Personal Information Leakage of Internet Users: A Study on the 12306 User Data Breach Incident

Ping He[1,a]*,Chang Liu[2,b]

[a]heping@guet.edu.cn, [b]815507915@qq.com

[1]Professor of Guilin University of Electronic Technology, Guilin,China
[2]Postgraduate Student of Guilin University of Electronic Technology, Guilin, China

Abstract—The number of Internet users worldwide continues to increase, and the security of personal information on the Internet has become a global concern.This article analyzes the 12306 website user information leakage incident as an example, emphasizes the shortcomings in network security governance practices .Our research findings indicate that the 12306 website user data breach was caused by hackers launching a credential stuffing attack on its database.This article analyzes the process of hackers stealing user information from a technical perspective and analyzes the risks of privacy policy in 12306 website and its actual implementation. We have outlined the process of such an attack to provide a comprehensive understanding and proposes some constructive suggestions from the perspectives of both enterprises and users to improve network security.

Keywords-Network security; Information safety; Privacy Protection; Internet

1. Introduction

The importance of personal information is becoming increasingly prominent in today's society, With the development of technology and the acceleration of digitization, people's lives are becoming more and more intertwined with the digital world. In this process, personal information plays a crucial role. It is not only a symbol of people's identities, but also the foundation of their actions and interactions in the digital world.A survey covering 25 countries shows that compared with the results of a survey a year ago, more than half of the respondents are more concerned about their online privacy[1].

Datareportal, Meltwater, and We Are Social collaborated to produce the 《Digital 2023: Global Overview Report》, which states that there are 5.16 billion Internet users in the world today, which means that 64.4% of the world's total population is now online. There are currently 4.76 billion social media users worldwide, accounting for nearly 60% of the world's population[2].China Internet Network Information Center (CNNIC) released the 52nd 《Statistical Report on Internet Development in China》on August 28, 2023 in Beijing.The report shows that as of June 2023, the number of Internet users in China reached 1.079 billion, an increase of 11.09 million compared to December 2022, and the Internet penetration rate reached 76.4%[2].There is a lot of evidence that more and more people consume and socialize from web browsers and mobile applications, and a large amount of digital funds and personal information are flooded in the network virtual platform. The Fortinet 2023 Global Operational

PMBDA 2023, December 15-17, Nanjing, People's Republic of China

DOI 10.4108/eai.15-12-2023.2345345

Technology and Cybersecurity Situation Research Report was recently released.The report is based on a special survey conducted by a third-party authoritative research institution on 570 operational technology (OT) professionals worldwide.The results show that with the continuous integration of IT/OT networks, the attack momentum of cybercriminals targeting the OT environment has increased, although the number of organizations that have not suffered from cyber threat intrusion has increased significantly year-on-year (from 6% in 2022 to 25% in 2023), but there is still huge room for improvement in the organization's cybersecurity protection capabilities.The data shows that three-quarters (75%) of the respondents said they had suffered at least one threat intrusion last year[4].Therefore, in the digital age, the importance of personal information is self-evident. People need to recognize the importance of protecting personal information and take active measures to safeguard their privacy and security. Kevin Macnish and Jeroen van der Ham argue in their paper that current ethical oversight methods for cybersecurity ethics are insufficient[5].At the same time, governments and various sectors of society should strengthen the laws, regulations, and oversight of personal information protection to ensure the lawful use and secure storage of personal information. Only in this way can we better enjoy the convenience and opportunities brought by digitalization while also safeguarding individual rights and social stability.

The academic community has conducted extensive research on personal privacy issues in the network.research in academia on personal privacy is mainly focused on ten directions, namely Medicine and Dentistry, Social Sciences, Computer Science, Business, Management and Accounting, Engineering, Psychology, Nursing and Health Professions, Decision Sciences, Environmental Science, and Energy. Each direction has its own unique focus of research.

- Medicine and Dentistry：The total number of research papers is 25992; The main research directions are 1. Data privacy protection in the medical process. 2. Privacy protection in medical devices.

- Social Sciences：The total number of papers is 23142, mainly focusing on the relationship between privacy policies and users.

- Computer Science：The total number of papers is 21518, and the main research direction is to explore the development of privacy protection programs from a technical perspective.

- Business, Management and Accounting：There are a total of 9936 papers, mainly focusing on the impact of personal privacy protection on business behavior, such as consumer purchasing intentions and the quality of online app services.

