Hindawi
Security and Communication Networks
Volume 2021, Article ID 5518220, 11 pages
https://doi.org/10.1155/2021/5518220

Research Article

An Intelligent Detection Method of Personal Privacy Disclosure for Social Networks

Haiyan Kang[1], **Yanhang Xiao**[1,2] **and Jie Yin**[1,3]

[1]*Department of Information Security, Beijing Information Science and Technology University, Beijing 100192, China*
[2]*Faculty of Engineering, University of New South Wales, Sydney 2052, Australia*
[3]*School of Computer Science, University of Sydney, Sydney 2006, Australia*

Correspondence should be addressed to Haiyan Kang; kanghaiyan@126.com

Received 5 February 2021; Revised 15 March 2021; Accepted 31 March 2021; Published 24 April 2021

Academic Editor: Hao Peng

With the increase of the number of users in the current social network platform (taking WeChat as an example), personal privacy security issues are important. This paper proposes an intelligent detection method for personal privacy disclosure in social networks. Firstly, we propose and construct the eigenvalue in social platform. Secondly, by calculating the value of user account assets, we can obtain the eigenvalue to calculate the possibility of threat occurrence and the impact of threat. Thirdly, we analyse the situation that the user may leak the privacy information and make a score. Finally, SVM algorithm is used to classify the results, and some suggestions for warning and modification are put forward. Experiments show that this intelligent detection method can effectively analyse the privacy leakage of individual users.

1. Introduction

Today's society is developing rapidly, and with the popularity of smartphones, the amount of private information they generate is increasing. With the occurrence of "PRISM," Facebook user personal information leaked, and other incidents, the issue of private security has begun to attract people's attention.

In recent years, WeChat is one of the most popular apps in China, and as of June 2019, the number of monthly active accounts on WeChat reached 1.13 billion. The huge number of users will contain a large amount of user privacy information. However, most users do not have relevant expertise or neglect information management. Therefore, when using WeChat, they do not pay attention to the protection of private information. After investigation and analysis, disclosure of account passwords, the addition of friend settings, location information, etc. during the chat process will pose a threat to user data if they are leaked and may cause economic loss or even personal harm. Because of the above problems, there have been related studies. Reference [1] conducted a comprehensive evaluation of the apps in the mobile app market but did not consider the risk of social platforms. Reference [2] aimed at Facebook to collect a large amount of user data and analyse them using a questionnaire, but the number of users on WeChat in China far exceeds Facebook. Reference [3] summarized the abnormal account detection scheme based on the four aspects of behaviour characteristics, content, graphs, and unsupervised learning. The detection scheme is rich, but the feature values involved are relatively small, making the data analysis insufficient and not targeting personal social accounts to make appropriate adjustments. Reference [4] fully considered the issue of trajectory privacy leakage and protected it with a prefix tree, but it was not enough to consider all aspects of personal social account security. Reference [5] made a formal description of the malicious use of the address book matching function and made corresponding protection measures, but it was also considered incomplete. In [6], a lot of malicious programs and risk programs on the Android side were analysed in depth, but only the leakage of information through phone calls and text messages was analysed. It did not analyse the

social platform and failed to give a more intuitive evaluation system. In [7], a high scientific and rigorous static analysis, dynamic analysis, and network data model were used for multidimensional analysis. However, only 30 apps were tested, and it was concluded that the current application market software leaks user privacy. No risk assessment is performed for each user. In [8], the risk analysis was based on association rules and game theory, but the selected feature values were few and were not targeted at social platforms. Moreover, it is learned that the privacy leak detection systems in the current environment are oriented to companies and enterprises and are not suitable for analysing individual users.

The innovations proposed in this article are as follows: (1) Through the investigation and analysis of WeChat, eight characteristic items in social networks are proposed and constructed, which are account passwords during chats, WeChat wallet consumption records (not friends), and WeChat wallet transfer records (friends), Moments settings of strangers, settings for nearby people, settings for adding friends, Moments settings of friends, and information acquisition of mini-program. The intelligent detection system uses these filtered feature items to calculate the risk value more efficiently and accurately. (2) Use the operations of asset identification, threat identification, and vulnerability analysis to calculate the comprehensive threat value. (3) An intelligent detection method based on SVM (Support Vector Machine) is proposed to divide the data more accurately. (4) After investigation, most of the detection software with similar functions today is oriented to enterprises, and this system is a rare intelligent detection system for individual users on the market.

2. Principles of Intelligent Detection System

The intelligent detection method of personal privacy leakage for social networks proposed in this article is always for users. There is a risk of personal privacy leakage. By obtaining user WeChat settings, asset identification, threat identification, and vulnerability analysis are performed, and the matrix is compared to obtain security. For event risk value, calculate information leakage risk coefficient according to weight. Reference [9–12] pointed out that machine learning has been widely applied in the fields of healthcare, cybersecurity, etc. due to its powerful data mining capabilities, where SVM is one of the most popular machine learning algorithms; therefore use SVM algorithm to divide information leakage risk coefficient and get a final evaluation.

2.1. Risk: Risk Is the Effect of Uncertainty on a Goal. The risks explored in this article refer to the risks of information security breaches, human or natural threats, and the use of vulnerabilities in information systems and their management systems to cause security incidents and their impact on organizations. In the current environment of high information transparency, private information cannot be in a state of zero risks [8].

2.2. Asset Identification

2.2.1. Assets and Their Value. Assets refer to any information or resources that are valuable to the unit. The value of assets does not refer to the economic value of the information system but is closely related to the business work of the organization. Asset value is the importance and sensitivity of assets and the main content of asset identification.

