



**National Conference on
Futuristic Development of Computing Technologies in
Science, Arts and Humanities**

9 to 11 Feb. 2023

**MATS School of Information Technology
(Information Technology & Computer Science)
MATS UNIVERSITY, RAIPUR (C.G.)**

SECURITY AND PRIVACY ISSUES IN IoT GENERATED BIG DATA

R.P. Waghmare¹, Dr. Gitanjali Sinha², Dr. B.T. Jadhav³

1. Research Student, MATS School of Information Technology, MATS University, Raipur (C.G.)
2. MATS School of Information Technology, MATS University, Raipur (C.G.)
3. Yashwantrao Chavan Institute of Science (Autonomous), Satara (MH)

*rpwaghmare@ycis.ac.in



Abstract:

- Now days Big Data (BD) and Internet of Things (IoT) are popular and mostly used terms. As secure , very important and sensitive data share on internet through IoT privacy and security are the significant challenges. There are number of devices connects with internet using IoT physically or virtually. Internet of Things (IoT) is one of the supreme technological revolutions of today's era and its feasible for handling is huge. Big Data refers to a massive set of data that no conventional data management tool can handle. Speed, Variety and volume are important features of BD and due to this big data, we need to provide more security and privacy for such big data generated by IoT. There was need a security and privacy model and the number of research study was concerned with the developed of a new IoT model that can enhance the security and privacy of the users of the IoT. Authentication, Identification and device heterogeneity are the major security and privacy concerns in IoT. Basic IoT system requires to become a secure system are 1. Authentication 2. Authorization 3. Confidentiality 4. Integrity 5. Non-Repudiation. This paper aims to improve understanding the security and privacy challenges to secure future data generated in IoT.
- *Keywords: Big Data, IoT, Security and Privacy*

Introduction:

- Big Data (BD) and Internet of Things (IoT) are popular in recent era. The number of devices and applications are also in ceases day by day. Due to every human, organization, applications generated large amount of data every moment. Number of applications present on Internet like Facebook, Tweeter etc are also increases and every one shares their data on it.
- As secure data gathered by number of IoT devices , applications and organizations Privacy and Security Issues are increased. We need to improve these issues and our paper focus on it.

- **Big Data:**

- Data which is huge, difficult to store, manage and analyse through traditional databases is termed as “Big Data”.^[1] Day by day number of devices increases which connect to internet and they generated huge amount of data every second. The volume and information captured from various IoT devices and multimedia by human and organizations is increasing every moment and has almost doubled every year. The data generated due to IoT devices can be categorized as structured or unstructured data that cannot be easily loaded into regular relational databases. This big data requires pre-processing to convert the raw data into clean data set and made feasible for analysis. [1]

Big Data:

- **The Characteristics of big data are. [2]**
- **1.) Volume** – It refers to **unbelievable amount of data generated each second** from different sources such as social media, mobile phones, cars, credit cards, M2M sensors, photographs and videos which would allow users to data mine the hidden information and patterns found in them.
- **2.) Velocity** – It refers to the **speed at which data is being generated, transferred, collected and analysed**. Data generated at an ever-accelerating pace must be analysed and the speed of transmission, and access to the data must remain prompt to allow for real-time access to different applications that are dependent on these data.
- **3.) Variety** – It refers to data generated in **different formats either in structured or unstructured format**. Structured data such as name, phone number, address, financials, etc can be organized within the rows and columns of a database. This type of data is relatively easy to enter, store, query, and analyse. Unstructured data which contributes to eighty percentage(80%) of today's world data are more difficult to sort and extract value. Unstructured data include text messages, audio, blogs, photos, video sequences, social media updates, log files, machine and sensor data.
- **4.) Variability** – It refers to the **high inconsistency in data flow and its variation during peak period**. The variability is due to gathering of data dimensions resulting from multiple different data types and sources. Variability can also refer to the inconsistent speed at which big data is consumed into the data stores.
- **5.) Value** – It refers to the **hidden value discovered from the data for decision making**. Significant value can be found in big data, including understanding your customers better, targeting them accordingly, optimizing processes, and improving machine or business performance.
- **6.) Veracity** – **It refers to the quality and reliability of the data source**. Its importance is in the context and the meaning it adds to the analysis. The knowledge of the data's veracity in turn helps in better understanding the risks associated with analysis and business decisions based on data set.
- **7.) Validity** – It refers to **the accuracy of the data been collected for its intended use**. Proper data governance practices need to be adopted to ensure consistent data quality, common definitions, and metadata.
- **8.) Vulnerability** – It refers to the **security aspects of the data been collected and stored**.
- **9.) Volatility** – It refers to **how long data is valid and the duration** for which it needs to be stored historically before it is considered irrelevant to the current analysis.
- **10.) Visualization** – It refers to **data being made understandable to nontechnical stakeholders** and decision makers. Visualization is the **creation of complex graphs that transforms the data into information, information into insight, insight into knowledge, and knowledge into advantage for decision making [2]**.

