Chapter 22
Preserving the Healthcare Data by Using PPDM Technique

K. Renuka Devi

iD https://orcid.org/0000-0001-8202-2155

Dr. Mahalingam College of Engineering and Technology, India

ABSTRACT

The recent advancement in technology catalyse the processing of vast amount of data. It leads to the purpose of extracting and storing those data in an electronic form. Data mining contributes a significant part in fulfilling this process. In the realm of data mining, privacy and security are two important problems. To resolve the above problems, the need of privacy preserving in data mining (PPDM) has been arisen. Under healthcare domain, the data mining will process the sensitive information which includes name, age, health records etc., Hence, PPDM plays a key role in securing that sensitive information and protecting those data from the intruders. Individual data is not only protected by PPDM, but it is also protected from leakage. In this paper, the PPDM technique such as anonymization technique and cryptographic technique has been analysed deeply with the view of its execution time. It is also focused that based on comparison, which method would be more efficient for protecting data from unauthorized access in the healthcare domain.

INTRODUCTION

In the growing era of technological world, the process of storing information and accessing that information increased rapidly. The information normally falls under the category of business related as well as transactional data (Gupta & Chandra, 2020). Those collected information might be containing sensitive information as well as private data which leads to the privacy issues. It is the technique of uncovering hidden patterns and trends in data. Those information will be gathered from databases and warehouses (Gheware et al., 2014). In this technological era, various business organizations rely on different data mining tools. It involves the process of collecting both structured and unstructured data. The process available will be suitable to process structured data but tedious for unstructured data (Kiranmai & Damodaram, 2014).

Ghate and Ingle (2015) suggested that by applying generalisation and suppression processes to the dataset, the Single Pass K Means Anonymization Algorithm can get an anonymised representation of the data and Canbay et al. (2020) demonstrated a new technique called a Mondrian based utility optimization model. This model will be utilized as a trustworthy tool for outlier detection of the given dataset to promote privacy preserving of those data.

Wang et al. (2020) introduced enhanced Mondrian anonymization model based on Self-Organizing Map (SOM Mondrian). This method aims on reducing the time complexity of Anonymization method and it also helps to increase the accuracy of data mining results. Ouyang, Huang, and Xin (2008) illustrated randomization approach for privacy-preserving mining for sequential patterns. This enhanced to identify the true frequent sequences by modifying every frequent sequence of the dataset. Ong and Miri (2007) have defined a strategy to ensure privacy during testing phase of the database. So, a different approach has been introduced by utilizing a privacy filter to make decision on the results gained by the queries examined. This has been implemented with query modification and data randomization techniques.

Gokulnath et al. (2015), explained about preservation of privacy by utilizing PCA (Principal Component Analysis) based perturbation technique. This method has been employed by K-means cluster algorithm and machine learning based clustering methodology. This results in improvising privacy and performance of data.

Kaur (2017) have exhibited a new method on hybrid approach based on perturbation and suppression technique. This paper enhanced the privacy by overcoming the limitation of information loss by hiding the information of users. Charu, Aggarwal, and Philip (2004) introduced new specific approach for PPDM which does not require problem – solving algorithm because it maps the original dataset into anonymized dataset. This paper illustrates that condensation technique preserves the inter-attribute correlations of the dataset.

Borhade and Shinde (2014), proposed a new approach by utilizing Association Rule with Condensation Approach. This helps to maximize privacy and minimize the data loss by sharing the information but without disclosing the individual's identity. Patel et al., (2012), introduced a new approach to extract association rule for horizontally partitioned data. This approach is highly useful for protecting our data from intruders and also it provides authentication within the parties involved. Rahman et al. (2010) proposed a new approach called "A Real-Time Privacy-Sensitive Data Hiding Approach Based on Chaos Cryptography" which is used to protect the privacy sensitive data. This approach is based on multiple levels of data hiding models which is highly reliable for protecting the sensitive data of users.

Nowadays, Data mining has been emerged in different areas like healthcare, financial domain, banking sector, security etc., The emergence of the data mining field is also getting broader as well as the privacy issues is also increasing (Kaur, Sharma, and Taruna, 2019). Data mining plays a significant role in dealing with sensitive information, whereas it also used to preserve the data. Many banking domain deals with sensitive data like credit card or bank account details, employee's salaries, login information, and the credentials that the customer don't want to disclose (Wang et al., 2009). For daily business transaction, the banking sector deals with huge amount of confidential data. Due to the reason of dealing with large amount of sensitive data, this field is one of the main targets for data breaches (Qi & Zong, 2011; Taric & Poovammal, 2017). So, the customers never want their information to be disclosed to hackers/intruders. As a result, a model must be developed to protect an individual's information while also maintaining the information's efficiency. Privacy preserving data mining plays a vital role in satisfying these conditions (Senosi & Sibiya, 2017). There are different PPDM techniques are available like Anonymization based PPDM, Perturbation based PPDM, Randomized Response based PPDM, Condensation approach based

Related Content

How Spiritual Leadership Affects Job Satisfaction and Workplace Deviant Behavior (Study at the Regional Secretariat of the City of Palangkaraya)
Benny Hutahayan (2020). *International Journal of Applied Management Theory and Research (pp. 61-73).*
www.irma-international.org/article/how-spiritual-leadership-affects-job-satisfaction-and-workplace-deviant-behavior-study-at-the-regional-secretariat-of-the-city-of-palangkaraya/244220

The Market for CSR Practices: Issues and Challenges for the Food Sector
Irene Daskalopoulou (2016). *International Journal of Food and Beverage Manufacturing and Business Models (pp. 28-39).*
www.irma-international.org/article/the-market-for-csr-practices/145323

An ANP Approach for Prioritizing the Agile Project Management Criteria in Industry 4.0 Transition
Seda Yankand Erkan Ikl (2019). *Agile Approaches for Successfully Managing and Executing Projects in the Fourth Industrial Revolution (pp. 165-187).*
www.irma-international.org/chapter/an-anp-approach-for-prioritizing-the-agile-project-management-criteria-in-industry-40-transition/223383

Managing Uncertainties in the Project 4.0 Lifecycle
Timo Klünder, Jan Niklas Dörselnand Marion Steven (2019). *Agile Approaches for Successfully Managing and Executing Projects in the Fourth Industrial Revolution (pp. 119-140).*
www.irma-international.org/chapter/managing-uncertainties-in-the-project-40-lifecycle/223381

Women in Organizations: Challenges for Management
Basak Ucanok Tan (2019). *Handbook of Research on Contemporary Approaches in Management and Organizational Strategy (pp. 207-225).*
www.irma-international.org/chapter/women-in-organizations/217185