- Engineering：The total number of papers is 8691, and the research direction is divided into two aspects: 1. Analyzing the impact of privacy policies on ethics from the perspective of engineering ethics. 2. Analyze current privacy protection technologies.

- Psychology：There are a total of 7113 papers, with research directions divided into 1. Analyzing the impact of privacy clauses on users from a psychological perspective. 2. Privacy protection during psychotherapy.

- Nursing and Health Professions：The total number of papers is 6735, with a research focus on privacy protection of patient information.

- Decision Sciences： The total number of papers is 4513, with a research focus on analyzing online privacy policies.
- Environmental Science： The total number of papers is 4373, with a research focus on the impact of privacy policies on environmental policies.
- Energy： The total number of papers is 3199, and the research direction is the relationship between users' perception of privacy and energy use.

Through these studies, we found that with the widespread use of intelligent software, there are significant vulnerabilities in users' personal information preservation.Many of the privacy risks that a user's privacy is exposed to stem from the authentication and management mechanisms of published information[6].For example, in March 2018, the 12306 website was hacked by hackers, resulting in the leakage of user personal information. The hackers used technical means to enter the database of the 12306 website and steal user personal information. This incident affected many users who used the 12306 website for ticket purchases, potentially leading to their personal information being used by lawbreakers for their own purposes and causing losses. It also raised users' concerns about the security of their personal information. This article aims to analyze this incident and read the privacy policy of 12306 in detail to provide protection measures for the privacy of Internet users.

In the second part of this article, we describe the causes and consequences of the 12306 website user information leakage incident.

In the third part, we analyze from a technical perspective the reasons behind this user information leakage incident, how hackers stole user information through password cracking. We focus on analyzing the privacy policy of the 12306 website, by comparing the user privacy policy of the 12306 website with its actual implementation, we analyze the risks of user privacy protection in the 12306 website.

Through the analysis of the 12306 incident, we provide some suggestions for improving user information protection and self-protection of users on Internet websites in the fourth part of this article, in order to enhance user information security.

2.12306 Software Information Leakage Event

2.1.Course of Events

A netizen posted an article titled "A large amount of 12306 user data is crazily spreading on the Internet, including user accounts, plaintext passwords, ID cards, emails, etc. (the leakage route is currently unknown)" on the vulnerability platform "WooYun" in December 2014. The netizen claimed that hackers have obtained a large amount of user information from 12306 and are circulating and trading it within certain hacker groups. It is understood that the sample data file contains a total of 131,653 records and has a file size of 14MB. According to the vulnerability details published on "WooYun," it is currently impossible to confirm whether the leakage was from the official 12306 platform or a third-party ticketing platform. It is hoped that the official authorities will immediately intervene in the investigation and notify the leaked users to change their passwords.

This vulnerability report shows a high level of harm and the type of vulnerability is "large-scale leakage of user data." This means that sensitive information such as user accounts, plaintext passwords, ID cards, and emails of registered 12306 users may have been leaked. The vulnerability has been reported to the National Internet Emergency Center for handling. The director of the research department at the network security company Qihoo 360 stated in an interview that after careful analysis, the batch of 131,653 12306 user data is authentic.

2.2.The Result of the Event

On December 15, 2014, the Chinese railway public security organ arrested two suspect in the leak of 12306 website, and the public security confirmed that the suspect illegally obtained 12306 user data by "hitting the database".According to the 360 Internet Security Center, the reason why 12306 was hacked is likely due to a vulnerability in its mobile app. The 12306 mobile app login interface can be maliciously exploited by hackers to attempt unlimited attempts to hack the database.

2.3.The Impact of the Event

After the news was released, it attracted the attention of numerous netizens. Many of them expressed their intention to quickly log in to the 12306 website to change their passwords to prevent their personal information from being used for illegal purposes.

Some netizens believe that because their 12306 website booking information contains phone numbers, ID cards, and other personal information, not only could this information be sold to organizations for sending spam messages and making sales calls, but more seriously, many people's website information also includes information left for booking tickets for family, friends, and colleagues. If used by criminals, there is a risk of falling victim to telecommunications fraud.

3.Analysis Of The 12306 Information Leakage Incident

The fundamental reason why the 12306 website was hacked successfully by "crashing the database" is that its account security system has defects.The technical staff of the “Bu Tian” Vulnerability Platform found that there is a vulnerability in the login interface of the 12306 mobile APP, and hackers can easily bypass its account security protection measures and attempt automatic login indefinitely.More than 130,000 user passwords of 12306 that were circulated online were obtained by hackers through "crashing the database". Such a huge number of login requests, 12306 did not detect and block in time.The 360 Security Center has reported the vulnerability of the 12306 mobile app to the China Railway Customer Service Center for repair, but the security crisis caused by the 12306 data leak is still continuing to ferment.