2.2.2. Asset Identification. Asset identification includes two steps: "asset classification" and "asset assignment." This article explores the classification of application software. Based on asset classification, further semiqualitative and semiquantitative analysis of assets is performed; that is, asset valuation is performed, to have a scientific and rational understanding of asset value. Assets are broken down into three security attribute assignments: "confidentiality assignment," "integrity assignment," and "availability assignment."

2.2.3. Confidentiality. It is the feature that prevents the information from being leaked to unauthorized individuals, entities, processes, or makes it useless.

2.2.4. Integrity. It protects the accuracy and completeness of information and processing methods.

2.2.5. Usability. It is a feature that can be accessed and used by authorized entities once they are needed [13].

2.3. Threat Identification

(1) Threat: Potential cause of an accident that may cause damage to assets or units.

(2) Threat identification: Referring to the process of analysing the potential cause of an accident. Threat identification is divided into "threat classification" and "threat assignment" [13].

2.4. Vulnerability Analysis

(1) Vulnerability: Weakness in assets or assets that can be threatened. Compared with threats, threats are the external cause of risk, and vulnerability is the internal cause of risk. The two together form a risk.

(2) Vulnerability identification: Referring to the process of analysing and measuring the weak links of assets that may be threatened to use [13].

2.5. Basic Introduction of SVM Algorithm. SVM refers to support vector machine, which is a common method of discrimination. In the field of machine learning, it is a supervised learning model, which is usually used for pattern recognition, classification, and regression analysis.

The main idea of SVM can be summarized as two points:

(1) It analyses linearly separable cases. For linearly inseparable cases, by using a nonlinear mapping algorithm, a linearly inseparable sample from a low-dimensional input space is transformed into a high-dimensional feature space to make it linearly separable. It is possible to perform a linear analysis of the nonlinear features of the sample using a linear algorithm in the feature space.

(2) It constructs the optimal hyperplane in the feature space based on the structural risk minimization theory, so that the learner is globally optimized, and the expectations in the entire sample space meet a certain upper bound with a certain probability [14].

3. Intelligent Detection Model Design

3.1. Basic Architecture of Intelligent Detection Model. This intelligent detection model is divided into a data source layer, an analysis layer, and a calculation layer, as shown in Figure 1. Among them, after the user source of WeChat data is obtained by the data source layer, eight characteristic values are selected for analysis and calculation; the analysis layer performs asset identification in turn for the characteristic values, and threat calculation and vulnerability analysis, respectively, obtain calculation tables. Asset identification selects three security attributes of asset confidentiality, integrity, and availability, calculates the asset value, and divides the asset value into five levels to obtain a quantitative asset value table. Threat identification is to classify threats into five levels based on the frequency of threats to obtain a table of the frequency of threats. Vulnerability analysis is to calculate the fragility property calculation table by calculating the basic measurement group, time measurement group, and environmental measurement group in turn; at the calculation layer, the three calculation tables in the analysis phase are combined with the security event to compare the two-dimensional matrix table to obtain from each eigenvalue's data the risk value of the security event.

Then the sum of the weight values of each risk value is used to obtain the risk value, and the risk value is brought into the corresponding SVM classifier to obtain the final result.

3.2. Eigenvalue Construction. Based on the investigation and analysis of WeChat, we selected the following conditions as the eigenvalues. The intelligent detection system uses these filtered feature items to calculate the risk value more efficiently and accurately:

3.2.1. Account Password in the Chat Process. The account and password are directly mentioned during the chat. If the chat history is stolen, the account and password information is leaked, and the entire account will be lost, with more illegal acts.

3.2.2. WeChat Wallet Consumption Records (Non-Friends). They require money to communicate with each other without knowing too much about the identity of the other party, have lack of security protection, and may cause economic losses.

3.2.3. WeChat Wallet Transfer Records (Friends). The transfer security between friends is higher than the transfer between non-friends, but if the identity of the friend is impersonated, the identity of the transfer counterparty is unknown, so even the transfers between friends will be at risk.

3.2.4. Setting up a Circle of Strangers. The setting of a circle of strangers is divided into invisible to strangers, ten photos visible to strangers, and unlimited. If the attacker continuously obtains the user circle information for a long time, the stranger can see that the ten photos are not much different from unlimited, which will cause a large amount of information leakage for the user.

3.2.5. Settings for Nearby People. If the nearby people are not closed, the real-time location of the user will be exposed and used by criminals.

3.2.6. Add Friend Settings. The related settings include whether you need to verify when adding as a friend. The way to search for users is divided into WeChat, mobile phone number, and QQ number, in addition to business card. Too many permissions in this regard will increase the possibility of being disturbed by strangers.

(7) Location of Moments: The attacker can further commit a crime based on the obtained positioning information, causing the user's personal safety to be threatened

(8) Mini-Program Information Acquisition: Mini-programs usually obtain user information. If the mini-programs are used by criminals, arbitrating user information will lead to user information leakage.

3.3. Asset Identification. Assets have security attributes such as confidentiality, integrity, and availability, which reflect the characteristics of the asset in different aspects. By quantifying the three security attributes, one can calculate a value that reflects the asset [15].

$$\text{AssetValue} = \text{INT}\left[\log_2 (2\text{Conf} + 2\text{Int} + 2\text{Avail})\right]. \quad (1)$$

Among them, Conf represents confidentiality assignment; *Int* represents integrity assignment; *Avail* represents availability assignment; INT represents rounding processing and rounding. The three security attributes are divided into 5 levels. The higher the level, the greater the impact on assets. There are 5 levels of corresponding security attributes, and the level of asset value is also divided into 5 levels. The greater the level is, the more important the asset is.

