Media and Communication Research (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/mediacr.2025.060103
ISSN 2523-2584 Vol. 6 Num. 1

Research on Big Data Privacy Protection Challenges and Solution Paths

Junyu Shi

Linyi Vocational College, Linyi, Shandong, China

Keywords: Big data; privacy protection; optimization countermeasures; laws and regulations

Abstract: The advent of the era of big data has introduced a wide array of application scenarios and immense value derived from data utilization. However, it has also brought increasingly prominent issues related to privacy protection, which urgently require systematic research into their challenges and potential solutions. By analyzing the importance of privacy protection in the context of big data, the existing challenges, and feasible optimization strategies, this study aims to elucidate the primary risks associated with privacy protection from both theoretical and practical perspectives and to propose targeted solutions. Research findings indicate that security vulnerabilities in data collection, storage, and transmission, as well as the lagging development of technology and regulatory frameworks, constitute the core issues in privacy protection. Recommended optimization strategies include enhancing the application of privacy protection technologies, improving legal and regulatory systems, and fostering collaborative innovation between technological advancements and regulatory frameworks. The study underscores the necessity of advancing privacy protection through a multi-dimensional approach that integrates technology, regulation, and societal efforts. Strengthening the construction of data security technologies and legal compliance frameworks while fostering the development of a sustainable and trustworthy big data ecosystem represents a critical direction for future advancements.

1. Introduction

With the widespread application of big data technologies, the capacity of various industries to collect, store, and analyze data has significantly improved, thereby amplifying the economic and societal value of data. However, the expansion of data volume and the broadening scope of its applications have given rise to increasingly severe issues related to privacy breaches. The improper use of personal information and frequent occurrences of data misuse not only infringe upon individual rights but also profoundly undermine societal trust and the healthy development of the data ecosystem. In recent years, global data privacy protection laws and regulations have gradually been refined, and privacy protection technologies have advanced continuously. Nevertheless, numerous challenges remain at the current stage, including security risks in data collection and storage, the lag in privacy protection technologies, and the insufficiencies of legal and regulatory frameworks. Therefore, researching the challenges of big data privacy protection and exploring

potential solutions to address these issues holds significant theoretical and practical value. This paper seeks to examine the importance of privacy protection in the context of big data, analyze the pressing issues currently faced, and propose optimization strategies, aiming to provide valuable references for the theoretical research and practical development of privacy protection.

2. The Importance of Big Data Privacy Protection

2.1. The growing risk of data privacy breaches

With the rapid proliferation and extensive application of big data technologies, the phenomenon of data privacy breaches has shown an increasingly alarming trend. In modern society, vast amounts of data flow frequently across domains such as personal life, corporate operations, and government administration. However, the frequent occurrence of data breaches poses a significant threat to the security of users' personal information. From users' browsing habits to sensitive financial data, a wide range of private information faces risks of being illegally accessed, traded, and misused. While the openness and sharing of data have undoubtedly enhanced efficiency, they have also rendered information more susceptible to becoming a target for cyberattacks. Malicious actors often employ technical means to illegally infiltrate databases, stealing substantial volumes of personal information that are subsequently exploited for fraudulent, extortion-related, and other criminal activities. Privacy breaches not only infringe upon individual rights but also inflict severe damage on corporate reputations and public trust, while simultaneously challenging the stability and order of the economic and social systems. Globally, the frequent occurrence of data breach incidents has become a focal point of public concern and industry attention, underscoring the urgency and necessity of enhancing data privacy protection[1].

2.2. Impact of privacy protection on social trust

Data privacy protection serves as a crucial cornerstone for maintaining social trust and represents a key factor in establishing reliable relationships between individuals, as well as between individuals and organizations, within the context of a modern, information-driven society. The widespread utilization of personal information in commercial and societal activities necessitates that users provide a certain degree of private data to platforms or institutions. However, the absence of robust privacy protection measures or the ineffective enforcement thereof can directly undermine public trust in these platforms or organizations. When personal information is exposed to breaches or misuse, the consequences extend beyond the erosion of user confidence in specific enterprises or institutions, potentially casting doubt on entire industries and hindering the broader development of a digitalized society. Public concerns regarding insufficient privacy protection can significantly impact their willingness to participate in the digital economy and the intelligent society, thereby stifling the innovation and dissemination of big data technologies. Only when adequate privacy protection measures are in place can users feel secure in engaging with data sharing and utilization, a trust that forms the foundation for fostering the healthy growth of the digital economy. Protecting data privacy is not merely a matter of safeguarding individual rights; it is also integral to ensuring the sustainable development of enterprises and society at large[2].

