

Chapter 5

Security and Privacy Challenges in Big Data Environment

M. Manikandakumar

 <https://orcid.org/0000-0003-2648-7139>

Thiagarajar College of Engineering, India

E. Ramanujam

Thiagarajar College of Engineering, India

ABSTRACT

In this information era, big data is revolutionizing business. The data are generated by each and every user from servers, terminals, smart phones, appliances, satellites, and a range of other sensors on vehicles: military, agriculture, and the like. Anything the end users does online can be traced, stored, and analyzed. It is also possible to analyze from various diverse sources such as social media postings, credit card or e-cash purchases, internet searches, mobile phone locations, etc. Users are willing to provide their private information, linked to their real-life identities, in exchange for faster or better digital services. But, the companies yet may not have the fundamental rights of the user from a security perspective. More risks are associated with big data security. The main purpose of this chapter is to explore the security concerns and privacy issues in big data environments.

INTRODUCTION

Almost all industries largely implementing the concept of Big Data to store peta bytes data for analyzing, marketing and research based on users perspective to earn superior insights concerning the customers and their business. Since many industries noticed that, storing confidential or private information of the end users in cloud storage (Manikandakumar et al., 2015) or in the dedicated server, those companies have the sole responsiveness for protecting private data from both a legal and a commercial perspective. As with all the new technologies, security and protection seems to be addendum at best. As the definition of Big Data (Gandomi & Haider, 2015), the breaches are also too large, with the possibility of high severe reputational hurt and legal consequence than these recent times.

DOI: 10.4018/978-1-7998-7705-9.ch005

The new types of data in the organizations that need to analyze the following.

- Web streams such as e-commerce, weblogs and social network analysis data.
- Business specific transaction data like Tele Caller records, Geo-spatial data, Temporal and Customer Relational Transaction data.
- Data generated from various types of sensor such as motion, weather, energy, wind, humidity, vibration, airflow, liquid flow, pressure and health-oriented sensors.
- Textual data from the large volume of archived documents, external sources or customer interaction data, e-mails, messages, and tweets.

According to a study by the Economist Intelligence Unit (Unit, 2007)

- Nearly 90% of Industrial experts and business leaders believe that the data is now the raw and fundamental source of production to business such as land, labor, and capital.
- Using Big data in businesses improves the performance by 26% and it is predicted that it may grow up to 41% in next few years.
- 58% of business sectors say that they are planning to increase the investment in big data over the next three years.

According to Gartner (Beyer, 2011)

- About 42% of Information Technology experts stated that, they had invested in big data technology or were planning to do so within a year.
- Organizations those integrate large volume, various information types and sources into a rational information management communications will outperform their industry peers by more than 20%.

Four V's

Big data involves the “four Vs” for the efficient storage of data - volume, velocity, variety and the value. When defining the Big Data it is common to discuss the three major V's: Volume, Velocity, and Variety.

Volume

The main characteristics of “Big” are the absolute volume of data to be stored. There is no effect focusing on minimum storage as the technology and the information utilized to rapidly increasing in a fraction of time even in seconds or lesser. Thomson Reuters (2010) quoted in their 2010 annual report that, the world was “awash with over 800 Exabyte of data growing”. Likely another data storage company named EMC reveals it was about to 900 Exabyte and might raise up to 50% per every year.

Velocity

Big Data Applications used in research (Swan, 2013), telecommunications (Hashem et al., 2015), emergency care (Ramanujam et al. 2016; Padmavathi et al. 2015), Agriculture and the like has given much more attention to handling both real-time and simulated data for data access and processing. At the same

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/security-and-privacy-challenges-in-big-data-environment/270593

Related Content

Intuitionistic Fuzzy Possibility Degree Measure for Ordering of IVIFNs with Its Application to MCDM

Animesh Biswas and Samir Kumar (2019). *International Journal of Fuzzy System Applications* (pp. 1-24).
www.irma-international.org/article/intuitionistic-fuzzy-possibility-degree-measure-for-ordering-of-ivifns-with-its-application-to-mcdm/239874

Telehomecare in The Netherlands: Barriers to Implementation

H.S.M. Kort and J. van Hoof (2012). *International Journal of Ambient Computing and Intelligence* (pp. 64-73).
www.irma-international.org/article/telehomecare-netherlands-barriers-implementation/66860

Intelligent Decision Support System for Osteoporosis Prediction

Walid Moudani, Ahmad Shahin, Fadi Chakik and Dima Rajab (2012). *International Journal of Intelligent Information Technologies* (pp. 26-45).
www.irma-international.org/article/intelligent-decision-support-system-osteoporosis/63350

A Semantic Approach for Multi-Agent System Design

Rosario Girardi and Adriana Leite (2018). *Intelligent Systems: Concepts, Methodologies, Tools, and Applications* (pp. 1711-1739).
www.irma-international.org/chapter/a-semantic-approach-for-multi-agent-system-design/205854

Design of Cognitive Healthcare System for Coronary Cardiac Disease Detection

Mihir Narayan Mohanty (2019). *Emerging Trends and Applications in Cognitive Computing* (pp. 1-34).
www.irma-international.org/chapter/design-of-cognitive-healthcare-system-for-coronary-cardiac-disease-detection/218993