It can be seen from Table 1 that the disclosure of the account password during the chat process will lead to the loss of the entire account information, so its three

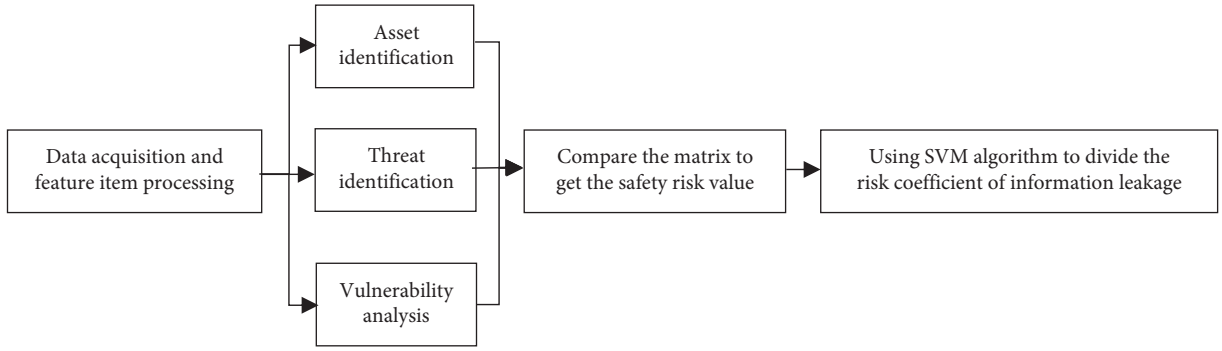

Figure 1: Schematic diagram of detection system structure.

Table 1: Quantification of asset identification.

Characteristic values	Security attributes			
	Conf	Int	Avail	Asset value
f1: Account password revealed during chat	5	5	5	5
f2: WeChat wallet consumption records (non-friends)	3	3	1	4
f3: WeChat wallet transfer records (friends)	3	1	1	3
f4: Moments permissions settings for strangers	3	3	4	4
f5: Setting nearby people	3	4	4	4
f6: Adding friends	3	4	3	4
f7: Moments location targeting	1	1	3	3
f8: Using mini-program	3	4	3	4

assignments and the calculated asset value are 5, which is the highest. Compared with WeChat wallet transfers between friends, the required protection information and processing methods are more accurate and complete than the WeChat wallet transfers between friends; that is, the integrity assignment is relatively high. For feature items that are likely to come in contact with strangers (Moments permissions settings for strangers, setting nearby people, adding friends, and using mini-program), we have assigned more average values. The location of the Moments is mostly limited to friends, so the value is lower.

3.4. Threat Identification. According to the frequency of threats, the possibility of threats is defined and divided into 5 levels. The higher the level, the higher the probability of threats.

It can be seen from Table 2 that the number of account password disclosures and WeChat wallet consumption records between non-friends during the chat process has a greater impact on each leak, so the interval assignment frequency of threats of different levels is smaller. The remaining eigenvalue assignment intervals are larger or assigned according to the settings in the specific WeChat.

3.5. Vulnerability Analysis. This paper uses the Common Weak Evaluation System (CVSS). The CVSS evaluation system consists of three measurement groups: the basic measurement group, the time measurement group, and the environment measurement group [15].

Basic metric = round_to_1_decimal (10 ∗ access vector ∗ access complexity ∗ authentication ∗ ((confidentiality impact ∗ confidentiality impact weight value) + (consistent impact ∗ consistency impact weight value) + (availability impact ∗ availability impact weight value)))

The values in Table 3 were selected according to Table 4 [16]. Since personal privacy leaks are based on local information, all access vectors are selected locally. WeChat has official protection measures, so the complexity of access is all high. Authentication refers to verifying whether the user has the right to access the system. Authentication is only required for special operations, so all selections are not required. If the account password disclosed in the chat is leaked, it will cause the user to lose all his accounts, so only the confidentiality impact, consistency impact, and availability impact of this feature item are selected all, and the rest are selected all or according to the impact. The confidentiality impact weight value, consistency impact weight value, and availability impact weight value are assigned according to the proportion of each characteristic item affected by the three attributes. Finally, the basic measurement value is calculated.

Time metric = round_to_1_decimal (basic metric ∗ available for use ∗ grade that can be repaired ∗ confidentiality of the report)

The values in Table 5 were selected according to Table 6 [16]. The leakage of the account password during the chat is most likely to be used, so this feature item can be selected for high utilization. The transfer records have low availability, so the selection is not confirmed. The availability of location selection in Moments is theoretically proven to be practical and feasible for the remaining feature items. The level that can be repaired is assigned according to the featured item according to whether it is

Table 2: Frequency of threats.

Characteristic values	Level 5	4	3	2	1
f1: Account password revealed during chat	10 times or above	7~9	5~6	3~4	0~2
f2: WeChat wallet consumption records (non-friends)	10 times or above		5~6	3~4	0~2
f3: WeChat wallet transfer records (friends)	21 times or above	16~20	11~15	6~10	0~5
f4: Moments permissions settings for strangers	Allow strangers to see ten Moments				Not allowing strangers to see ten Moments
f5: Setting nearby people	Open				Close
f6: Adding friends	No verification required; can be searched through WeChat, QQ number, mobile phone number; can be added through group chat, QR code, business card	No verification required; can be searched through WeChat, QQ number, mobile phone number; not added through group chat, QR code, business card	Requires verification; can be searched by WeChat, QQ number, mobile phone number; can be added through group chat, QR code, business card	Requires verification; can be searched by WeChat, QQ number, mobile phone number; not added by group chat, QR code, business card	Requires verification; cannot be searched by WeChat, QQ, or mobile phone number; can be added through group chat, QR code, business card
f7: Moments location targeting	10 times or above	7~9	5~6	3~4	0~2
f8: Using mini-program	10 and above	7~9	5~6	3~4	0~2

Table 3: Calculation table of basic metrics.