3. Problems in big data privacy protection

3.1. Risk of privacy leakage in the data collection process

In the process of big data utilization, the data collection stage constitutes the first and most

critical barrier for ensuring privacy protection. However, this phase is fraught with significant risks of privacy breaches. At present, many data collection practices lack transparency, with numerous platforms and institutions employing default authorizations, concealed clauses, or unreasonable permission requirements to obtain user information. Often, users unwittingly provide excessive private data without full awareness of the scope or purpose of such collection. To maximize profits, certain organizations engage in data collection practices that go beyond what is necessary, employing technologies such as location tracking and behavioral monitoring to extract sensitive user information. This excessive data collection not only increases the likelihood of privacy breaches but also exposes users to substantial potential risks.

Unauthorized data collection activities are also alarmingly prevalent. Organizations and individuals frequently exploit illegal technical methods, such as web scraping and phishing attacks, to acquire data resources, resulting in the collection and usage of vast amounts of private data without authorization. Moreover, the absence of encryption, anonymization, and other protective measures during the data collection process further amplifies the risk of information leaks. Compounding these risks, companies and institutions often exhibit weak awareness of security management in data collection, lacking effective risk control measures and technical capabilities to mitigate the associated vulnerabilities. The complexity and opacity of data sources further create fertile ground for the misuse of data, rendering the issue of privacy breaches increasingly difficult to control. Insufficient policies and regulatory enforcement exacerbate privacy protection challenges during data collection. In many countries and regions, comprehensive regulatory frameworks have yet to be established, leaving the scope, methods, and responsibilities associated with data collection inadequately defined under legal provisions. As a result, violations of data collection norms frequently go unpunished, creating significant gaps in privacy protection and further eroding public trust in data service providers. These deficiencies not only perpetuate severe vulnerabilities in privacy safeguards but also weaken the foundation of trust necessary for the sustainable development of data-driven services and technologies.

3.2. Security risks in data storage and transmission

The storage and transmission stages of data handling represent critical areas for privacy protection; however, these processes are fraught with numerous potential security vulnerabilities. With the widespread application of big data technologies, the scale of data storage has expanded exponentially, surpassing the capacity of traditional storage technologies to address increasingly complex security demands. As a result, incidents of data breaches have become frequent. Many enterprises and institutions fail to implement adequate encryption measures when storing private data. Common issues, such as improper database configurations and overly permissive access controls, have become frequent vulnerabilities exploited by malicious actors. Once a database is compromised, the large-scale leakage of user privacy information can lead to severe negative consequences for both individuals and society at large[3].

The security challenges associated with data transmission are equally significant. During the process of data transfer over networks, information is highly susceptible to threats such as man-in-the-middle attacks and eavesdropping. Although data transmission protocols have been continuously optimized, the absence of key technologies and effective management measures still leaves data exposed to risks during transmission. For instance, in the absence of end-to-end encryption, data transmitted over networks can be illicitly intercepted or tampered with, resulting in the leakage of sensitive information or its improper utilization. Furthermore, issues such as transmission delays and data loss can compromise the integrity of the information, exacerbating the difficulty of ensuring privacy protection. The extensive adoption of cloud storage technologies has

introduced additional risks to data storage and transmission. While cloud technologies have significantly enhanced the efficiency of data storage and transfer, the inherent security vulnerabilities of cloud platforms remain a critical concern. Some cloud service providers fail to establish comprehensive data protection mechanisms, leaving them ill-equipped to address both external and internal threats effectively. This lack of robust security infrastructure substantially increases the risk of privacy data breaches during storage and transmission. Collectively, these challenges underscore the need for more sophisticated and robust solutions to safeguard privacy in the evolving landscape of data storage and transmission.