According to the monitoring of 360 Security Guard, there have been trojans disguised as 12306 data packets on the Internet, which are spread in online storage and chat group sharing. It is recommended that netizens should not easily download them.In addition, the 12306 has been exposed to vulnerabilities repeatedly. Users who have bought train tickets should take the tickets as soon as possible, beware of malicious refunds of accounts, and also be vigilant against incoming SMS messages from fake railway staff, so as to avoid fraud.

3.1.Credential stuffing attacks

A.Credential stuffing attack is a primary attack method targeting personal users on the internet. When personal users register identity verification information on different websites, they often use static passwords instead of complex dynamic passwords. This greatly benefits users by reducing the burden of remembering multiple passwords for the numerous websites and apps, thus providing convenience in usage. However, this practice also brings the hidden danger of information collision. Once hackers obtain user accounts and passwords from websites with low security, they can attempt to use the same credentials to attack websites with higher security, which is known as a Database Breach attack.

Database intrusion mainly consists of three steps: "database extraction", "Data Scrubbing" and "Database breach".

- "Database extraction": Dragging the database refers to hackers using various social engineering and technical means to illegally obtain sensitive information from the database. Generally, this sensitive information includes user account information such as usernames, passwords; identity information such as real names, identification numbers; contact information such as email addresses, phone numbers, addresses, etc.

- "Data Scrubbing": After dragging the database and obtaining a large amount of user data, hackers will use a series of technical means and the black market to monetize valuable user data, which is often referred to as "scrubbing". (1) Virtual currency in user accounts, game accounts, equipment, etc. can be monetized through trading, also known as "account theft". (2) Financial account information such as Alipay, online banking, credit cards, stock accounts and passwords can be used for financial crimes and fraud. (3) Finally, some classified user information, such as students, workers, bosses, etc., is mostly used for sending advertising, spam messages, and e-commerce marketing. There are also specialized advertising companies that pay money to purchase these categorized information.

- "Database breach": Database breach is when hackers collect leaked username and password information from the internet, generate corresponding dictionary tables, and then try to log in to other websites in bulk. Many users use the same account passwords across different websites, so hackers can use the obtained A website account information to attempt to log in to website B, which can be understood as a database breach attack.

The process of hackers stealing users is shown in “Fig. 1”.

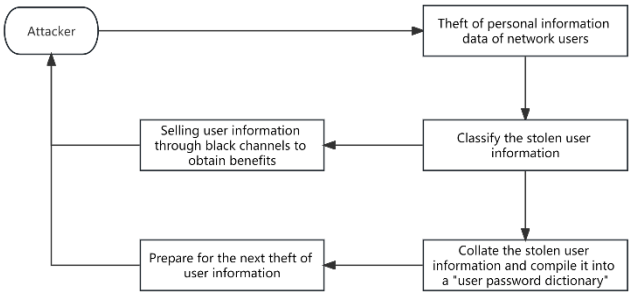

Figure 1 The process of hacking user information

3.2. Analysis of the 12306 Privacy Agreement

The privacy seal programs provide a set of guidelines and

enforcement mechanisms to assure that online services abide by their own privacy policy[7].Petra Saskia Bayerl and Gabriele Jacobs believe that privacy is now the joint responsibility of both first and third parties[8]. Through analysis of user data leakage incidents, we believe that analysis of website user privacy agreements is necessary.We analyzed the protocols for collecting, sharing, and saving user information in the 12306 software.

1) Collection of personal information and sensitive personal information

"Personal information" refers to various information recorded through electronic or other means that can identify the identity of a specific natural person or reflect the activities of a specific natural person or in combination with other information.

"Sensitive personal information" refers to personal information that, once disclosed, illegally provided, or abused, may endanger personal or property safety, easily damage personal reputation and physical and mental health, or lead to discriminatory treatment.

