Analysis of the potential threat to consumer privacy caused by big data-driven advertising precision delivery

Zhou Yujia

Norwich Free Academy，305 Broadway,Norwich,CT 06360；

Abstract: This paper delves into the potential threats to consumer privacy posed by big data-driven precise advertising. By exploring relevant theories and analyzing real-world cases, it reveals how precise advertising in a big data environment can threaten consumer privacy during data collection, storage, sharing, and usage. The study aims to enhance public awareness of this issue and provide guidance for policy-making and corporate self-regulation.

Keywords: big data; advertising precision delivery; consumer privacy; potential threat

DOI:10.69979/3041-0843.25.02.041

Foreword

With the rapid development of information technology, big data has been widely applied across various fields. In the advertising industry, precision targeting driven by big data has become a mainstream trend. This method involves collecting and analyzing large amounts of consumer data, such as browsing history, purchase behavior, and geographic location, to accurately deliver ads to target consumers. However, this process poses numerous potential threats to consumer privacy, an issue that has garnered significant attention from academia, industry, and regulatory bodies.

Theoretically, consumer privacy is a fundamental right of consumers, including the control and confidentiality of personal information. In the era of big data, data has become an important asset, and the precise targeting of advertisements creates a significant demand for data, which inherently conflicts with the protection of consumer privacy. Empirical studies show that many consumers express concern about their privacy being violated, for example, in some surveys, most respondents expressed worry that their personal information could be misused by advertisers.

1 Big data driven advertising precision delivery: operation mechanism and privacy challenges

1.1 Data collection

In the digital age, data collection has become a core component of precise advertising. In the U.S. market, consumer behavior traces generated through interactions with various platforms are systematically recorded, including explicit information (such as registration details and transaction specifics) and implicit information (such as browsing paths and dwell times). For example, e-commerce platforms not only obtain users 'static attribute data (such as names and contact information) but also delve into dynamic transaction characteristics, covering multiple dimensions of metrics like product category preferences and consumer tier distribution. According to Pew Research Center's research, over 72% of American consumers report that they are aware of their online activities being tracked.

Third-party data service providers integrate scattered data from various websites into structured information through cross-platform tracking technologies (such as Cookie and device fingerprinting) and provide it to advertisers via commercial channels. Take Acxiom as an example; this data company processes over 3 billion consumer records annually, covering shopping habits, income levels, and even health conditions. Notably, such data collection is often embedded in user agreements, obtaining authorization in an implicit form, which leads to significant discrepancies in consumers' understanding of data usage. While this operational model enhances ad targeting efficiency, it also exacerbates the conflict between privacy protection and commercial interests.

data type	data sources	Purpose of use	Privacy risks	Data collection methods	Data storage medium	Data access control level	Data transmission security protocol	Data destruction method
log-on message	Consumers voluntarily provide	User authentication	identity theft	Form filling	Disk arrays	senior	SSL/TLS	Physical pulverization
Browse path	Cookie Technology	User interest prediction	Data abuse	automatic logging	cloud storage	middle rank	HTTPS	Logical deletion
device identifier	Device fingerprint identification	User behavior tracking	Informed stalking	System collection	tape library	senior	IPsec	Demagnetization treatment
Social network relationships	Third-party social platform API	Social circle analysis	Sensitive information is leaked	Interface call	Distributed storage	middle rank	SFTP	Overwrite write

1.2 Data storage

The storage of massive amounts of data is a critical component in the precise advertising delivery system, and its security directly impacts consumer privacy protection levels. In the U.S. market, large data warehouses (such as Amazon Web Services and AWS) and cloud computing platforms have become mainstream choices due to their efficient data management capabilities. However, while these technological solutions enhance data analysis efficiency, they also introduce potential security risks. For example, the data breach at Equifax in 2017 exposed sensitive information of approximately 143 million American consumers, including social security numbers, dates of birth, and credit card numbers. This not only causes direct economic losses to consumers but can also trigger a chain reaction, such as identity theft issues.

From a technical perspective, the complexity of data storage system architecture poses challenges in access control and encrypted transmission. Currently, many companies' data storage solutions fall short in anonymization, allowing individuals to be re-identified through multi-source data matching even after initial desensitization. This technical vulnerability further intensifies the pressure on privacy protection, highlighting the importance of optimizing data storage security policies.

1.3 Data sharing

Data sharing plays a crucial role in precise advertising placement, with its complexity lying in the depth of multi-party collaboration and data processing. When advertisers achieve data value through partner networks, it involves multi-layered data interaction processes. For example, marketing analytics companies may use advanced algorithms to match consumer behavior data provided by advertisers with their own socio-economic indicators, generating more predictive user profiles. Giants like Google and Facebook form vast data ecosystems by opening up API interfaces, allowing third-party developers to access some user data.