3.3. Lagging privacy protection technologies and regulations

The rapid advancement of big data technologies has exacerbated the lagging state of privacy protection technologies and related regulations. From a technical perspective, existing privacy protection methods are insufficient to address the novel threats posed by the big data environment. For instance, traditional data encryption techniques exhibit low efficiency when handling vast amounts of data. Although more advanced privacy-preserving methods, such as differential privacy and homomorphic encryption, offer promising theoretical results, they remain immature in practical applications. Performance limitations and high implementation costs hinder their widespread adoption, leaving data vulnerable to persistent risks during collection, storage, and utilization. The lack of effective technical measures underscores a critical gap in mitigating privacy risks in the era of big data. From a legal and policy standpoint, the pace of privacy regulation updates has failed to keep up with the rapid evolution of big data technologies. Many privacy protection laws in various countries and regions were formulated based on traditional data usage scenarios, rendering them inadequate to address the multifaceted challenges of the big data landscape. Emerging data usage practices, such as algorithmic decision-making and cross-border data flows, introduce new privacy risks that existing legal frameworks have yet to fully address. Additionally, certain legal provisions suffer from vague definitions and insufficient enforcement, undermining their ability to provide robust support for privacy protection.

The lack of international coordination in privacy protection regulations further exacerbates the complexity of these challenges. In the context of increasingly frequent cross-border data flows, significant discrepancies between the legal systems of different countries and regions create substantial compliance challenges for enterprises. This lack of uniformity not only heightens the difficulty of implementing effective privacy protection measures but also provides opportunities for malicious actors to exploit legal loopholes and engage in data misuse. The lag in legal frameworks and the absence of international regulatory alignment significantly complicate the privacy protection landscape, constituting one of the primary obstacles to resolving the privacy challenges posed by big data.

4. Optimization Countermeasures for Big Data Privacy Protection

4.1. Strengthening Privacy Protection Mechanisms in the Data Collection Process

Protecting user privacy during the data collection process constitutes the foundational pillar of the entire privacy protection framework, necessitating the establishment of comprehensive and multi-layered mechanisms to mitigate risks. Data collection should strictly adhere to the principles of necessity and minimization, ensuring that only essential data is collected to achieve specific usage objectives while avoiding the acquisition of unnecessary information. To this end, explicit collection standards must be established to limit the scope of sensitive information collected and to ensure that the data collection process remains open and transparent. By clearly informing users

about the specific purposes of data collection and the protective measures in place, trust in data collection practices can be significantly enhanced. Enhancing user participation in privacy management is a critical pathway for optimizing the data collection process. For instance, implementing user authorization mechanisms prior to data collection can provide users with explicit options, granting them the autonomy to decide whether to authorize access to specific types of information. The application of technical measures can further strengthen the security of this process. Privacy-preserving technologies, such as secure computation techniques, can ensure that data is neither misused nor exposed during the user authorization process. These measures collectively aim to foster a secure and trustworthy data collection environment, thereby laying a solid foundation for broader privacy protection initiatives[4].

In terms of technological security, it is recommended to embed robust privacy protection technologies at the data collection endpoint, such as real-time encryption and anonymization techniques, to prevent sensitive data from being intercepted or misused before it reaches the server. By implementing a classification and hierarchical framework for data, different levels of sensitivity in data content can be addressed through tailored collection and protection strategies. For instance, highly sensitive data should be subjected to stricter collection protocols and advanced encryption techniques to minimize the risk of potential breaches. Regulatory authorities must also intensify their oversight of data collection practices. Through compliance audits and mandatory technical requirements, regulatory bodies can ensure that data collection activities align with relevant privacy protection laws and regulations. A collaborative effort involving enterprises, users, technology, and regulatory agencies is essential for constructing a transparent and efficient privacy protection framework. This multi-stakeholder approach can effectively address the inherent risks associated with data collection processes, fostering a secure and trustworthy data ecosystem.