IoT and Data Generated by it:

- Internet of Things (IoT) is a concept that connects physical or virtual objects to the internet. The technology very often used is the sensor, allowing to link a physical object such as a watch, a drone or even a speaker, to the internet. If for a long time the few objects connected to the Internet were the telephone and the computer, this is no longer the case today and every year new types of objects incorporating IOT technology are born.
- **Internet of Things (IoT) is an interconnected wireless network where smart nodes (IoT devices) interact with each other in order to exchange data through the communicating medium**[4]. So, there was need a security and privacy model and the number of research study was concerned with the developed of a new IoT model that can enhance the security and privacy of the users of the IoT [5].
- Connectivity and exchange of information lies at the core of smart grid functionality, which made connected devices a corner-stone for this technology. These devices are called the “internet of things (IoT)”, and enable the grid components to exchange data to maintain an up-to-date system status and receive commands to act as grid conditions change. [6]
- **IoT devices are increasing significantly in number each year, and are bringing unique opportunities and challenges with their wider implementation.** [6]
- **IoT devices generate a huge amount of data, which cannot be handled through conventional analysis techniques. This massive data is termed as “big data”, and it motivated the move towards new data analysis techniques. Big data generated from IoT devices are also exposed to security threats, and that have attracted a lot of attention as well.** [6]
- **Machine learning is a useful way to sift through big data, and extract useful information that can extensively aid in demand and generation pattern recognition, generation** [6]

Big Data Security:

- Big data are huge data sets that are very complex. The data generated is highly dynamic and this further adds to its complexity. **The raw data must be processed in order to extract value from it. This gives rise to challenges in processing big data and business issues associated with it[2]**
-

- As data shared by user may be sensitive eg: Medical Data

Security and Privacy Maintaining the security of the data stored in the cloud applications is very important. Sensitive and personal information that is kept in the cloud based applications should be defined as being for internal use only, not to be shared with third parties. This would be a major concern when providing personalized and location-based services as access to personal information are required to produce relevant results. Each operation such as transmitting data over network, interconnecting the systems over network or mapping virtual machines to their respective physical machines must be done in a secured way. [2]

Solutions/Proposals to Address Big Data Security and Privacy Challenges:

- Basic IoT system requires following to be fulfilled in order to become a secure system. [7]
 - Authentication ,
 - Authorization
 - Confidentiality
 - Integrity
 - Non-Repudiation
- There are various security principles which is helpful for proper working of these layers available in IOT.

Each layer has specific application.

- Application layer
- Perception layer
- Network layer
- Physical layer

Massive amount of device are connecting with the network communication device and IOT devices day by day [8].

A Case Study in A Secure Social Application:

- **Redbus:**
- Redbus is an online travel agency for bus ticket booking in India.
- Redbus decided to use Google data infrastructure for data processing and analysis in order to improve customer sales and management of the ticket booking system.
- Google BigQuery enabled RedBus to process massive amounts of booking and inventory data within seconds.

- Applications that reside on multiple servers continuously streams customer searches, seat inventory and booking information to centralized data collection system.
- This massive data collected is then sent to BigQuery, which runs complex queries and within seconds provide answers to various analytical queries such as how many times a customer searched for a destination and found very few bus services available, any technical issues that might arise during booking and notify the necessary team.
- This whole infrastructure helped RedBus to fix glitches quickly, minimize lost sales and improve customer service. [3]
- If there is no any privacy and security then all the data can be used for advertise of product and services and customer booking are affected using Redbus application. So security and privacy are important aspects in such applications.
- We need to improve security and privacy in such type of data, applications.

- **Conclusion:**

- There are large number of IoT devices and data generated by they are in huge amount. Here we conclude that data share by IoT devices is private, sensitive and by use of it by anyone can miss use it so Security and privacy is necessary for Big data generated by IoT.
-

- **Acknowledgement:**

- I would like to express my gratitude to Rayat Shikshan Sanstha's Yashavnatrao Chavan Institute of Science (Autonomous), Satara Maharashtra for supported me and offered deep insight into the study.
- I also like to my express my gratitude to Dr. Gitanjali Sinha ,MATS University Raipur C.G. and Prin. Dr. B.T. Jadhav , YC Inst of Science(Autonomous), Satara for continuous support and encourage me.

REFERENCES

-
- BIG DATA IN CLOUD COMPUTING REVIEW AND OPPORTUNITIES Manoj Muniswamaiah, Tilak Agerwala and Charles Tappert Seidenberg School of CSIS, Pace University, White Plains, New York, International Journal of Computer Science & Information Technology (IJCSIT) Vol 11, No 4, August 2019
 - Abadi, D. J. (2009). Data management in the cloud: Limitations and opportunities. IEEE Data Eng. Bull, 32(1), 3-12.
 - Sakr, S. & Gaber, M.M., 2014. Large Scale and big data: Processing and Management Auerbach, ed.
 - A survey on IoT platforms: Communication, security, and privacy perspectives Leonardo Babun, Kyle Denney, Z. Berkay Celik, Patrick McDaniel, A. Selcuk Uluagaca Computer Networks Volume 192, 19 June 2021, 108040
 - A survey of security and privacy issues in the Internet of Things from the layered context Samundra Deep, Xi Zheng, Alireza Jolfaei, Dongjin Yu, Pouya Ostovari, Ali Kashif Bashir First published: 16 March 2020 <https://doi.org/10.1002/ett.3935>
 - Risks and security solutions existing in the Internet of things (IoT) in relation to Big Data Yigliana Alvarez^{1§}, Miguel Angel Leguizamón-Páez¹, Tania J. Londoño¹ Ingeniería y competitividad ISSN: 0123-3033 ISSN: 2027-8284 Facultad de Ingeniería, Universidad del Valle 2020
 - Security and Privacy Challenges in Internet of Things Molugu Surya Virat; S.M. Bindu; B. Aishwarya; B.N. Dhanush; Maniunath R. Kounte Date of Conference: 11-12 May 2018
 - Security and Privacy Issues in IoT Environment Navod Neranjan Thilakarathne International Journal of Engineering and Management Research, Volume-10, Issue-1, February 2020 17 Apr 2020

Thank
you

