المؤتمر الوطني الخامس لكليات الحاسب الآلي بالجامعات السعودية
The Fifth National Conference of Saudi Computers Colleges 2022
21-22 December 2022

Security and Privacy Challenges in Big Data

Samiha Alarjani and Shakeel Ahmed
Computer Science Department
CCSIT, King Faisal University
Alhassa, Saudi Arabia

Introduction

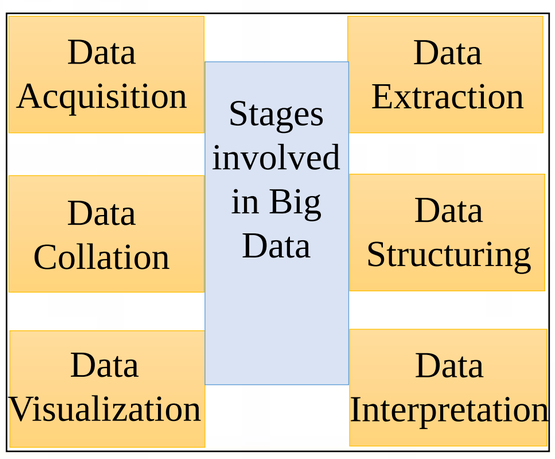
- The concept of Big Data came around 2005 and it refers to a broad range of huge datasets almost impossible to process and manage using classical data management tools due to their complexity.
- Even though most users enjoy the simplicity brought by Big Data, they also encounter a lot of inconveniencing issues.
- If Big data is not protected in a perfect way for user data, it will directly threaten the security and privacy of users' data.
- In this study, we investigate the security and privacy related issues in Big Data and the proposed protection mechanisms.

Literature Review

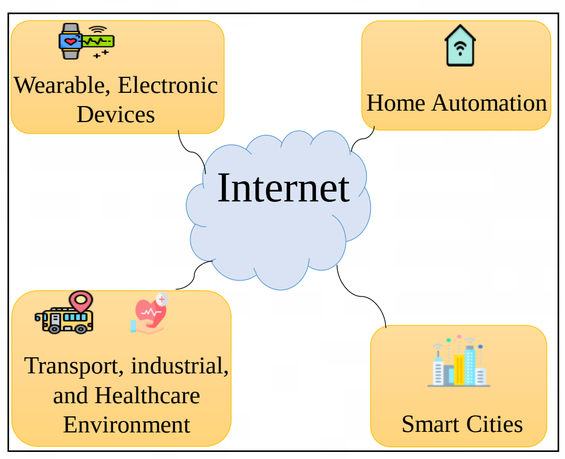
- Kupwade et al. presented the recent Big Data security and privacy issues in healthcare industry.
- Abouelmehdi et al. reviewed some related works and identified certain risks to the security and privacy of health-related data.
- Raghav et al. reviewed the security aspects of Big Data and the encryption rates of the most widely used encryption algorithms.
- Dongpo Zhang analyzed the security issues of Big Data and proposed some protection methods to be used for Big Data security and privacy.
- Yazan et al. presented the lifecycle of Big Data which is composed of four phases including: data collection, data storage, data analytics, and knowledge creation.
- LEI et al. identified four different types of users based on their roles in data mining applications which are data provider, data collector, data miner, and decision maker.

Background

Big Data Stages



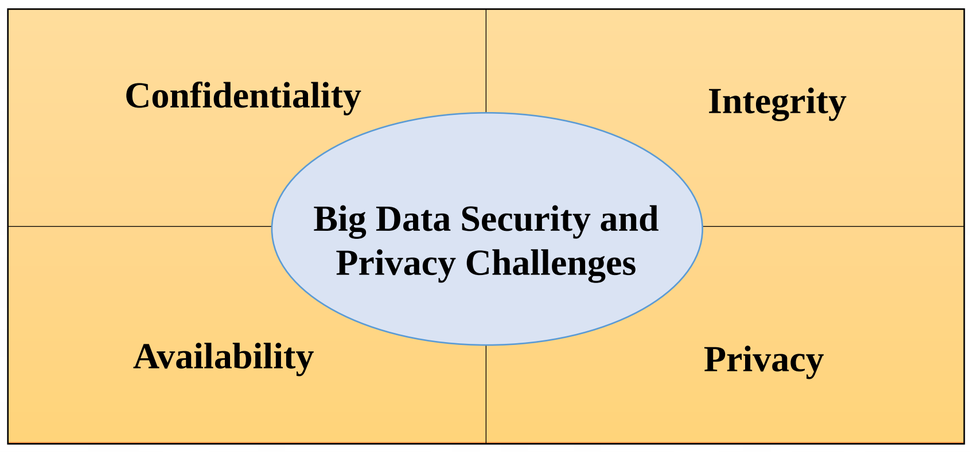
Big Data Applications



Threats to Big Data Security and Privacy

- In recent years, there has been an explosive growth in social networks which aims at discovering the interesting social patterns.
- To extract the social patterns, the organization makes use of an application that needs to share the users' data with a third party.
- Even though the identifiers of users are removed when the data is published, this publication might lead to exposures of users' sensitive information.
- Some examples of security and privacy threats are data breaches, account hijacking, insider threat, unauthorized access, and insecure interfaces.

Security and Privacy Challenges in Big Data



Security and Privacy Protection Mechanisms in Big Data

- Most widely used security and privacy protection technologies are:
  - User Authentication
  - Data Encryption
  - Data Masking
  - Access Control
  - Security Monitoring and Audit

Security Challenge	Protection Scheme
Confidentiality	- Encryption of sensitive files - User authentication - Secure dispose of data records
Integrity	- Always validate - Access control management - Always backup data
Availability	- Improve physical infrastructure - Speeding up the recovery time - Remove corrupted data
Privacy	- Establish a comprehensive privacy protection law - User authentication - Data encryption

Conclusion

- Even though Big Data has simplified our life, there are still some challenges related to Big Data security and privacy.
- Data confidentiality, integrity, availability, and privacy are among the security challenges in the context of Big Data.
- This study suggests some protection Methods such as data encryption, user authentication, data backup and access control management.
- It is recommended for future works to introduce an appropriate implementation of security and privacy mechanisms that can enhance the security and privacy of Big Data.