The 12306 website will collect the above two types of information from users after they sign the privacy agreement. According to relevant laws, regulations and national standards of the People's Republic of China, the 12306 website may collect users' personal information without their authorization or consent in any of the following circumstances:

- Information related to the fulfillment of obligations under laws and regulations with the 12306 website;
- Information directly related to the national security and defense security of the People's Republic of China;
- Information directly related to public safety, public health, and major public interests;
- Information directly related to criminal investigation, prosecution, trial, and execution of judgment.
- Information collected to protect the lives, property, and other significant legal rights and interests of users or other individuals when it is difficult to obtain the user's consent;
- Personal information disclosed to the public at random by users;
- Information required for signing and fulfilling contracts as requested by users;
- Personal information collected from publicly disclosed information through legal channels such as news reports or government information releases;
- Maintain the information required for the safe and stable operation of public relations

2) Sharing of personal information

12306 will share user personal information with suppliers and third-party merchants of various functional goods or technical services in the privacy policy. It is necessary to be vigilant:App users may suffer loss of personal privacy due to security breaches or common data sharing

practices between app developers and third parties[9].This sharing is conditional on the following:

- 12306 will share user information with third parties after signing confidentiality agreements in accordance with the terms of the confidentiality agreements.
- If the information sharing object is added or changed, 12306 needs to obtain user permission before continuing to share user information.
- Share information with the consent of the user or as required by policies and laws.

3) Preservation of personal information

The 12306 website established an information security team and established an information security assurance system suitable for the business development of the 12306 website.Access control mechanisms are deployed to ensure that only authorized personnel can access personal information.Users are also required to properly maintain their personal information themselves.

Through our actual login to the 12306 website and the 12306 app, we found that the current protection of user privacy in the 12306 website is implemented according to its website's privacy policy. However, based on this information leakage incident, it can be found that the third-party management of user information sharing in the 12306 website is not strict, which will cause users to be unable to specifically understand how their information is shared with third parties, what information is shared with third parties, and what responsibility third parties have for the protection of user information. The above risks will still lead to a greater risk of user information protection, potentially leading to another information leak in the future.

4.The Risk Of Personal Information Leakage On The 12306 Website And Suggestions

Like any large online transaction and service website, handling the security and privacy of a large number of users' personal information is an important task.Although the 12306 website has made considerable progress in focusing on database security, there are still some common cybersecurity issues and challenges.In response to the above information security issues, we have provided some practical suggestions to the website and users.

4.1.Some problems in the protection of personal information on the 12306 website

In response to this information leakage incident, we have summarized and found that the 12306 website currently has the following problems:

- Privacy Policy: Every user needs to read and agree to a privacy policy when registering for using the 12306 website.This policy stipulates how the website collects, stores, uses, and shares your personal data.The problem often lies in many users not fully understanding these terms, which may lead them to waive certain privacy rights without their knowledge.
- Insecure database: If the database security of the website is insufficient, attackers may use known or unknown vulnerabilities to steal user data.

- Website data encryption: For transactional websites, whether to encrypt sensitive information such as bank information and personal phone numbers is also a major issue. When attacked, unencrypted data is more likely to be used maliciously.

- Third-party sharing: whether the user's personal information is used to share with third parties is also a key issue, even if it is anonymized data, there is a risk of being reverse-analyzed

4.2.Suggestions on user information security

Based on the above research, this article proposes several targeted suggestions for Internet companies and their users on personal information security.

1) Suggestions for the company

a) Logging and auditing of user account usage behavior.The system server side should conduct detailed logging and auditing of user behavior based on account information, and conduct periodic audits (with relatively short time intervals) through the logging of the aforementioned factors to detect issues such as misuse and malicious use of user accounts and handle them as early as possible.

b) Detection, filtering and blocking of malicious user traffic.The system server side should deploy IDS intrusion detection system, IPS intrusion prevention system, firewall and other equipment or deploy current efficient and popular UTM equipment to detect and protect against various attack methods used by malicious users, focusing on filtering malicious traffic, burst traffic, etc.

c) Filtering and processing of abnormal application requests.The server side of the system, especially the database server side, should configure and add filtering and processing modules for user-specific application requests to avoid being subjected to popular SQL injection attacks due to the database's own vulnerabilities that have not been patched in a timely manner.

d) Load balancing and load protection mechanism.The system is facing a huge amount of service traffic, and server-side devices basically require multiple servers for business sharing in order to improve performance and avoid processing bottlenecks. Therefore, it is necessary to adopt reasonable load balancing and load protection mechanisms.Effectively sharing the traffic of each server can be achieved through load balancing methods such as Round Robin and LRU.The load protection mechanism requires real-time evaluation of each server's CPU resources, memory resources, etc. If the threshold is exceeded, overload protection will be immediately implemented to ensure the security of the server itself.