Characteristic values	Related parameters a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
f1	0.7	0.8	1.0	1.0	0.5	1.0	0.25	1.0	0.25	5.6
f2	0.7	0.8	1.0	0.7	0.333	0.7	0.333	0.7	0.333	3.92
f3	0.7	0.8	1.0	0.7	0.333	0.7	0.333	0.7	0.333	3.92
f4	0.7	0.8	1.0	0.7	0.333	0.7	0.333	1.0	0.333	2.3
f5	0.7	0.8	1.0	0.7	0.333	0.7	0.333	1.0	0.333	2.3
f6	0.7	0.8	1.0	0.7	0.333	0.7	0.333	1.0	0.333	2.3
f7	0.7	0.8	1.0	0.7	0.25	0.7	0.25	0.7	0.5	3.92
f8	0.7	0.8	1.0	0.7	0.333	0.7	0.333	1.0	0.333	2.3

a1: access vector, a2: access complexity, a3: authentication, a4: confidentiality impact, a5: confidentiality impact weight value, a6: consistency impact, a7: consistency influence weight value, a8: usability impact, a9: usability impact weight value, and a10: basic measure.

easy to recover after the leak. The transfer records and Moments positioning are better than other feature items. Therefore, high and theoretical are selected, and the rest are selected unconfirmed. The confidentiality of the report has been uniformly selected.

Environmental metric value = round_to_1_decimal ((time metric score + ((10-time metric score) ∗ incidental loss impact)) ∗ target distribution).

The values in Table 7 were selected according to Table 8 [16]. The impact of the loss of transfer records and location of the Moments is small, so the impact of incidental losses is selected as medium and low, and the rest are selected as high. The target distribution is assigned according to the distribution of the feature items. The account password and password in the chat are selected to be high, and the rest are selected to be low or medium. Calculate environmental metrics.

Vulnerability value calculation formula [16]: V = INT {(environmental measurement value ∗ 5)/10 + 0.5}, so vulnerability value is shown in Table 9.

Table 4: Basic metric value assignment table.

Parameters	Values corresponding to different degrees			
Access vector	Local: 0.7		Remote: 1.0	
Access complexity	Height: 0.8		Low: 1.0	
Authentication	Requires: 0.6		Not required: 1.0	
Confidentiality impact	None: 0	Section: 0.7	All: 1.0	
Confidentiality impact weight value	Normal: 0.333	Confidentiality: 0.5	Consistency: 0.25	Availability: 0.25
Consistency impact	None: 0	Section: 0.7	All: 1.0	
Consistency influence weight value	Normal: 0.333	Confidentiality: 0.25	Consistency: 0.5	Availability: 0.25
Usability impact	None: 0	Section: 0.7	All: 1.0	
Usability impact weight value	Normal: 0.333	Confidentiality: 0.25	Consistency: 0.25	Availability: 0.5

Table 5: Time measurement value calculation table.

Characteristic values	Related parameters			
	Exploitability	Repairable level	Confidentiality of the report	Time metric
f1: Account password revealed during chat	1.00	0.87	0.90	4.4
f2: WeChat wallet consumption records (non-friends)	0.85	1.00	0.90	3.0
f3: WeChat wallet transfer records (friends)	0.85	1.00	0.90	3.0
f4: Moments permissions settings for strangers	0.95	0.87	0.90	1.7
f5: Setting nearby people	0.95	0.87	0.90	1.7
f6: Adding friends	0.95	0.87	0.90	1.7
f7: Moments location targeting	0.90	0.90	0.90	2.9
f8: Using mini-program	0.95	0.87	0.90	1.7

Table 6: Time metric value assignment table.

Parameters	Values corresponding to different degrees			
Exploitability	Unconfirmed: 0.85	Proved by theory: 0.90	Practical: 0.95	High: 1.00
Repairable level	Unconfirmed: 0.87	Proved by theory: 0.90	Practical: 0.95	High: 1.00
Confidentiality of the report	Unconfirmed: 0.90	Unverified: 0.95	Confirmed: 1.00	

Table 7: Calculation table of environmental measures.

Characteristic values	Related parameters		
	Collateral loss effects	Target distribution	Environmental metric value
f1: Account password revealed during chat	0.5	1.00	7.2
f2: WeChat wallet consumption records (non-friends)	0.3	0.25	1.3
f3: WeChat wallet transfer records (friends)	0.3	0.25	1.3
f4: Moments permissions settings for strangers	0.5	0.75	4.4
f5: Setting nearby people	0.5	0.75	4.4
f6: Adding friends	0.5	0.25	1.5
f7: Moments location targeting	0.1	0.25	0.9
f8: Using mini-program	0.5	0.75	4.4

Note: round_to_1_decimal refers to rounding to one decimal place.

Table 8: Assignment table of environmental metrics.

	Values corresponding to different degrees			
Parameters	No	Low	Middle	High
Collateral loss effects	0	0.1	0.3	0.5
Target distribution	0	0.25	0.75	1.0

3.6. Risk Calculation: The Calculation of Risk Is as Follows. After completing asset identification, threat identification, and vulnerability identification, an appropriate model can be used to calculate the risk value of a security event caused by the vulnerability using threats. This article adopts the risk calculation model in Chinese National Standard GB/ *T* 20984 "Information Security Technology, Information Security Risk Assessment Specification".

The formula is expressed as risk value = R (A, T, V) = R (L (T, V), F (A, V)). Among them, R is the calculation function

Table 9: Vulnerability value calculation table.

