# Research on Privacy Security Steady State Evaluation Model of Mobile Application Based on Information Entropy and Markov Theory

Ming Yang[1], Li Jia[1], Tilei Gao[1], Tao Zhang[1], and Wanyu Xie[2]
*(Corresponding author: Tilei Gao)*
School of Information, Yunnan University of Finance and Economics[1]
Kunming 650221, China
Personnel Department, Kunming Metallurgy College[2]
Email: gtllei@ynufe.edu.cn
*(Received May 25, 2020; Revised and Accepted Dec. 10, 2020; First Online Aug. 14, 2021)*

## Abstract

In the process of using mobile applications, if users want to use customized personalized services, they need to provide personal privacy information to businesses, which constitutes a potential threat to their privacy security. In order to help users choose secure mobile applications reasonably, and ensure the security of users during use, this paper combines information entropy and Markov theory to study the measurement of user privacy security of mobile applications, and proposes an effective steady-state evaluation model of user privacy security. Finally, the model is put into a specific case for analysis. The analysis results show that the method can effectively evaluate the privacy security of users, and compared with the previous methods, the method is more real, objective and simple.

*Keywords: Information Entropy; Markov Theory; Mobile Application Security; Privacy Security; Risk Evaluation*

## 1 Introduction

With the popularity of smart phones, more and more mobile applications appear in front of people, providing users with a variety of accurate services. However, while enjoying the convenient services brought by these mobile applications, users' privacy security is also threatened. These leakage scenarios include wireless network connection, mobile payment, public equipment use, fingerprint recognition, face recognition, *etc.* [5]. These privacy leakage scenarios include wireless network connection, mobile payment, public equipment use, fingerprint recognition, face recognition, *etc.* In these scenarios, the privacy information of users is continuously collected, transmitted and stored, and is faced with the risk of being leaked, abused and stolen. These risks not only come from individuals, but also from service providers and terminal equipment. There are various ways to steal data, such as internal theft, external hacker intrusion and employee negligence. According to the information collected by the Identity Theft Resource Center and the U.S. Department of health and human services, more than 137 million records were leaked in 2019 [32], and the information leaked due to the privacy security problems of mobile applications accounted for a large proportion.

At present, most of the research on user privacy security is focused on cloud services and big data applications. It is known that, compared with the traditional information system security risk factors, the risk hierarchy of users privacy information in mobile applications is more complex, including traditional information system security risk, user behavior risk [11], third party application risk [6] and unique risk of mobile application service [10]. Therefore, considering the importance of security risk assessment to the ecosystem and sustainable development of mobile e-commerce platform [28], this paper proposes a risk evaluation model for privacy security of mobile applications based on the characteristics of mobile applications and the security model of information system [26].The evaluation model can provide users with practical and intuitive risk evaluation results and protect their privacy security in mobile commerce.

The organizational structure of this paper is as follows: Section 1 - Introduction: The background, content and significance of the research are presented. Section 2 - Related researches: This chapter discusses the privacy security risk attributes and evaluation methods in mobile applications, and puts forward the main problems to be solved in the current research. Section 3 - Research on privacy security measurement and evaluation of mobile applications Based on Information Entropy and Markov Theory: According to the characteristics of mobile application users' privacy security, this chapter puts information entropy and Markov into the research of user

privacy security, proposes a privacy security measurement method based on information entropy, and proposes a dynamic evaluation method of user privacy security based on Markov theory. Section 4 - Steady state evaluation model of mobile application privacy security based on information entropy and Markov: In this chapter, a user privacy security risk evaluation model of mobile application based on information entropy and Markov is established, and the evaluation steps of the model are introduced in detail. Section 5 - Case analysis: In this chapter, the proposed model is put into three different types of mobile applications to carry out case studies. Section 6 - Conclusion: This chapter summarizes the research work of this paper and points out the future research direction.

## 2 Relevant Researches

Recent researches on users' privacy security of mobile application are studied and summarized for this paper, as follows:

### 2.1 Research on Users Privacy Security Risk Attribute of Mobile Application

The privacy security research of mobile application is different from the traditional security research. It includes the influence of various risk factors, such as the technical defects of privacy protection, the user's own security weak consciousness, the application environment risk, the terminal equipment risk factor, the operator's management risk, and the privacy risk caused by laws and regulations.

1) Technology risk. Reference [23] studies the privacy security of mobile applications, emphasizing the importance of user authentication technology and protocol. Reference [18] pointed out that the privacy security of users in mobile applications is affected by many technical factors, including data encryption, intrusion detection, identity management, security awareness, *etc.* Literature [1] pointed out that whether to adopt anonymization technology will directly affect the privacy and security of mobile applications. Reference [4] analyzes the security problems of various levels of network physical system (CPS) architecture, and points out that to improve its security, attention should be paid to the influence of related technologies, such as access control, data encryption, attack detection, user authentication and authorization, *etc.* Reference [12] proposed a secure routing protocol with node self-sufficiency resistance, which improves the protection of mobile application privacy.

2) User vulnerability risk. Ampong [3] mentioned that in the process of using mobile applications, users' privacy awareness, privacy concerns and privacy intrusion experience will directly affect their personal privacy security. Literature [7, 34] thinks that location information is extremely sensitive in mobile commerce, and the exposure of location information may cause the risk of mobile application information abuse. Reference [33] analyzes the characteristics of user privacy security in the big data environment, and points out that some common user behaviors will directly cause the disclosure of personal privacy information, such as privacy Association setting, spatial location sharing, information behavior negligence and simple password setting.

