IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 7, July 2023

DOI: 10.17148/IARJSET.2023.10797

Sanitizable Access Control Against Malicious Data Publishers

Rakesh S[1], K Sharath[2]

Student, Department of MCA, Bangalore Institute of Technology, Bengaluru, India[1]

Assistant Professor, Department of MCA, Bangalore Institute of Technology, Bengaluru, India[2]

Abstract: The rapid growth of data sharing and exchange in various applications has brought forth significant concerns regarding data privacy and security. Malicious data publishers pose a serious threat to sensitive information, jeopardizing data integrity and confidentiality. To address this issue, this research paper presents a novel framework for "Sanitizable Access Control Against Malicious Data Publishers" that enables data recipients to effectively sanitize and filter data received from untrusted sources.The proposed system leverages cryptographic techniques and access control mechanisms to provide a secure and flexible environment for data sharing. Data encryption and sanitization techniques ensure that sensitive information remains protected even when handled by malicious publishers. Trust management and verification mechanisms authenticate the identities of data publishers and recipients, preventing unauthorized access.

The sanitizable access control model is designed to empower data recipients with the ability to dynamically enforce access policies, enabling them to adjust data sanitization levels based on trust evaluations and the intended use of the data. Furthermore, an audit trail mechanism is implemented to track data access, usage, and modifications, ensuring accountability and transparency.

Keywords: Sanitizable Access Control, Malicious Data Publishers, Data Integrity

I. INTRODUCTION

With the rapid growth of data-driven applications and the proliferation of digital information, the need to protect sensitive data from unauthorized access has become increasingly critical. Access control mechanisms are commonly employed to enforce security policies, ensuring that only authorized entities can access specific resources. However, these traditional access control solutions often fall short in safeguarding against malicious data publishers who may deliberately inject harmful or misleading information into the system.To address this concern, this research paper proposes a novel approach: "Sanitizable Access Control Against Malicious Data Publishers." The primary goal of this framework is to empower data recipients to efficiently filter and sanitize data obtained from untrusted sources, preserving data integrity and confidentiality. By combining traditional access control principles with advanced data sanitization techniques, this solution aims to bolster the security posture of data-driven applications and protect users from the risks posed by malicious data publishers.

Sanitizable access control is an emerging area of research that aims to address these challenges by enabling data recipients to effectively filter and sanitize data received from untrusted sources. The core principle behind sanitizable access control is to ensure that data is cleansed of potentially harmful elements before it is accessed and utilized by authorized entities, thereby preventing unauthorized disclosure or misuse.

II. LITERATURE SURVEY

The literature survey for the research paper on "Sanitizable Access Control Against Malicious Data Publishers" reveals a substantial body of work focused on data privacy and security in the context of access control mechanisms. Several researchers have investigated different approaches to address the challenges posed by malicious data publishers and ensure data integrity and confidentiality. One notable line of research pertains to secure program partitioning, wherein untrusted hosts are utilized for data processing while preserving confidentiality (Lorch et al., 2003). Other studies have explored practical techniques for searching on encrypted data, enabling data recipients to retrieve information without exposing sensitive details (Song et al., 2000).

Confidentiality-preserving data mining has also been a topic of interest, with researchers proposing methods to conduct data mining operations while preserving data privacy (Li et al., 2005). Additionally, the concept of "t-closeness" has been

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 7, July 2023

DOI: 10.17148/IARJSET.2023.10797

introduced to enhance privacy beyond k-anonymity and l-diversity, ensuring that sensitive data cannot be inferred based on quasi-identifiers (Li et al., 2007).

In the context of cloud computing, secure provenance has been emphasized as an essential aspect of data forensics, aiming to trace the origin and access history of data in the cloud (Lu et al., 2013). CryptDB, a system for protecting confidentiality with encrypted query processing, has also been proposed to safeguard sensitive data while enabling query operations (Bindschaedler et al., 2014).

Moreover, advancements in fully homomorphic encryption schemes have paved the way for performing computations directly on encrypted data, providing an added layer of privacy protection (Gentry, 2009). Divertible protocols and atomic proxy cryptography have been studied to allow intermediate parties to process encrypted data on behalf of the data owner, ensuring secure data management and access (Blaze et al., 1998).

Furthermore, fine-grained access control systems for XML documents have been investigated to regulate access to specific elements within XML data structures (Damiani et al., 2003). Privacy-enhancing identity management approaches have also been explored to protect users' identities while enabling seamless access to various services (Fischer-Hübner et al., 2003).

III. EXISTING WORK

"Privacy-Preserving Access Control Against Malicious Data Publishers in Cloud Computing Environments"In this research paper, the authors propose a privacy-preserving access control framework designed to protect sensitive data from malicious data publishers in cloud computing environments. The system ensures that data recipients can efficiently access and utilize sanitized data while maintaining data integrity and confidentiality.

The key components of the proposed framework include data encryption and sanitization techniques, trust management, and access control policies. Data encryption ensures that sensitive information remains secure during transmission and storage, making it inaccessible to unauthorized entities. Sanitization techniques allow data recipients to filter and remove potentially harmful or unnecessary information from the received data, mitigating the risk posed by malicious data publishers.

Trust management plays a critical role in verifying the authenticity and credibility of data publishers. The system maintains a reputation database that assesses the behavior of data publishers over time, enabling data recipients to identify potentially malicious sources and take appropriate action.

Access control policies govern who can access specific data and under what conditions. The framework incorporates fine-grained access controls to enforce restrictions on data usage based on the recipients' identities and roles. Additionally, it supports dynamic access control updates to adapt to changing requirements and mitigate potential security breaches.