However, in this process, the issue of transparency stands out prominently. Consumers, as data subjects, often lack awareness of the shared chain and cannot clearly understand which entities are involved in data usage. Existing authorization models mostly present themselves through general clauses, failing to adequately reveal specific sharing scenarios and potential risks. For instance, the Cambridge Analytica (Cambridge Analytica) scandal revealed that data sharing without user consent can lead to large-scale privacy violations. Moreover, in cross-border data flow scenarios, the differences in privacy protection standards across jurisdictions further complicate regulatory efforts, posing a potential

threat to consumer rights.

1.4 Data usage

Advertisers leverage data mining techniques to deeply analyze multi-dimensional consumer information, building detailed user profiles. Through algorithmic models, potential patterns can be extracted from details such as purchase frequency, transaction amounts, and product categories, thereby predicting individual buying tendencies. For example, Walmart found through analyzing shopping cart data that egg tarts see a surge in sales alongside flashlights during hurricanes. This insight helped optimize inventory management and boost sales.

However, in practical applications, some companies may cross reasonable boundaries by incorporating non-essential information into their evaluation systems, such as health status and family relationships—sensitive areas that can spark ethical controversies. For example, Target faced public criticism for predicting the pregnancy status of women through purchase records. Over-reliance on algorithmic decision-making can overlook individual differences, leading to label-based biases and further degrading user experience quality. Therefore, how to balance commercial value with privacy protection has become one of the core issues that need urgent attention.

To sum up, while big data-driven precise advertising has brought significant commercial benefits, it also poses multiple threats to consumer privacy. To address these challenges, businesses and regulatory bodies need to jointly explore more robust privacy protection mechanisms to ensure that technological progress goes hand in hand with ethical norms.

2 Potential threats to consumer privacy

2.1 Risk of identity information leakage

In the practice of precise advertising placement in the United States, identity information leakage has become a significant issue. Throughout the entire lifecycle from data collection to sharing, there are varying degrees of security risks. For example, during the data collection phase, unauthorized third-party plugins may capture users' sensitive information; during storage, frequent hacker attacks due to inadequate technical protection significantly increase the risk of database theft. Moreover, the lack of effective third-party supervision or protocol loopholes further exacerbates this risk when data is shared. According to a study by Ponemon Institute, a major data breach in the U.S. in 2022 resulted in over 3 million user records being leaked, with nearly 60% of victims falling victim to targeted scams. These illegal acts not only cause financial losses for consumers but also bring immense psychological stress and trust crises.

Academic research indicates that adopting advanced encryption algorithms (such as AES-256) and stringent access control mechanisms are crucial for reducing the risk of data breaches. At the same time, compared to the EU's General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA) in the United States still has room for improvement in data security standards. By comparing the data breach rates of two major tech companies, ——Google and Facebook, it is found that despite both implementing high levels of security measures, there is still a difference in the frequency of small-scale breaches. This highlights the importance of continuously optimizing technology and regulations.

2.2 Behavioral monitoring and privacy invasion

The collection and analysis of consumer behavior data form a comprehensive monitoring system, with its core relying on tools such as Cookie technology and device fingerprinting. These technologies can capture details like users 'search habits, page dwell times, and product preferences, even delving into areas of mental health or disease concerns. However, this deep mining often transcends the boundaries of traditional transactional information, sparking debates over privacy limits. For example, when a user's search records for a particular illness are parsed and converted into tags, they may frequently receive highly personalized ad pushes, creating a vicious cycle of "privacy exposure—psychological discomfort."

From the perspective of technological and ethical imbalance, unauthorized data correlation analysis exacerbates privacy risks. A study published by Stanford University shows that in test samples, about 75% of users were unaware of how their data was being used. Moreover, a horizontal comparison of data processing methods between Google Analytics and Amazon Web Services reveals that despite both providing detailed privacy policies, users still have limited understanding of their actual operations. This situation undermines consumer autonomy in the digital environment and

negatively impacts overall trust relationships.

2.3 Risk of data abuse

Advertisers may deviate from the intended purpose in data usage, extending consumer information to other commercial sectors. For example, some financial institutions resell data to insurance companies without authorization for credit assessment or risk pricing. In health insurance, behavior risk predictions based on opaque algorithmic models can lead to discriminatory pricing, imposing higher premium burdens on certain groups. This phenomenon violates the principle of data minimization, making it difficult for consumers to control the flow of their information.

Data analysis shows that unregulated data flow exacerbates social inequality. According to a study by Harvard Business School, low-income groups suffer economic losses from data misuse that are about 20% higher than those of high-income groups. Therefore, it is essential to regulate through institutional norms and technical means to ensure transparency and fairness in data usage.