3) Application scenario risk. In some mobile applications, there are usually mobile advertisements, which are intrusive to users' privacy, and even forcibly obtain users' personal location information [25]. In addition, the authorization of some application rights will also affect the privacy and security of users. Reference [9] mentioned that users will be forced to agree to open some application permissions before using some mobile commerce applications, resulting in users having no autonomy in whether to share their own information.

4) Mobile terminal device risk. References [13,16,17,19] respectively discussed various privacy risks existing in mobile terminals, including the protection of sensitive data by the device, the location tracking function of the device, the management of the authority of the device, and the malicious monitoring function of the device itself.

5) Management and legal risk. As an important role in the process of mobile application interaction, the improper management of service providers and the restrictions of laws will affect the privacy and security of users. Reference [14] mentioned that establishing a standard privacy policy for mobile applications is the key to solve their privacy security issues. Literature [20] pointed out that the common management risks include imperfect disclosure standards of privacy information, lack of supervision and punishment system, malicious disclosure of internal personnel, *etc.* Literature [27] analyzes the consumption behavior characteristics of mobile application users, and establishes corresponding risk evaluation indicators, which include privacy management mechanism, platform privacy protection investment, information-sharing risk, third-party information collection, privacy law differences and other related factors.

### 2.2 Research on Privacy Security Risk Ealuation Method

At present, the researches on privacy security evaluation methods for mobile applications are rare. Literature [8, 21, 22, 31] proposed some effective measurement methods for the security of cloud services using

the method of information entropy, but did not specifically evaluate the privacy security. In Reference [24], a privacy-considered information security evaluation model was built with the risk recommendation system based on the identifiability, context of use, quantity, sensitivity, And freshness of the personal identity information data. Lo [15] proposed a user privacy analysis framework called LRPdroid for Android platform, which realized information leakage detection, user privacy leakage assessment and privacy risk assessment of applications installed on mobile devices based on Android. According to the characteristics of mobile application permission, Reference [30] proposed a mobile application risk evaluation strategy based on mobile application permission characteristics. Reference [29] also proposed a risk evaluation method for mobile applications in Android environment based on its permission characteristics.

Summing up the above methods, it can be found that these assessment methods are not targeted for the evaluation of user privacy security, and most of the evaluation objects are only limited to one kind of risk, without considering the interaction between various risk classes, and also not combining with the actual risk environment for dynamic assessment.

Therefore, in order to solve the above problems, this paper proposes a special evaluation method for mobile application privacy security, and establishes the corresponding model to realize the multi-level and multi angle evaluation of mobile application privacy security.

# 3 Research on Privacy Security Measurement and Evaluation of Mobile Applications based on Information Entropy and Markov Theory

In order to realize the dynamic evaluation of privacy and security of mobile applications, two following research contents have been carried out:

## 3.1 Research on Privacy Security Measurement Method of Mobile Application based on Information Entropy

Measurement is the premise of evaluation, an objective and accurate measurement result will directly affect the evaluation results. However, the privacy security of mobile applications is an abstract concept, and only be measured by specific methods. Therefore, in order to effectively measure the privacy and security of mobile applications, this paper proposes to use "the uncertainty of risk" to quantitatively describe "the level of privacy security of mobile applications" from the opposite perspective. As shown in Figure 1.

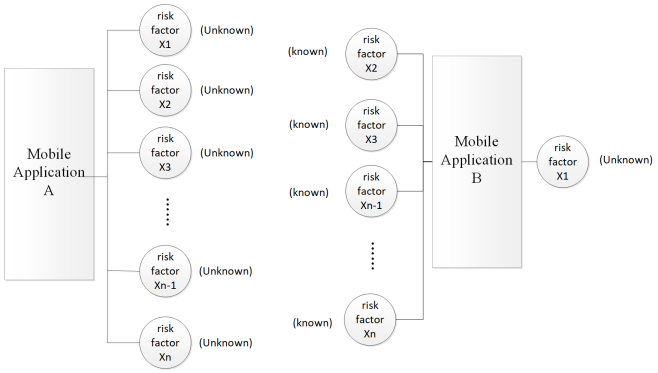


Figure 1: Two extreme mobile application risk environments

According to the theory of information entropy, it is assumed that both A and B contain some unknown risk factors $X_i, X = \{X_1, X_2, \ldots, X_n\}$.

1) The entropy of application A is maximal, and its risk is the highest. There are $n$ unknown risks in application A, and the probability of the risk occurrence $P(X_1) = P(X_2) =, \ldots, = P(X_n)$, the entropy value $H(X) = \log_2 n$ will reach its maximum according to the information entropy equation. At this time, the system contains many risk factors, and its controllable degree will reach the lowest.

2) The entropy of application A is minimal, and its risk is the lowest. There is only one unknown risk in the application B. According to the equation of information entropy, its entropy $H(X)$ will reach the lowest. At this time, the uncertainty of risk factors contained in the system is the lowest, and the risk is completely controllable, that is to say, the security of the system reaches the highest.