Characteristic values	Vulnerability value
f1: Account password revealed during chat	4
f2: WeChat wallet consumption records (non-friends)	1
f3: WeChat wallet transfer records (friends)	1
f4: Moments permissions settings for strangers	3
f5: Setting nearby people	3
f6: Adding friends	1
f7: Moments location targeting	1
f8: Using mini-program	3

of security risk, A is the value of the asset, T is the threat, V is the vulnerability, L is the possibility of threatening the use of the vulnerability of the asset to cause a security event, and F is the loss caused by the security event.

In the specific calculation of risk, there are three key calculation links.

3.6.1. Calculate the Probability of a Security Incident. According to the frequency and vulnerability of threats, calculate the probability that a threat will cause a security event using vulnerability, that is, the probability of a security event = L (frequency of threats, the severity of vulnerability) = L (T, V).

This system uses a two-dimensional matrix algorithm to calculate the probability of a security event, as shown in Table 10 [15].

3.6.2. Calculate Losses Caused by Security Incidents. According to the value of the asset and the severity of the vulnerability, calculate the loss caused by the security event once it occurs, that is, the loss caused by the security event = F (asset value, severity of vulnerability) = F (A, V).

This system uses a two-dimensional matrix method to calculate the loss of security events, as shown in Table 11 [15].

3.6.3. Calculating the Value at Risk. According to the calculated probability of the security event and the loss caused by the security event, calculate the risk value, that is, risk value = R (the probability of the security event, the loss caused by the security event) = R (L (T, V), F (A, V)).

The system uses the two-dimensional code matrix method to calculate the risk value of security events, as shown in Table 12 [15].

4. Sum Based on Weights

The risk value of each data security event is obtained from Table 12, and each risk value is multiplied by the weight value of Table 13 to obtain the final risk value.

$$T = INT\left\{\left[\frac{\sum_{i=1}^{N}(t_i)}{2N + 0.5}\right]\right\}. \quad (2)$$

Table 10: Two-dimensional matrix of security event probability calculations.

Severity of vulnerability	Frequency of threats				
	1	2	3	4	5
1	2	4	7	9	12
2	4	6	9	13	16
3	6	9	13	17	21
4	8	11	14	21	23
5	9	13	18	23	25

Table 11: Two-dimensional matrix table of security event loss calculation.

Severity of vulnerability	Asset value				
	1	2	3	4	5
1	2	3	6	9	11
2	3	6	9	12	16
3	5	8	12	16	20
4	7	10	13	19	22
5	9	13	18	23	25

Table 12: Two-dimensional matrix table for calculating the risk value of security events.

Loss caused by the security event	Probability of the security event				
	1~5	6~10	11~15	16~20	21~25
1~5	3	6	9	14	13
6~10	6	11	17	21	21
11~15	11	18	22	30	30
16~20	15	21	31	40	55
21~25	22	35	55	85	100

Table 13: Comprehensive threat calculation table.

Characteristic values	Related parameters			
	T_s	T_i	t	Weights
f1: Account password revealed during chat	4	5	9	0.225
f2: WeChat wallet consumption records (non-friends)	3	3	6	0.15
f3: WeChat wallet transfer records (friends)	2	2	4	0.1
f4: Moments permissions settings for strangers	2	2	4	0.1
f5: Setting nearby people	2	3	5	0.125
f6: Adding friends	2	2	4	0.1
f7: Moments location targeting	2	1	3	0.075
f8: Using mini-program	2	3	5	0.125

All comprehensive threat value calculation formulas [15–17].

The calculation formula for the comprehensive calculation value of a single threat to an information asset: $t = T_s + T_i$.

Among them, t is a single threat comprehensive value, T_s is a threat source value, defined as a value between 1 and 5, and T_i is an impact degree value and is also defined as a value between 1 and 5.

4.1. SVM Algorithm Application. The intelligent detection system uses the SVM algorithm to divide the comprehensive

threat value (as shown in Figure 2) and further divides the risk level of the user account more accurately.

The specific process is as follows.

Based on the comprehensive threats mentioned above, it is worth calculating the risk value. The obtained risk values are divided into two categories. The scores of 1 to 40 are low-risk areas, and the scores of 40 to 100 are high-risk areas. Among them, 1 to 20 in the low-risk areas are defined as safe, 21 to 40 are defined as basic safety, 41 to 59 in the high-risk areas are defined as higher risks, and 60 to 100 parts are expressed as high risks.

The feature quantities of two types of risk values defined as safety and basic safety are recorded in the initial feature vector set 1. The feature quantities defined as two types of risk values of higher-risk and high-risk are recorded in the initial feature vector set 2.

Normalize the feature data items to remove the extreme data. Convert the processed two types of data formats into an input format acceptable to the SVM classifier (class vector Y, feature vector Xi)

The corresponding classifier 1 is trained using data defined as safe and basic safety as training samples, and the corresponding classifier 2 is trained using data defined as safe and basic safety as training samples.

Set the SVM parameters and use the K-fold cross-validation algorithm to find the optimal parameters. Perform asset identification, vulnerability analysis, and threat identification from the characteristic values read by the user. After risk calculation, determine the low-risk area or high-risk area based on the score and enter the corresponding risk area as a test sample. The SVM classification model performs classification judgment. Substitute the results obtained by the SVM into the Naive Bayes formula to obtain the security risk probability, and send feedback of the final results to the user.

4.2. SVM Algorithm Training. The SVM calculation process is shown in Figure 3. The format of the training data and test data is:

<label> <index (1)>: <value1> <index (2)>: <value2> ...

