained and equipped in utilizing the three figures which are, volume, variety and velocity. Moreover they should be well familiar with the current technologies related to the large data technology. Thorough understanding and the enacting of this kind of technology by the high ranked professionals always play a significant role in the decision making process and the establishment of better aiding policies. With the rise of the large data technology the future will be able to witness significant use in its application in most of the business institutions. There has been kick offs whereby most of the BDM institutions have amalgamated allowing most of the skilled and trained personnel to trace some of the government models. This aids in achieving quality decision making processes in relation to the accompanying expenses of the IT sector. Since it's not a new idea that requires amassing, most of the business organizations always tend to utilize algorithms that are always difficult in nature. Additionally, most of these devices always entail the structuring of data and the meaning of extraction of various composites thus achieving better results. This study will help the researcher to uncover the critical areas of Cyber Security that many researchers were not able to explore. Thus a new theory on Big Data may be arrived at.

REFERENCES

1. Jiang, Y.G. and J. Wang, 2016. Partial copy detection in videos: A benchmark and an evaluation of popular methods. *IEEE Trans. Big Data*, 2: 32-42.
2. Anandakumar, H. and K. Umamaheswari, 2017. Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers. *Cluster Comput.*, 20: 1505-1515.
3. Kaseb, S.A., A. Mohan, Y. Koh and Y.H. Lu, 2017. Cloud resource management for analyzing big real-time visual data from network cameras. *IEEE Trans. Cloud Comput.* 10.1109/TCC.2017.2720665.
4. Zhang, C., W. Shang, W. Lin, Y. Li and R. Tan, 2017. Opportunities and challenges of TV media in the big data era. *Proceedings of the IEEE/ACIS 16th International Conference on Computer and Information Science*, May 24-26, 2017, Wuhan, China, pp: 551-553.
5. Lecuyer, M., R. Spahn, R. Geambasu, T.K. Huang and S. Sen, 2017. Pyramid: Enhancing selectivity in big data protection with count featurization. *Proceedings of the IEEE Symposium on Security and Privacy*, May 22-26, 2017, San Jose, CA, USA., pp: 78-95.
6. Li, T., J. Tang and J. Xu, 2015. A predictive scheduling framework for fast and distributed stream data processing. *Proceedings of the IEEE International Conference on Big Data*, October 29-November 1, 2015, Santa Clara, CA, USA., pp: 333-338.
7. Park, G., L. Chung, L. Khan and S. Park, 2017. A modeling framework for business process reengineering using big data analytics and a goal-orientation. *Proceedings of the 11th International Conference on Research Challenges in Information Science*, May 10-12, 2017, Brighton, UK., pp: 21-32.
8. Shamoto, H., K. Shirahata, A. Drozd, H. Sato and S. Matsuoka, 2016. GPU-accelerated large-scale distributed sorting coping with device memory capacity. *IEEE Trans. Big Data*, 2: 57-69.
9. Suganya, M. and H. Anandakumar, 2013. Handover based spectrum allocation in cognitive radio networks. *Proceedings of the International Conference on Green Computing, Communication and Conservation of Energy*, December 12-14, 2013, Chennai, India, pp: 215-219.
10. Xia, F., H. Liu, I. Lee and L. Cao, 2016. Scientific article recommendation: Exploiting common author relations and historical preferences. *IEEE Trans. Big Data*, 2: 101-112.
11. Arulmurugan, R. and H. Anandakumar, 2018. Early Detection of Lung Cancer Using Wavelet Feature Descriptor and Feed Forward Back Propagation Neural Networks Classifier. In: *Computational Vision and Bio Inspired Computing*, Jude Hemanth, D. and S. Smys (Eds.). Springer, New York, pp: 103-110.
12. Chu, W.W., 2014. Erratum: Data Mining and Knowledge Discovery for Big Data. In: *Data Mining and Knowledge Discovery for Big Data*, Chu, W.W. (Ed.). Springer, Heidelberg, Germany, pp: 305-308.
13. Arulmurugan, R., K.R. Sabarmathi and H. Anandakumar, 2017. Classification of sentence level sentiment analysis using cloud machine learning techniques *Cluster Comput.*, 1: 1-11.
14. Schmidt, D., W.C. Chen, M.A. Matheson and G. Ostrouchov, 2017. Programming with BIG data in R: Scaling analytics from one to thousands of nodes. *Big Data Res.*, 8: 1-11.
15. Anandakumar, H. and K. Umamaheswari, 2017. A bio-inspired swarm intelligence technique for social aware cognitive radio handovers. *Comput. Electr. Eng.*, 10.1016/j.compeleceng.2017.09.016.
16. Yusuf, I.I., I.E. Thomas, M. Spichkova and H.W. Schmidt, 2017. Chimney: Connecting scientists to hpc, cloud and big data. *Big Data Res.*, 8: 39-49.
17. Anandakumar, H. and K. Umamaheswari, 2017. An efficient optimized handover in cognitive radio networks using cooperative spectrum sensing. *Intell. Autom. Soft Comput.*, 10.1080/10798587.2017.1364931.
18. Zong, Z., R. Ge and Q. Gu, 2017. Marcher: A heterogeneous system supporting energy-aware high performance computing and big data analytics. *Big Data Res.*, 8: 27-38.
19. Anandakumar, H. and K. Umamaheswari, 2014. Energy efficient network selection using 802.16G based GSM technology. *J. Comput. Sci.*, 10: 745-754.
20. Shmueli, G., 2017. Research dilemmas with behavioral big data. *Big Data*, 5: 98-119.
21. Dumbill, E., 2013. Making sense of big data. *Big Data*, 1: 1-2.
22. Haldorai, A. and U. Kandaswamy, 2018. Cooperative Spectrum Handovers in Cognitive Radio Networks. In: *Cognitive Radio, Mobile Communications and Wireless Networks*, Rehmani, M.H. and R. Dhaou (Eds.). Springer International Publishing, USA., pp: 47-63.