However, in practice, a complex mobile application is bound to be affected by a number of different risk factors, and the probability of these factors is different too, so that its entropy is bound to be between the maximum value and the minimum value. Therefore, according to the theory of information entropy, the level of mobile application security in the actual situation can be described by the size of entropy.

## 3.2 State Description of Mobile Application Risk Environment Based on Markov

On the basis of information entropy measurement, this paper will further combine Markov theory to describe the risk environment state of mobile application, and express it with the form of mathematical matrix, so that to provide support for the follow-up steady-state evaluation research [2].

It is assumed that there are $n$ risk factors $X_i$ in a mobile application environment, so its complex risk environment can be described as the following matrix:

$$R = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{51} & X_{52} & \cdots & X_{nn} \end{bmatrix} \quad (1)$$

Matrix $R$ is the privacy security risk matrix of mobile applications, in which:

1) Element $X$ on diagonal $X_{ii}$ represent the separate occurrence of risk factors $X_i$;

2) Non diagonal element $X_{ij}$ represent simultaneous occurrence of risk factors $X_i$ and $X_j$.

It is known that in the actual use of mobile applications, their risk factors have different probability of occurrence, which leads to they have a variety of possible privacy security states. Therefore, in order to objectively evaluate the privacy security of mobile applications, it is necessary to effectively restore their random risk state. Using $P(X_{ij})$ represents the probability of different risks, and substituting it into matrix (1), the risk state transition matrix as follows will be obtained:

$$(R) = \begin{bmatrix} P(X_{11}) & P(X_{12}) & \cdots & P(X_{1n}) \\ P(X_{21}) & P(X_{22}) & & P(X_{2n}) \\ \vdots & & \ddots & \vdots \\ P(X_{n1}) & P(X_{n2}) & \cdots & P(X_{nn}) \end{bmatrix} \quad (2)$$

### 3.3 Research on Steady State Evaluation Method Based on Information Entropy and Markov

After proposing the privacy security measurement method and risk state description matrix of mobile application, this paper will establish a special privacy risk attribute model to realize a multi-level and multi-dimensional evaluation of the whole mobile application. As shown in Figure 2:

In this paper, the privacy security attribute model of mobile applications is divided into three levels: target layer A, risk class layer $\beta$ and risk factor layer $X$. In order to objectively reflect the real risk environment, this paper uses cross lines to describe the relationship between layer 2 and layer 3, which is used to reflect the influence of different risk factor $X_i$ on different dimension $\beta$.

As shown in the following example, suppose a mobile commerce contains two types of risk $\beta_1$ and $\beta_2$. The risk factors involved are shown in Table 1:

As shown in Table 1, risk classes A and B contain a common risk factor $X_2$, so their transition state matrix

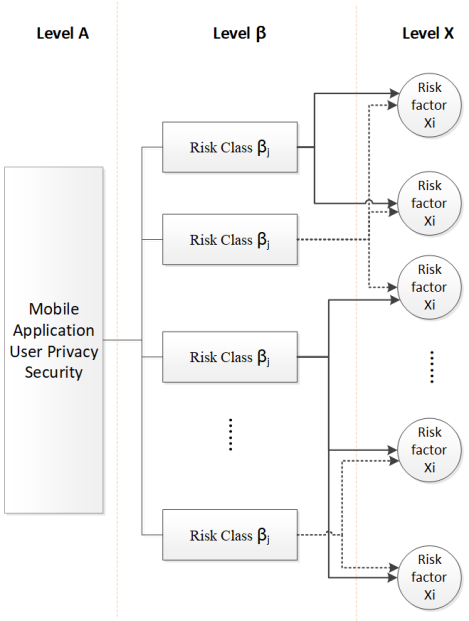


Figure 2: The abstract privacy security risk attribute model of mobile application

Table 1: Risk classes $\beta_1$ and $\beta_2$ and the risk factors they contain

| Risk Class | Risk factors included Risk Class |
|---|---|
| $\beta_1$ | $X_1, X_2$ |
| $\beta_2$ | $X_2, X_3, X_4$ |

$P(R)$ can be calculated as follows:

$$P(R) = \begin{bmatrix} P(\beta_{11}) & P(\beta_{12}) \\ P(\beta_{21}) & P(\beta_{22}) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sum_{i=1}^{3} P(X_i)} P(X_1) & \frac{1}{\sum_{i=1}^{3} P(x_i)} P(X_2) \\ \frac{1}{\sum_{i=3}^{4} P(\alpha_i)} P(X_2) & \frac{1}{\sum_{i=3}^{4} P(\alpha_i)} \{P(X_3) + P(X_4)\} \end{bmatrix} \quad (3)$$

According to the above calculation method, it is assumed that a mobile application has m risk classes $\beta_i$ in the long-term use process, and the steady-state probability of each class is $\widehat{P}(\beta_i)$, which represents the convergence probability of a risk class in the long-term use process. $\widehat{P}(\beta_i)$ can objectively reflect the stable probability of a certain risk class in the long-term use process. It is the result of dynamic evaluation of different risk states through Markov matrix. The relationship between $\widehat{P}(\beta_i)$ and $P(R)$ is as follows:

$$\begin{cases} \widehat{P}(\beta_1) = P(X_{11})\widehat{P}(\beta_1) + P(X_{12})\widehat{P}(\beta_2) + \ldots + P(X_{1m})\widehat{P}(\beta_m) \\ \widehat{P}(\beta_2) = P(X_{21})\widehat{P}(\beta_1) + P(X_{22})\widehat{P}(\beta_2) + \ldots + P(X_{2m})\widehat{P}(\beta_m) \\ \widehat{P}(\beta_3) = P(X_{31})\widehat{P}(\beta_1) + P(X_{32})\widehat{P}(\beta_2) + \ldots + P(X_{3m})\widehat{P}(\beta_m) \\ \vdots \\ \widehat{P}(\beta_m) = P(X_{n1})\widehat{P}(\beta_1) + P(X_{n2})\widehat{P}(\beta_2) + \ldots + P(X_{nm})\widehat{P}(\beta_m) \\ 1 = \widehat{P}(\beta_1) + \widehat{P}(\beta_2) + \ldots + \widehat{P}(\beta_m) \end{cases} \quad (4)$$