For example: 0 1 : 1 2 : 1 3 : 1 4 : 1 5 : 2 6 : 2 7 : 2 8 : 2.

Among them:

<label> is the category identifier of the training data set, set to 0 and 1,0 for security, and 1 for basic security.

<index> refers to 8 feature quantities of the input algorithm, which are integers.

<value> is the value of the feature code for each item and is an integer.

SVM_train implements training on training samples to obtain SVM models.

SVM classification is a prediction of the classification result of the data set according to the trained model.

Use SVM_train to train the input training data set to obtain the SVM model file. The SVM algorithm maps each input training sample, that is, an n-dimensional vector into a high-dimensional space, forming multiple scattered points, and passing the aggregation of points. The region simulates the classification hyperplane and continuously uses the newly input training sample data to make corrections, generates template files, and records the classification features.

In this paper, the K-fold cross-validation method is used to obtain the optimal parameters by verifying the accuracy of the results. The main purpose of the verification algorithm is to divide the data set A into a training set B and a test set C. When the sample size is small, the data set A can be randomly divided into k packets, and one of the packets is used as the test set at a time. The remaining k-1 packets are trained as a training set. The cross-validation method is used to prevent overfitting caused by the model being too complicated [18]. By constantly transforming two important parameters of the SVM: the penalty factor C and the kernel function parameter g, the optimal parameters $C = 2048$ and $g = 0.0078$ are determined.

4.3. SVM Algorithm Processing

SVM classifier 1:

Input $x = \{a_1, a_2, \cdots, a_m\} y \{y_0, y_1\}$, x represents the feature value set of each sample in the test sample, y represents the categories are 0 and 1, which represent safety and basic safety respectively;

Output The user's security risk probability is less than or equal to 50% as safe, and greater than 50% as basic safety.

Step 1: Normalize the feature data

Step 2: Convert the processed feature data into an input format acceptable by the classifier (feature vector x, category vector y) to obtain training samples

Step 3: Set the SVM type to 0-SVM and the kernel function type to radial basis function (RBF)

Step 4: Set the penalty factor C and kernel function parameter G

Step 5: Set the K value of the K-fold cross-validation algorithm

Step 6: Use the SMO algorithm to find the support vector

Step 7: Build a hyperplane model from training samples

Step 8: Enter the test samples for classification, and get the classification result y

Step 9: Calculate the $P(a_i|y)$ to obtain the conditional probability ratio of each feature attribute in the result classification y

Step 10: Calculate $p(y)$ to get the probability of category y appearing

Step 11: Calculate $p(a_i)$ to get the probability of each characteristic attribute

Step 12: Substitute the formula

$$P(y|x) = \frac{P(x|y)P(y)}{P(a_i)P(a_2)\cdots P(a_m)} \cdot a. \tag{3}$$

Step 13: return $P(y|x)$

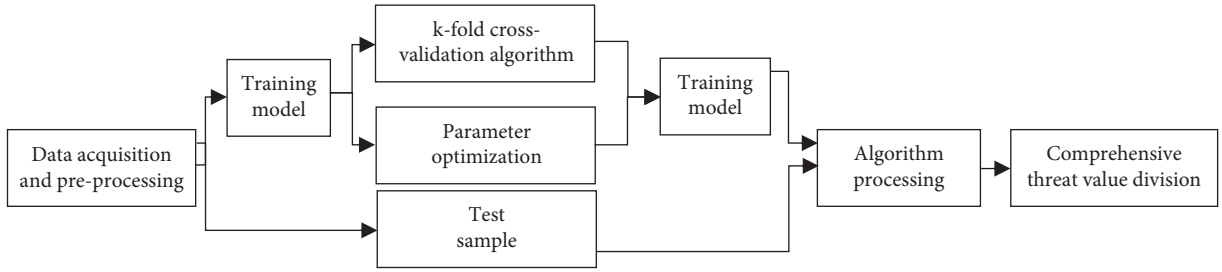

Figure 2: SVM algorithm application diagram.

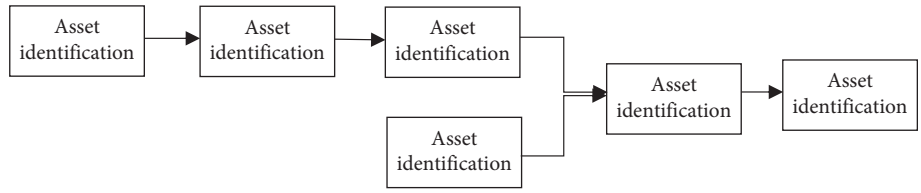

Figure 3: SVM calculation process.

SVM classifier 2:

Input $x = \{a_1, a_2, \cdots, a_m\} y\{y_0, y_1\}$, x is the feature value set of each sample in the test sample, y is the category is 0 and 1, which means high-risk and higher-risk respectively

Output The user's security risk probability is less than or equal to 50% as a high-risk and greater than 50% as higher-risk.

SVM classifier 2 process is the same as SVM classifier 1.

5. Experiment and Analysis of Intelligent Detection System

5.1. Environmental Configuration and Data Acquisition. This test system is designed to run on the Android platform. During the test phase, Android Studio is used to simulate the Android platform for various tests.

Due to the inconvenience of directly obtaining the personal privacy data of the user's WeChat, a questionnaire was used at this stage to collect the WeChat usage of 149 users as a training sample for the SVM classifier. The specific content of the questionnaire is shown as Appendix in Supplementary Materials (available here).