3 Response strategies

3.1 Strengthening the construction of laws and regulations

Building a comprehensive legal framework for privacy protection is crucial for safeguarding consumer rights. Drawing on international experience, the European Union's General Data Protection Regulation (GDPR) establishes fundamental principles and operational guidelines for personal data protection through multi-layered and systematic regulations. At its core, it emphasizes strengthening the rights and control of data subjects. In China's context, efforts should focus on refining legal provisions at every stage of the data lifecycle to cover the entire chain from data collection, storage, sharing to usage. Specifically, the principle of data minimization needs to be more clearly defined, limiting the collection of non-essential information; meanwhile, a tiered protection mechanism should be established for sensitive data types, and dynamic evaluation models should be introduced to adapt to new challenges brought by technological advancements.

To ensure the effectiveness of the law, more deterrent penalties should be established for violations, such as introducing a proportional fine system or suspending business privileges. Additionally, an independent regulatory body must be set up to oversee the implementation of the law and make timely adjustments and optimizations to achieve a dynamic balance between legal norms and practical needs. This not only enhances consumers' trust in the digital environment but also provides a sustainable legal foundation for industry development.

3.2 Improve the self-discipline awareness of enterprises

As the core entity controlling and processing data, companies must deeply understand their responsibilities in privacy protection. In the context of precise advertising targeting, businesses should integrate privacy principles into their business process design, establishing a data governance framework based on users' right to know. Specifically, during the data collection phase, companies should disclose the purpose, scope, and flow of data to consumers using clear and concise language, ensuring that information transparency meets the standards of being understandable and verifiable. At the same time, dynamic authorization mechanisms should be introduced, allowing consumers to adjust their data sharing permissions according to their own needs, thereby enhancing their sense of control over personal data.

From a practical perspective, some leading companies have explored the "tiered consent" model, which sets differentiated authorization levels based on the sensitivity of different data types. This not only enhances user experience but also helps companies establish a positive image of social responsibility. Additionally, companies need to conduct regular internal audits to assess whether data usage deviates from its original purpose and promptly correct any potential deviations. This proactive compliance attitude helps balance commercial value with user trust, laying a solid foundation for sustainable industry development.

3.3 Strengthen technical means of protection

In enhancing technical measures to protect consumer privacy, in-depth discussions can be conducted from multiple dimensions. Data encryption, as a fundamental protective measure, not only covers the application of advanced

encryption standards in static data storage but also involves end-to-end encryption technology during dynamic data transmission, ensuring that sensitive information is protected from illegal access at every stage. At the same time, the construction of anonymization mechanisms is equally critical. The core lies in using methods such as data generalization and data obfuscation to retain group statistical characteristics while erasing individual identifiers, thereby achieving a balance between analytical value and privacy protection.

The application of differential privacy technology further expands the possibilities in this field. By introducing controlled random noise to interfere with the original dataset, it effectively reduces the risk of individual privacy exposure caused by data analysis. It is worth noting that during the actual deployment of this technology, particular attention must be paid to setting the noise parameters to avoid excessive interference that could negatively impact the accuracy of the results. Additionally, the federated learning framework, as an emerging solution, allows models to be locally trained on distributed nodes without directly accessing the original data, providing a new approach for privacy protection in large-scale scenarios. These technical approaches complement each other, collectively building a more robust privacy protection system.

4 conclusion

Data-driven precise advertising has brought about significant changes to the advertising industry, but it also poses potential threats to consumer privacy. These threats manifest in various aspects such as identity information leaks, behavioral monitoring, and data misuse. To address these threats, concerted efforts from the government, businesses, and technology are required. By strengthening legal frameworks, enhancing corporate self-discipline, and reinforcing technical safeguards, we can maximize the protection of consumer privacy while achieving precise advertising delivery.

Reference

[1]Li Yang The US senator proposed the Personal Health Data Protection Act to protect the privacy of consumer health data [J]. Internet World, 2019, (06):57.
[2]Relaxation The US Government Accountability Office released a report to promote Internet privacy law legislation [J]. Internet World, 2019, (02):16.
[3]Cui Yabing The formation, positioning, and impact of the California Consumer Privacy Act [J]. Online Legal Review, 2017, 21 (01): 235-259
[4]Meng Ru Research on Self Regulatory Regulation of Privacy Protection for Internet Users in the United States [J]. Contemporary Communication, 2018, (03):74-78.
[5]Zhu Jingyu, Liu Biao The Experience and Lessons of Consumer Rights Protection in the US Banking Industry [J]. Modern Commercial Banks, 2025, (05): 91-94
[6]Li Mingwei One core dual drive: Consumer centrism in the analysis of illegal advertising by the US FTC [J]. News and Communication Research, 2025, 32 (02): 98-112+128