Solving the Equation (4), it can obtain the

steady-state probability of various risks $\hat{P}(\beta_i) = \left\{\widehat{P}(\beta_1), \widehat{P}(\beta_2), \ldots, \widehat{P}(\beta_m)\right\}, \sum_{i=1}^{m} \widehat{P}(\beta_i) = 1$.

Further, substituting $\widehat{P}(\beta_i)$ into the information entropy equation and using the reciprocal form, it can quantitatively describe the privacy security of the entire mobile application, the equation is as follows:

$$E = 1/H = 1/-\sum_{i=1}^{m} \widehat{P}(\beta_i) \log_2 \widehat{P}(\beta_i) \tag{5}$$

$E$ represents the security evaluation result of the whole mobile application. The higher the value is, the higher the privacy security of the mobile application is. Similarly, if the probability of occurrence of risk factors contained in class $\beta_i$ is normalized, the security evaluation result $E(\beta_i)$ of the risk class can be obtained. The higher the value, the higher the privacy security of this risk class is.

As mentioned above, around the hierarchical structure shown in Figure 2, this paper proposes a bottom-up privacy security risk evaluation method for mobile applications based on information entropy and Markov.

# 4 Steady State Evaluation Model of Mobile Application Privacy Security Based on Information Entropy and Markov

In this chapter, the proposed hierarchical structure in chapter 3 will be embodied, and the calculation steps of the proposed method will be described in detail, so as to build a privacy security steady state evaluation model of mobile application based on information entropy and Markov.

## 4.1 Risk Attribute Model for Privacy Disclosure of Mobile Commerce Users

In this regard, this paper selects a total of 24 mobile application privacy risk factors, and divides these factors into 5 classes, namely technical risk class $\beta_1$. Application risk class $\beta_2$. Management and legal risk class $\beta_3$. User risk class $\beta_4$. Mobile terminal equipment risk class $\beta_5$. The risk attribute model established according to the above division is shown in Figure 3.

## 4.2 The Calculation Steps of Privacy Security Evaluation Method

According to the evaluation method proposed in chapter 3, the calculation is carried out from bottom to top, which is divided into five steps.

**Step 1:** According to the evaluation standard of risk level shown in Table 2, the occurrence probability grade of each risk factor in the third layer is evaluated, and the probability of single risk factor is obtained by normalization.

Table 2: The level of probability of risk factors occurrence

| level | Definition and description |
|---|---|
| [8, 10) | This factor has a great risk and a direct threat to the user's privacy |
| [6, 8) | This risk has a high probability of occurrence and exists in most mobile business environments |
| [4, 6) | This risk is a common risk, which exists in some mobile commerce |
| [2, 4) | This risk exists and only occurs when special conditions are met |
| (0, 2) | This factor has high security and hardly causes user privacy risk |

**Step 2:** According to the membership relationship shown in Figure 3, the state transition matrix $P(R)$ is calculated.

**Step 3:** Solve Equation (4) and calculate the steady-state probability $\widehat{P}(\beta_i)$ of each risk class.

**Step 4:** According to Equation (5), the privacy security evaluation result of the whole mobile application is calculated.

**Step 5:** The evaluation results $E(\beta_i)$ of different risk categories are calculated respectively, and the calculation equation is as follows:

$$E(\beta_i) = \frac{\log_2 m}{-\sum_{j=1}^{m} P(X_{j,\beta_i}) \log_2 P(X_{j,\beta_i})} \tag{6}$$

In Equation (6), $m$ is the total number of risk factors included in each risk class, $P(X_j, \beta_i)$ represents the influence weight of risk factor $X_j$ on risk class $\beta_i$, which is the result of normalization treatment.

# 5 Case Study

## 5.1 Evaluation Process

In order to verify the proposed evaluation model, this paper selects three mobile application products with long operation time in the market. Among them, product A is a mobile application for financial business; product B is a mobile application for catering delivery; product C is a mobile application for map navigation. For these 3 applications, this paper carried out a detailed evaluation process, as follows:

**Step 1:** According to the definition of Table 2, a total of 10 experts were invited to evaluate the underlying risk factors of three different applications by AHP