5.2. Functional Test. User test assignment table is shown in Table 14. For the functional test of this system, we first obtained the WeChat related records of a user for testing, as sample 1. The user has been tested and calculated a comprehensive threat value of 26. After obtaining the comprehensive threat value, the data format of this sample is converted into an input format acceptable to the classifier. Based on the user's comprehensive threat value of 26, the sample should be determined. Enter SVM classifier 1. The input format is 0 1:3 2:1 3:2 4:1 5:1 6:1 7:4 8:2, and processing of sample 1 is complete.

After removing the extreme data from the remaining samples, the above steps are processed and sent to the corresponding SVM classifier. The training samples are used to build a hyperplane model. When the system intelligently detects the risk leakage probability, it will automatically obtain the feature quantity, calculate the comprehensive threat value after calculation, and send it to the corresponding SVM classifier to obtain the final security risk probability.

According to this method, we processed the results of 149 user questionnaires and calculated the number of scores for each segment. The results are shown in the following Table 15.

From Table 15, we can see that a total of 89 users are in the low-risk area and 34 users are in the risk area, of which 26 users are in the security zone. This shows that the security awareness education has been effective, and people have realized that personal privacy is important, but there are also many users in high-risk areas, indicating that there is still a need to continue with efforts to expand coverage and increase everyone's security awareness.

5.3. Performance Testing. Obtain user WeChat related information through a questionnaire. As a sample, test the personal privacy leak detection value of a user's social network, and give a warning or suggestion to get the percentage of people at each risk level, and then get the current data of whether people know and implement the degree of privacy protection in place, which aspects are of importance to people, and which aspects are ignored by people, and provide directions for the promotion of privacy protection awareness in the future. The findings are shown in Figure 4.

At the same time, we counted the number of occurrences of high threats for each feature item (that is, the number of times assigned 4 or 5).

In Figure 5, we can see that most people have a certain awareness of self-privacy protection, but many people ignore the function of "people nearby" and allow strangers to view the private information that may be leaked in Moments. A system that can protect the privacy of the user's privacy is

Table 14: User test assignment table.

Characteristic values	User test related assignments
f1: Account password revealed during chat (6 times)	3
f2: WeChat wallet consumption records (non-friends) (3 times)	1
f3: WeChat wallet transfer records (friends) (10 times)	2
f4: Moments permissions settings for strangers (close)	1
f5: Setting nearby people (close)	1
f6: Requires verification; cannot be searched by WeChat, QQ, or mobile phone number; can be added through group chat, QR code, business card	1
f7: Moments location targeting (8 times)	4
f8: Using mini-program (2)	2

Table 15: Number distribution of each risk area.

Level	Number
Safety	60
Basic safety	29
High-risk	26
Higher-risk	34

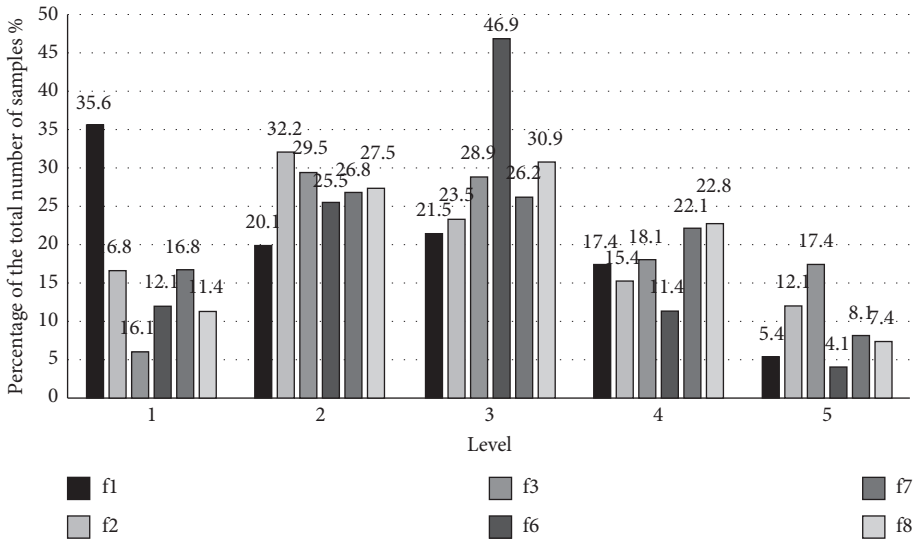

Figure 4: Statistics of the percentage of occurrences of each feature.

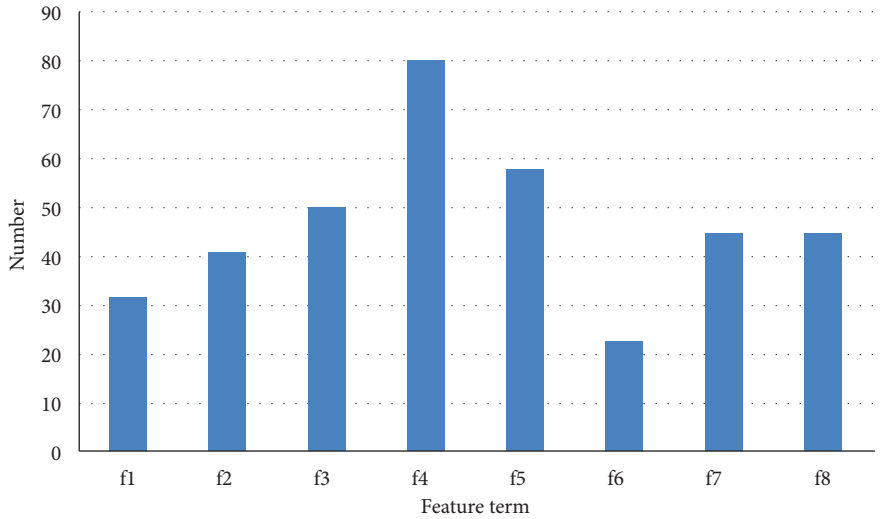

Figure 5: Statistics of occurrence times of high threats for each feature.

essential. Through this intelligent detection system for personal privacy leaks for social networks, users can clearly understand their negligence in the process of using WeChat and correct them to prevent problems before they occur.