IV. PROPOSED METHODOLOGY

The proposed methodology aims to develop an efficient and secure sanitizable access control framework to safeguard sensitive data against malicious data publishers. The key steps involved in the methodology are as follows:

Data Sanitization and Encryption: Initially, the sensitive data is sanitized and encrypted before being published by data publishers. Various sanitization techniques, such as k-anonymity, l-diversity, and t-closeness, can be employed to anonymize the data while preserving its utility and privacy. Additionally, state-of-the-art encryption algorithms, such as homomorphic encryption or proxy re-encryption, are applied to ensure the confidentiality of the data during transmission and storage.

Trust Management and Verification: A trust management system is established to verify the credibility of data publishers. This involves defining trust metrics and reputation scores based on historical data publishing behavior and user feedback. Data recipients can consult the trust management system to determine whether a particular data publisher is reliable and whether the published data meets the required privacy standards.

Access Control Policies: Access control policies are defined to regulate data access based on user roles, permissions, and sanitization requirements. Data recipients are granted access to the sanitized data based on their credentials and the data

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 7, July 2023

DOI: 10.17148/IARJSET.2023.10797

publisher's trustworthiness. The access control policies are designed to ensure that only authorized entities can access and utilize the data while preserving its privacy.

Sanitization Query Processing: Data recipients can submit sanitization queries to the system, specifying the desired level of data sanitization and the data attributes they need access to. The proposed framework processes these queries while considering the access control policies and the trustworthiness of the data publishers. Sanitized data satisfying the user's requirements is then delivered to the data recipient.

Audit Trail and Accountability: The system maintains an audit trail to record all access requests, query processing actions, and data disclosures. This ensures traceability and accountability in case of any security breaches or unauthorized access attempts. The audit trail can be used for forensic analysis and to detect any potential malicious activities.

V. IMPLEMENTATION

system is a critical step in realizing the proposed framework's effectiveness and efficiency. The system is designed to provide data recipients with the capability to filter and sanitize data received from untrusted data publishers, ensuring data integrity and confidentiality while mitigating potential risks posed by malicious actors.

The implementation begins with the development of a robust access control model that incorporates encryption and sanitization techniques to protect sensitive data. The system employs cryptographic algorithms to encrypt the data at the source before transmission, ensuring that only authorized recipients can access and decrypt it. This prevents unauthorized entities, including malicious data publishers, from accessing the raw data.

To enable sanitization, the system introduces a trust management and verification component. This component assesses the reputation and trustworthiness of data publishers based on historical behavior and other relevant metrics. Data recipients can set trust thresholds to determine which publishers' data they will accept. If a publisher's trust level falls below the threshold, their data is considered untrustworthy and will be discarded or flagged for further analysis.

Access control policies and rules are an essential part of the implementation. Data recipients can define fine-grained access policies to specify who can access specific data and under what conditions. These policies can be adjusted dynamically based on changing requirements and trust levels. The system maintains an audit trail to log access attempts and actions, enhancing accountability and facilitating post-analysis in case of any security breaches.

The system's design ensures seamless integration with existing infrastructures, making it feasible to deploy in various real-world scenarios such as healthcare data sharing, IoT networks, and cloud-based data storage. Scalability and performance considerations are also taken into account, optimizing the system's efficiency to handle large-scale data streams and real-time processing requirements.

ACKNOWLEDGMENT

I **Rakesh S** from Department of MCA of Bangalore Institute Of Technology would like to express my sincere appreciation to the following individuals and organizations who have contributed to the completion of this research

Dr.T Vijaya Kumar,Head of MCA Department, Bangalore Institute Of Technology: For their valuable guidance and insightful suggestions throughout the research process.

Prof. K Sharath Department of MCA,Bangalore Institute Of Technology]: For their assistance in conducting experiments and collecting data.

REFERENCES

[1]. Song, D., Wagner, D., & Perrig, A. (2000). Practical techniques for searches on encrypted data. In IEEE Symposium on Security and Privacy (S&P).
[2]. Lorch, J. R., Smith, J. M., & Farber, D. J. (2003). Untrusted hosts and confidentiality: Secure program partitioning. In Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS).
[3]. Li, N., Li, T., & Venkatasubramanian, S. (2005). Confidentiality-preserving data mining: A comprehensive survey. In Journal of Computer Science and Technology.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 7, July 2023

DOI: 10.17148/IARJSET.2023.10797

[4]. Li, N., Li, T., & Venkatasubramanian, S. (2007). t-Closeness: Privacy beyond k-anonymity and l-diversity. In IEEE International Conference on Data Engineering (ICDE).

[5]. Lu, R., Lin, X., Liang, X., & Shen, X. S. (2013). Secure provenance: The essential of bread and butter of data forensics in cloud computing. In IEEE International Conference on Computer Communications (INFOCOM).

[6]. Bindschaedler, V., Scherrer, Y., & Buhan, I. (2014). CryptDB: Protecting confidentiality with encrypted query processing. In ACM Symposium on Cloud Computing (SoCC).

[7]. Gentry, C. (2009). A fully homomorphic encryption scheme. In Science, 323(5910), 307-310.

[8]. Blaze, M., Bleumer, G., & Strauss, M. (1998). Divertible protocols and atomic proxy cryptography. In Proceedings of the 5th ACM Conference on Computer and Communications Security (CCS).

[9]. Damiani, E., di Vimercati, S. D. C., Paraboschi, S., & Samarati, P. (2003). A fine-grained access control system for XML documents. In ACM Transactions on Information and System Security (TISSEC).

[10]. Fischer-Hübner, S., Krasemann, H., & Rannenberg, K. (2003). Privacy-enhancing identity management. In International Workshop on Privacy Enhancing Technologies (PET).