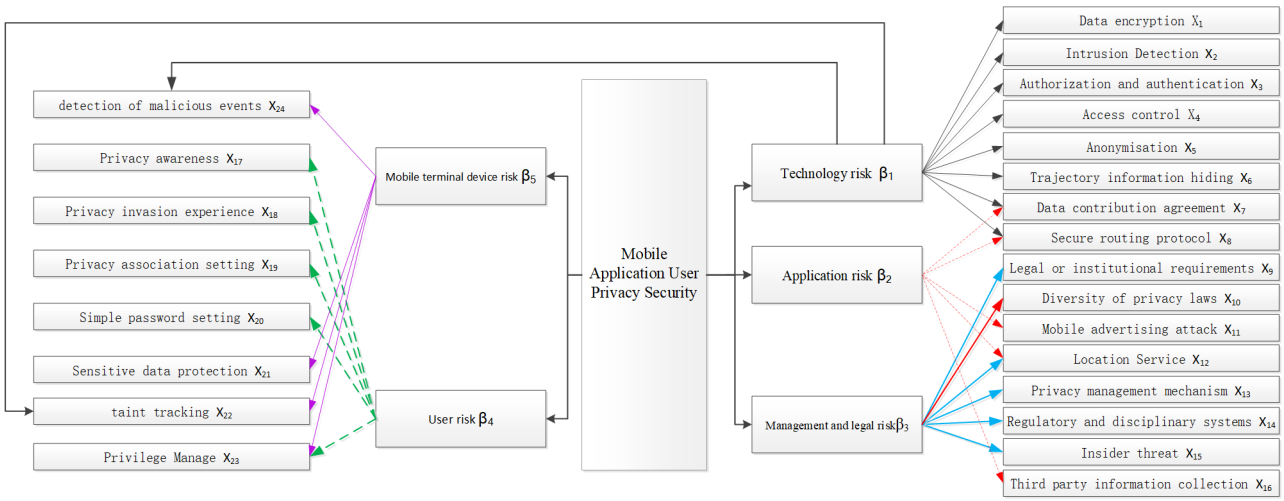


Figure 3: The privacy security risk attribute model of mobile application

method. Finally, the level of each risk factor is calculated, and the probability is obtained by further normalization. The results are shown in Table 3.

**Step2:** According to Equation (3) and the membership relationship shown in Figure 3, the privacy security risk state transition matrix $P^A(R), P^B(R), P^C(R)$ of the 3 applications are calculated as follows:

$$P^A(R) = \begin{bmatrix} 0.595 & 0.214 & 0.000 & 0.000 & 0.190 \\ 0.450 & 0.450 & 0.100 & 0.000 & 0.000 \\ 0.000 & 0.133 & 0.867 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.794 & 0.206 \\ 0.250 & 0.000 & 0.000 & 0.350 & 0.400 \end{bmatrix}$$

$$P^B(R) = \begin{bmatrix} 0.452 & 0.290 & 0.000 & 0.000 & 0.258 \\ 0.300 & 0.467 & 0.233 & 0.000 & 0.000 \\ 0.000 & 0.389 & 0.511 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.838 & 0.162 \\ 0.263 & 0.000 & 0.000 & 0.316 & 0.421 \end{bmatrix}$$

$$P^C(R) = \begin{bmatrix} 0.548 & 0.262 & 0.000 & 0.000 & 0.190 \\ 0.355 & 0.355 & 0.290 & 0.000 & 0.000 \\ 0.000 & 0.346 & 0.654 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.784 & 0.216 \\ 0.158 & 0.000 & 0.000 & 0.421 & 0.421 \end{bmatrix}$$

**Step 3:** The data of $P^A(R), P^B(R), P^C(R)$ are respectively substituted into the Equation (4), and the steady-state probability of each risk is calculated, as shown in Table 4.

**Step 4:** Substitute the results of table4 into Equation (5) to obtain the privacy security evaluation results of the 3 applications, the results are shown in Figure 4.

**Step 5:** According to Equation (6), the privacy security evaluation results of various risks are calculated respectively.

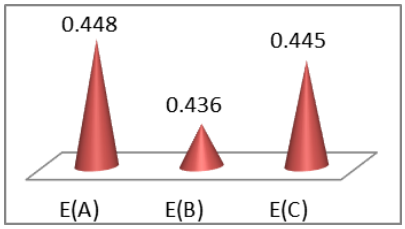


Figure 4: Comparison of evaluation results of three applications

Given that each risk class and its risk factors are as shown in Table 5, the result calculated by Equation (6) is shown in Figure 5.

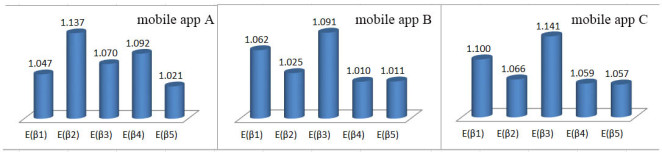


Figure 5: Comparison of entropy values of each risk class in three applications

## 5.2 Analysis of Evaluation Results

1) Analysis of top-level evaluation results. The comparison of Figure 4 shows that $E(A) > E(C) > E(B)$, which indicates that the financial business application A has the highest privacy security, and the catering delivery application B has the lowest privacy security. But from the data size difference comparison, it can be found that the privacy security performance of the 3 applications is not much different, indicating that the privacy security performance of the 3 applications is similar.