6. Concluding Remarks

The system proposed in this article is based on reading multiple characteristic values of personal WeChat and establishing a model based on three aspects of asset identification, threat identification, and vulnerability analysis. According to the risk calculation models and methods in the national standards of information security risk assessment standards, the dimension matrix table calculates the possibility of security events, the loss of security events, and the risk value of security events, determines the risk level according to the magnitude of the risk, evaluates the personal privacy leakage of the user's online social software, gives a score, and informs the user about source of risk.

This article only mentions the scoring function in the system and the function of displaying the risk source of personal privacy leakage. In the future, more functions will be added to improve the entire system, which will also make the judgment more accurate and create a more accurate situation for the individual users, creating safe environment to use social networks.

Data Availability

Due to the inconvenience of directly obtaining the personal privacy data of the user's WeChat, a questionnaire was used at this stage to collect the WeChat usage of 149 users as a training sample for the SVM classifier. The specific content of the questionnaire is given in Supplementary Materials.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by humanities and social sciences research project of the Ministry of Education (No. 20YJAZH046) and Higher Education Research Projects (No. 2020GJZD02), "Practical training plan" (Entrepreneurship) for cross training of high-level talents.

Supplementary Materials

The detailed information of the questionnaire is given in Appendix I. According to the number, it is divided into five levels: 5 is the highest, it is gradually decreased in order, and 1 is the lowest. (*Supplementary Materials*)

References

[1] G. Dini, F. Martinelli, I. Matteucci, M. Petrocchi, A. Saracino, and D. Sgandurra, "Risk analysis of Android applications: a user-centric solution," *Future Generation Computer Systems*, vol. 80, pp. 505–518, C, 2018.

[2] P. Van Schaik, J. Jansen, J. Onibokun, J. Camp, and P. Kusev, "Security and privacy in online social networking: Risk perceptions and precautionary behaviour," *Computers In Human Behavior*, vol. 78, pp. 283–297, 2018.

[3] Y.-Q. Q. Zhang, S.-Q. Lv, and D. Fan, "Anomaly detection in online social networks," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 38, no. 10, pp. 2011–2027, 2015.

[4] Z. F. Huo, X.-F. Meng, and Y. Huang, "PrivateCheckIn: trajectory privacy-preserving for check-in services in MSNS," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 36, no. 4, pp. 716–726, 2013.

[5] Y. Y. Cheng, L.-Y. B. Ying, S.-B. R. Jiao, P.-R. G. Su, and D.-G. Feng, "Research on user privacy leakage in mobile social messaging applications," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 37, no. 1, pp. 87–100, 2014.

[6] K. Wang, *Research and Application of Android Platform Application Risk Detection*, Beijing University of Posts and telecommunications, Beijing, China, 2012.

[7] T. X. Li, Y.-X. Q. Xing, A.-Q. J. Hu, and Y.-J. Wang, "Research on multi-dimensional privacy leakage evaluation model for mobile terminals," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 41, no. 9, pp. 2134–2147, 2018.

[8] Q. Kuang, *Risk Assessment Based on Personal Privacy Disclosure [D]*, Guizhou University, Guizhou, China, 2016.

[9] H. Chen, A. B. Ünal, M. Akgün et al., "Privacy-preserving SVM on outsourced genomic data via secure multi-party computation," in *Proceedings of the Sixth International Workshop on Security and Privacy Analytics*, pp. 61–69, New Orleans, LA, USA, October 2020.

[10] S. Park, J. Byun, J. Lee, J. H. Cheon, and J. Lee, "HE-friendly algorithm for privacy-preserving SVM training," *IEEE Access*, vol. 8, pp. 57414–57425, 2020.

[11] J. Liang, Z. Qin, J. Ni et al., "Efficient and privacy-preserving outsourced SVM classification in public cloud," in *Proceedings of the ICC 2019-2019 IEEE international conference on communications (ICC)*, pp. 1–6, IEEE, Shanghai, China, May 2019.

[12] Y. Xu, C. Wu, K. Zheng, Xu Wang, X. Niu, and T. Lu, "Computing Adaptive Feature Weights with PSO to Improve Android Malware Detection," *Security and Communication Networks*, vol. 2017, Article ID 3284080, 14 pages, 2017.

[13] H. Xiang, O. Fu, and B. Zhan, *Information Security Measurement and Risk Assessment*, Electronic Industry Press, Beijing, China, 2014.

[14] CSDN: (Support Vector Machine) https://blog.csdn.net/android_ruben/article/details/78308868?utm_source=debugrun&utm_medium=referral.

[15] M. Shang and M. Chen, "A calculation method of risk value of information assets," *Network security technology and application*, vol. 5, pp. 163-164, 2014.

[16] Y. Wu, X. Li, and K. Lu, *Information Security Risk Assessment*, pp. 70–83, Tsinghua University Press, Beijing, China, 2007.

[17] 2007 Information Security Technology— Risk Assessment Specification for Information Security: General Administration of quality Supervision, Inspection, and Quarantine of the People's Republic of China and China National Standardization Administration.

[18] N. Jia, S. Fu, and M. Xu, *Privacy-Preserving Blockchain-Based Nonlinear SVM Classifier Training for Social Networks*, Security and communication networks, vol. 2020, Article ID 8872853, 10 pages, 2020.