4.2. Improving the security of data storage and transmission

Data storage and transmission constitute critical stages in privacy protection, and enhancing the security of these processes is of paramount importance for building a robust privacy protection framework. During the storage phase, the application of advanced encryption technologies is a vital measure to ensure data security. Cutting-edge encryption algorithms should be employed to encrypt stored data, ensuring that even if the storage system is compromised, the data content remains inaccessible to unauthorized entities. Additionally, the widespread adoption of distributed storage technologies is recommended, where data is fragmented and stored across multiple locations. This approach mitigates the risk of a complete information breach caused by the compromise of a single storage point. Access control is another essential component of data storage security. It is crucial to implement stringent data access permissions, granting access to sensitive data only to authorized personnel and systems with a demonstrated need. Furthermore, the security of the access process should be reinforced through technologies such as multi-factor authentication, which adds an additional layer of protection. In the context of data storage lifecycle management, clear data retention and deletion policies must be established and strictly adhered to. For example, retaining data for excessive periods increases its exposure to security risks. Therefore, sensitive data should be promptly and securely destroyed once it has served its intended purpose, thereby reducing the risks associated with long-term storage. By integrating advanced encryption, distributed storage, strict access controls, and lifecycle management policies, the security of the data storage phase can be significantly enhanced, contributing to a comprehensive privacy protection system.

In the process of data transmission, the comprehensive implementation of end-to-end encryption technologies is essential to ensure that data remains protected from eavesdropping or tampering during transit. To further enhance the security of data transmission, encrypted communication

protocols such as HTTPS and SSL/TLS should be widely adopted, with encryption algorithms periodically updated to counteract emerging attack methods. Integrating secure authentication mechanisms, including technologies such as digital signatures, is critical for maintaining the integrity and authenticity of data during transmission, thereby preventing potential man-in-the-middle attacks or the forgery of data packets. To address the growing sophistication of cyberattacks, it is imperative to establish real-time monitoring and early warning systems that enable the timely detection and mitigation of abnormal activities within the data storage and transmission processes. For instance, artificial intelligence technologies can be employed to conduct real-time monitoring of data flows, enabling the identification of potential threats and the immediate implementation of countermeasures. Regular security assessments and vulnerability remediation efforts are equally important to ensure that storage and transmission systems remain in an optimal state of defense. By integrating advanced encryption technologies, secure access management, and real-time monitoring, a comprehensive security framework for data storage and transmission can be established. This integrated approach significantly reduces the likelihood of privacy breaches and strengthens the overall resilience of data protection mechanisms. Through these measures, the risks associated with data transmission can be effectively mitigated, contributing to a robust and secure privacy protection system.

4.3. Improving the synergistic innovation of privacy protection regulations and technologies

The inherent complexity of privacy protection issues necessitates a close integration of technology and regulatory frameworks, as both are mutually reinforcing and collectively drive the comprehensive development of privacy protection. From a technological perspective, sustained investment in the research and development of privacy-preserving technologies is essential to facilitate their application in real-world scenarios. For instance, differential privacy techniques can effectively safeguard individual privacy during data analysis, while homomorphic encryption enables computations to be performed directly on encrypted data, mitigating security risks associated with decryption. Additionally, privacy-preserving computation and federated learning technologies allow for collaborative multi-party operations without sharing raw data, thereby enhancing data utilization efficiency while simultaneously reducing the risk of privacy breaches. Standardization efforts in privacy protection technologies are equally critical. It is imperative to advance the standardization and formalization of privacy-preserving technologies to ensure consistency and reliability in their application. Industry standards and technical guidelines should be established to define clear privacy protection requirements for various scenarios, promoting the interoperability and mutual recognition of technologies across regions and industries. Furthermore, governments and industry associations play a pivotal role in driving pilot programs and demonstration applications of privacy-preserving technologies, accelerating their adoption and widespread dissemination. By combining advancements in cutting-edge technologies such as differential privacy, homomorphic encryption, and federated learning with the establishment of standardized frameworks, privacy protection efforts can achieve greater efficacy and scalability. This collaborative approach not only addresses the technical challenges but also creates a robust foundation for ensuring privacy in diverse and evolving data-driven environments[5].