Table 3: Scoring results of probability of occurrence of underlying risk factors

| ***Mobile App*** | ***risk factor*** | **level** | $\boldsymbol{P}(\boldsymbol{x}_i)$ | ***risk factor*** | **level** | $\boldsymbol{P}(\boldsymbol{x}_i)$ | ***risk factor*** | **level** | $\boldsymbol{P}(\boldsymbol{x}_i)$ |
|---|---|---|---|---|---|---|---|---|---|
| A | $X_1$ | 2 | 1.905% | $X_9$ | 2 | 1.905% | $X_{17}$ | 9 | 8.571% |
| | $X_2$ | 8 | 7.619% | $X_{10}$ | 1 | 0.952% | $X_{18}$ | 8 | 7.619% |
| | $X_3$ | 7 | 6.667% | $X_{11}$ | 1 | 0.952% | $X_{19}$ | 9 | 8.571% |
| | $X_4$ | 4 | 3.810% | $X_{12}$ | 2 | 1.905% | $X_{20}$ | 1 | 0.952% |
| | $X_5$ | 2 | 1.905% | $X_{13}$ | 5 | 4.762% | $X_{21}$ | 5 | 4.762% |
| | $X_6$ | 2 | 1.905% | $X_{14}$ | 3 | 2.857% | $X_{22}$ | 4 | 3.810% |
| | $X_7$ | 5 | 4.762% | $X_{15}$ | 2 | 1.905% | $X_{23}$ | 7 | 6.667% |
| | $X_8$ | 4 | 3.810% | $X_{16}$ | 8 | 7.619% | $X_{24}$ | 4 | 3.810% |
| | | | | | | | | | |
| B | $X_1$ | 1 | 0.840% | $X_9$ | 2 | 1.681% | $X_{17}$ | 9 | 7.563% |
| | $X_2$ | 1 | 0.840% | $X_{10}$ | 2 | 1.681% | $X_{18}$ | 9 | 7.563% |
| | $X_3$ | 3 | 2.521% | $X_{11}$ | 3 | 2.521% | $X_{19}$ | 9 | 7.563% |
| | $X_4$ | 7 | 5.882% | $X_{12}$ | 9 | 7.563% | $X_{20}$ | 2 | 1.681% |
| | $X_5$ | 2 | 1.681% | $X_{13}$ | 8 | 6.723% | $X_{21}$ | 3 | 2.521% |
| | $X_6$ | 9 | 7.563% | $X_{14}$ | 3 | 2.521% | $X_{22}$ | 4 | 3.361% |
| | $X_7$ | 8 | 6.723% | $X_{15}$ | 2 | 1.681% | $X_{23}$ | 8 | 6.723% |
| | $X_8$ | 3 | 2.521% | $X_{16}$ | 8 | 6.723% | $X_{24}$ | 4 | 3.361% |
| | | | | | | | | | |
| C | $X_1$ | 1 | 0.901% | $X_9$ | 2 | 1.802% | $X_{17}$ | 9 | 8.108% |
| | $X_2$ | 1 | 0.901% | $X_{10}$ | 2 | 1.802% | $X_{18}$ | 9 | 8.108% |
| | $X_3$ | 3 | 2.703% | $X_{11}$ | 2 | 1.802% | $X_{19}$ | 9 | 8.108% |
| | $X_4$ | 2 | 1.802% | $X_{12}$ | 9 | 8.108% | $X_{20}$ | 2 | 1.802% |
| | $X_5$ | 1 | 0.901% | $X_{13}$ | 8 | 7.207% | $X_{21}$ | 3 | 2.703% |
| | $X_6$ | 9 | 8.108% | $X_{14}$ | 3 | 2.703% | $X_{22}$ | 4 | 3.604% |
| | $X_7$ | 8 | 7.207% | $X_{15}$ | 2 | 1.802% | $X_{23}$ | 8 | 7.207% |
| | $X_8$ | 3 | 2.703% | $X_{16}$ | 7 | 6.306% | $X_{24}$ | 4 | 3.604% |

Table 4: The steady-state probability of each risk

| ***Mobile App*** | $\widehat{\boldsymbol{P}}(\boldsymbol{\beta}_1)$ | $\widehat{\boldsymbol{P}}(\boldsymbol{\beta}_2)$ | $\widehat{\boldsymbol{P}}(\boldsymbol{\beta}_3)$ | $\widehat{\boldsymbol{P}}(\boldsymbol{\beta}_4)$ | $\widehat{\boldsymbol{P}}(\boldsymbol{\beta}_5)$ |
|---|---|---|---|---|---|
| A | 0.095 | 0.145 | 0.242 | 0.297 | 0.234 |
| B | 0.141 | 0.179 | 0.198 | 0.267 | 0.222 |
| C | 0.115 | 0.156 | 0.178 | 0.297 | 0.261 |

Table 5: The level of probability of risk factors occurrence

| **Risk class $\boldsymbol{\beta_i}$** | **risk factor $\boldsymbol{x_j}$ contained in $\boldsymbol{\beta_i}$** |
|---|---|
| $\beta_1$ | $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_{22}, X_{24}$ |
| $\beta_2$ | $X_7, X_8, X_{11}, X_{12}, X_{16}$ |
| $\beta_3$ | $X_9, X_{10}, X_{12}, X_{13}, X_{14}, X_{15}$ |
| $\beta_4$ | $X_{17}, X_{18}, X_{19}, X_{20}, X_{23}$ |
| $\beta_5$ | $X_{21}, X_{22}, X_{23}, X_{24}$ |

2) Comparative analysis of steady-state probability results of risk classes. It can be seen from Table 4 that among the three applications, the steady-state probability $\widehat{P}(\beta_4)$ is the largest, and $\widehat{P}(\beta_1)$ is the lowest. This result shows that the user risk $\beta_4$ is the most likely to cause privacy security risk in the long-term use of these 3 applications; on the contrary, technical risk $\beta_1$ is the least likely to occur.