e) Management standardization.The system has complex functions, sensitive business data, and a high level of confidentiality, and the requirements for different administrators' permissions and roles vary. ethical decision-making is an unavoidable aspect of the everyday practices of cybersecurity and ICT professionals[10]. To ensure security management and avoid security issues in internal management, the following requirements are recommended:

- Strictly divide the roles of management personnel and their corresponding permissions to avoid one-person monopoly and potential safety hazards;

- Manage the physical conditions of the server room to avoid electronic leakage and malfunctions caused by static electricity;

- Server administrator's account/password management should be well done, requiring the use of strong passwords to avoid internal personnel theft;
- Implement port minimization management for servers to avoid unnecessary open ports and vulnerabilities that can be scanned by internal personnel, resulting in internal attacks;
- Manage the log and patch management of server system software and application software to facilitate auditing and avoid internal attacks due to security vulnerabilities;
- Strictly divide the security domain of the server according to the confidentiality level requirements of business and data to avoid information leakage.

f) Self-excavation and protection of website vulnerabilities.Using vulnerability scanning and mining equipment, conduct periodic scans of various servers on the intranet, and promptly patch the risks and vulnerabilities found during the scans to achieve the goal of self-repairing the vulnerabilities before they are used by hackers.

2) Suggestions for users

For users, actively setting and managing their own passwords can maximize the protection of their personal information security and property security.Therefore, we offer the following suggestions for users:

- Avoid using the same password everywhere, and set up separate accounts and passwords for different websites.
- Use a strong password in the form of "letters + numbers + special characters" as much as possible, with a length of at least 8 characters.
- It is necessary to change each set of passwords regularly, and we recommend changing them every two months.

5. Conclusions

After the 12306 user information leak incident, the importance of personal information security in the Internet era has once again been highlighted.This incident should also arouse our concern and attention to the protection of personal information security.In the Internet era, the protection of personal information is particularly important, as this information may involve the privacy, property security, and even national security of users.Therefore, we must take effective measures to protect the privacy and security of users and prevent similar incidents from happening again.

References

[1] Akram, W, & Kumar, R. (2018). A study on positive and negative effects of social media on society. International Journal of Computer Sciences and Engineering, 5(10):351–354. https://doi.org/10.26438/ijcse/v5i10.351354.

[2] Datareportal, Meltwater, and We Are Social collaborated ,2023.“DIGITAL 2023: GLOBAL OVERVIEW REPORT, ” https://datareportal.com/reports/digital-2023-global-overview-report

[3] China Internet Network Information Center,2023.“Statistical Report on Internet Development in China， ” https://www.cnnic.net.cn/n4/2023/0828/c88-10829.html

[4] Fortinet, 2023.“Research Report on Operational Technology and Cybersecurity Situation in 2023,” https://www.fortinet.com/content/dam/fortinet/assets/reports/report-state-ot-cybersecurity.pdf

[5] Kevin Macnish , Jeroen van der Ham. (2020) Ethics in cybersecurity research and practice.Technology in Society, 63:2 https://doi.org/10.1016/j.techsoc.2020.101382

[6] Constantinos Patsakis , Athanasios Zigomitros, Achilleas Papageorgiou ,Edgar Galván-López . (2014)Distributing privacy policies over multimedia content across multiple online social networks.Computer Networks75 : 531–543 https://doi.org/10.1016/j.comnet.2014.08.023

[7] Hui Na Chua , Anthony Herbland , Siew Fan Wong , Younghoon Chang .(2017) Compliance to personal data protection principles: A study of how organizations frame privacy policy notices. Telematics and Informatics 34:157–170 https://doi.org/10.1016/j.tele.2017.01.008

[8] Petra Saskia Bayerl , Gabriele Jacobs .(2022)Who is responsible for customers’ privacy? Effects of first versus third party handling of privacy contracts on continuance intentions Technological Forecasting & Social Change, 185 :1 https://doi.org/10.1016/j.techfore.2022.122039

[9] Lisa Parkera,, Vanessa Halterb, Tanya Karliychukc, Quinn Grundya. (2019) How private is your mental health app data? An empirical study of mental health app privacy policies and practices.International Journal of Law and Psychiatry 64:198–204, https://doi.org/10.1016/j.ijlp.2019.04.002

[10] Paul Formosa , Michael Wilson , Deborah Richards .(2021)A principlist framework for cybersecurity ethics.computers & security,109:11 https://doi.org/10.1016/j.cose.2021.102382