From a legal perspective, the rapid updating and refinement of privacy protection regulations are essential to address gaps in existing legal frameworks that fail to account for the new challenges posed by big data in emerging scenarios. For instance, specific regulatory requirements must be developed to address privacy risks associated with cross-border data flows and novel algorithmic decision-making processes. Additionally, the enforcement of privacy laws must be strengthened by introducing stricter punitive measures for data breaches and violations, thereby enhancing the

deterrent effect of legal frameworks against malicious actions. International cooperation is another critical aspect of privacy protection. Given the inherently global nature of data flows, harmonization of privacy protection regulations across countries is increasingly urgent. Greater collaboration between nations in the areas of privacy-related laws, standards, and technologies is necessary to foster the development of a global privacy protection ecosystem. For example, bilateral or multilateral agreements could establish unified privacy standards to regulate cross-border data transfers and ensure consistency in data protection practices. Coordinated innovation necessitates the synergy of technology, law, and policy, facilitated by the establishment of interdisciplinary collaboration mechanisms that promote comprehensive governance. Enterprises, governments, and academic institutions should collectively engage in privacy protection efforts by contributing to technological research and development, legal frameworks, and the promotion of industry standards. Policy guidance and resource allocation can play pivotal roles in accelerating advancements in privacy protection. The deep integration of technology and regulation will be instrumental in building a secure, efficient, and responsible big data environment. This integrated approach will provide a robust foundation for safeguarding data privacy while simultaneously supporting the sustainable development of society and the economy. Through a combination of international cooperation, regulatory rigor, and technological innovation, a balanced and resilient privacy protection system can be realized on a global scale.

5. Conclusions

The protection of data privacy in the context of big data has emerged as a critical issue for the development of digital societies, with its effectiveness directly influencing public trust in data-driven services and the sustainability of the data ecosystem. This study examines the significance of big data privacy protection, highlighting the high risks associated with data breaches, the erosion of social trust, and the mounting pressures of legal compliance, all of which underscore the urgent need to strengthen privacy safeguards. Furthermore, the study delves into key challenges, including excessive data collection and insufficient transparency during the data acquisition process, security vulnerabilities in data storage and transmission, and the lagging progress of privacy protection technologies and regulatory frameworks. On this foundation, a series of optimization strategies are proposed. These include refining data collection mechanisms to adhere to principles of necessity and transparency, enhancing the security of data storage and transmission through advanced encryption and monitoring technologies, accelerating the development of cutting-edge privacy protection technologies, and fostering synergistic innovation between legal frameworks and technological advancements. The findings suggest that effective privacy protection requires multidimensional support encompassing technology, legal frameworks, and international collaboration. It necessitates strengthening data governance, advancing the standardization of privacy-preserving technologies, and enhancing the enforcement of policies to establish a reliable data protection system. Looking ahead, further efforts should be directed toward deepening applied research on privacy-preserving technologies and bolstering international cooperation in the harmonization of privacy-related legal and regulatory frameworks. These efforts aim to provide secure and sustainable privacy protection solutions that support the continued development of the digital society. By integrating technological innovation, regulatory advancement, and global coordination, a comprehensive and resilient framework for data privacy protection can be achieved, ensuring the long-term trust and stability of digital ecosystems.

References

[1] Ding Zhiping, Lin Kun. Research on key technologies for privacy protection in the context of big data[J]. Computer

Programming Skills and Maintenance, 2022(2):72-74.
[2] Lina Duo. Legal Supervision Challenges and Responses to Personal Privacy Protection in the Age of Big Data[J]. Regional Governance, 2024(7):0108-0110.
[3] Jin Lianzeng. Challenges and responses to privacy protection and information security in the era of big data[J].IT Manager World, 2024(4):92-94.
[4] Ma Yiqun. Research on privacy protection and encryption algorithm for big data communication[J]. Communication Power Technology, 2024, 41(5):203-205.
[5] MU Jianlai, SHI Yajing, WANG Jiurong. Research and progress of privacy protection technology in big data environment [J]. Microcomputer, 2024(6):64-66.