3) Comparative analysis of privacy security evaluation result of each risk class. It can be seen from Figure 5 that in application B and Application C, the values of $E(\beta_4)$ and $E(\beta_5)$ are relatively low, indicating that for these two applications, the most difficult to control are the user's own risk $\beta_4$ and the terminal equipment risk $\beta_5$. These two risks are the main reasons for the privacy security of such applications. On the contrary, the value of $E(\beta_3)$ is the highest, which indicates that this kind of risk is basically controllable and is not easy to cause privacy security problems.

However, in application A, the values of $E(\beta_2)$ and $E(\beta_4)$ are higher, which indicates that financial applications have strict application standards and user behavior control. Compared with other types of applications, the application environment risk $\beta_2$ and user risk $\beta_4$ are easier to control in financial applications. The most important problem that leads to the privacy security of financial applications is focused on $\beta_3$ and $\beta_5$.

4) Analysis of risk factors layers. Through the above comparison, combined with table 5, it can be found that the user privacy risk $\beta_4$ and $\beta_5$ have the highest probability of occurrence in application A and application B. Observe the contained factors of risk class $\beta_4$ and $\beta_5$, it can be found that the security problems

of the most mobile applications are mainly caused by the factors $\{x_{17}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}, x_{24}\}$.

In the financial application A, risk class $\beta_2$ has the greater probability of occurrence. Observe the contained factors of risk class $\beta_2$, it can be found that the security problems of the financial application are mainly caused by the factors $\{x_7, x_8, x_9, x_{10}\}$

### 5.3 Countermeasures and Suggestions

To sum up, technical risk is not the main reason leading to the privacy security of mobile applications at this stage. To fundamentally improve the privacy security of applications, we need to shift the focus to the control of users' own risks and mobile terminal equipment risks. For the application providers, it is necessary to strengthen the management and control of user risk, remind users of the existing privacy security risks, standardize the user's operation behavior, and provide users with some suggestions for self-security protection; on the other hand, users themselves need to strengthen their own privacy security awareness and do their own security precautions. For financial applications, the application providers should further clarify the confidentiality agreement with users, reduce the access to users' use rights, and clarify the ownership of relevant responsibilities, and ensure the information security of users through laws and regulations.

### 5.4 Comparison with other Evaluation Methods

In order to explain the characteristics of the proposed method more intuitively, this paper compares the proposed method with AHP and CMM / CMMI. Among them, AHP (analytic hierarchy process) is a qualitative and quantitative evaluation method, which is easy to operate and has certain objectivity, but it is not suitable for the system with random state; CMM / CMMI is a set of evaluation method based on software process, and its evaluation results have long-term reference value for the management and improvement of the whole software process, but the establishment of its evaluation model needs a lot of manpower and financial resources.

Finally, this paper compares the 3 methods from four aspects which are usability, objectivity, decision support and cost. The results are shown in Table 6.

Table 6: Comparison with AHP and CMM/CMMI

| | Usability | Objectivity | Decision support | Cost |
|---|---|---|---|---|
| **Our Method** | High | High | Modest | Modest |
| **AHP** | Modest | Modest | Low | Low |
| **CMM/CMMI** | Low | Low | High | High |

In Table 6, usability refers to the ease of use of the method. Objectivity refers to the evaluation objective degree of the method. Decision support refers to the support degree of the method to the decision. Cost refers to the cost of adopting this method.

Through the comparison, it can be found that the method proposed in this paper has advantages in usability and objectivity. It is a simple, easy-to-use and moderate cost evaluation method, which can directly and truly reflect the privacy security degree of the evaluation object. Its evaluation results can provide practical data support for the privacy security protection of mobile applications.

## 6 Conclusions

In this paper, the privacy security attribute model of mobile applications is established, and the privacy security measurement is carried out based on information entropy, and the privacy risk environment is described based on Markov theory. A steady-state evaluation model of mobile application user privacy security based on information entropy and Markov is proposed. Finally, through the case study, the evaluation results show that the model can achieve multi-level and multi-dimensional analysis of mobile application privacy security, so as to provide the basis for privacy protection of mobile applications. The model is simple, easy to use, and has good objectivity, which is of great significance to the research on privacy security of mobile application. In the future research, the attribute model of risk is a content that needs to be further studied. Only by constantly improving the attribute model of risk can we provide more accurate reference for decision-making. On the other hand, it is necessary to strengthen the dynamic evaluation of security, and introduce the concept of time to dynamically describe the privacy security of mobile applications, so as to provide more realistic security evaluation results for decision makers.

## Acknowledgments

This research is supported by National Natural Science Foundation Project (No. 71462036), the Scientific Research Foundation of Yunnan Education Department (No. 2020J0377, No. 2020J0392) and School-level Project of Yunnan University of Finance and Economics (No. 2017D29).

The authors would like to thank the anonymous reviewers and the editors for their suggestions.

## References

[1] E. Aghasian, S. Garg, and J. Montgomery, "A privacy-enhanced friending approach for users on multiple online social networks," *Computers*, vol. 7, no. 3, pp. 42, 2018.

[2] M. H. R. Al-Shaikhly, H. M. El-Bakry, and A. A. Saleh, "Cloud security using Markov chain and genetic algorithm," *International Journal of Electronics and Information Engineering*, vol. 8, no. 2, pp. 96–106, 2018.

[3] G. O. A. Ampong, A. Mensah, A. S. Y. Adu, J. A. Addae, O. K. Omoregie, and K. S. Ofori, "Examining self-disclosure on social networking sites: A flow theory and privacy perspective," *Behavioral Sciences*, vol. 8, no. 6, pp. 58, 2018.

[4] Y. Ashibani and Q. H. Mahmoud, "Cyber physical systems security: Analysis, challenges and solutions," *Computers & Security*, vol. 68, pp. 81–97, 2017.

[5] B. Chen and Z. Wu, "Research on personal privacy protection in big data environment," *Journal of Jilin Normal University of engineering and technology*, vol. 35, no. 8, pp. 68–70, 2019.

[6] T. T. El-Khazendar and T. S. M. Barhoom, "Protect local data on personal devices: Third party application," 2015.

[7] M. Fodor and A. Brem, "Do privacy concerns matter for millennials? Results from an empirical analysis of location-based services adoption in germany," *Computers in Human Behavior*, vol. 53, pp. 344–353, 2015.

[8] T. Gao, T. Li, R. Jiang, M. Yang, and R. Zhu, "Research on cloud service security measurement based on information entropy.," *International Journal Network Security*, vol. 21, no. 6, pp. 1003–1013, 2019.

[9] A. Gutierrez, S. O'Leary, N. P. Rana, Y. K. Dwivedi, and T. Calle, "Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: Identifying intrusiveness as the critical risk factor," *Computers in Human Behavior*, vol. 95, pp. 295–306, 2019.

[10] B. N. Jagdale and J. W. Bakal, "Controlled broadcast protocol for location privacy in mobile applications," *Procedia Computer Science*, vol. 78, pp. 782–789, 2016.

[11] M. J. Keith, S. C. Thompson, J. Hale, P. B. Lowry, and C. Greer, "Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior," *International Journal of Human-Computer Studies*, vol. 71, no. 12, pp. 1163–1173, 2013.

[12] C. T. Li, C. C. Yang, and M. S. Hwang, "A secure routing protocol with node selfishness resistance in manets," *International Journal of Mobile Communications*, vol. 10, no. 1, pp. 103–118, 2012.

[13] H. LI, B. Wang, W. Zhang, Q. Tang, and Y. Zhang, "X-Decaf: Detection of cache file leaks in android social apps," *Journal of Electronics & Information Technology*, vol. 39, no. 1, pp. 10, 2017.

[14] Z. Li, Y. Tian, W. Zhang, and Y. Liu, "Research on china mobile application privacy policy," *Cyberspace Security*, vol. 11, no. 6, pp. 11, 2020.

[15] N. W. Lo, K. H. Yeh, and C. Y. Fan, "Leakage detection and risk assessment on privacy for android applications: Lrpdroid," *IEEE Systems Journal*, vol. 10, no. 4, pp. 1361–1369, 2014.

[16] Y. Nan, Z. Yang, M. Yang, S. Zhou, Y. Zhang, G. Gu, X. Wang, and L. Sun, "Identifying user-input privacy in mobile applications at a large scale," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 647–661, 2016.

[17] A. Ruiz-Heras, P. García-Teodoro, and L. Sánchez-Casado, "Adroid: Anomaly-based detection of malicious events in android platforms," *International Journal of Information Security*, vol. 16, no. 4, pp. 371–384, 2017.

[18] F. Shirazi and A. Iqbal, "Community clouds within m-commerce: A privacy by design perspective," *Journal of Cloud Computing*, vol. 6, no. 1, pp. 22, 2017.

[19] Y. A. Tan, Y. Xue, C. Liang, J. Zheng, Q. Zhang, J. Zheng, and Y. Li, "A root privilege management scheme with revocable authorization for android devices," *Journal of Network and Computer Applications*, vol. 107, pp. 69–82, 2018.

[20] B. Tian, Y. Zheng, P. Liu, and C. Li, "The evaluation index and empirical study on risk of privacy informationdisclosure of mobile app users," *Library and Information Service*, vol. 62, no. 19, pp. 101–110, 2018.

[21] G. Tilei, L. Tong, Y. Ming, and J. Rong, "Research on a trustworthiness measurement method of cloud service construction processes based on information entropy," *Entropy*, vol. 21, no. 5, pp. 462, 2019.

[22] J. Wang, J. Liu, and H. Zhang, "Access control based resource allocation in cloud computing environment.," *International Journal Network Security*, vol. 19, no. 2, pp. 236–243, 2017.

[23] C. H. Wei, M. S. Hwang, and A. Y. H. Chin, "An improved authentication protocol for mobile agent device in rfid environment," *International Journal of Mobile Communications*, vol. 10, no. 5, pp. 508–520, 2012.

[24] Y. C. Wei, W. C. Wu, G. H. Lai, and Y. C. Chu, "pISRA: Privacy considered information security risk assessment model," *The Journal of Supercomputing*, vol. 76, no. 3, pp. 1468–1481, 2020.

[25] V. M. Wottrich, E. A. van Reijmersdal, and E. G. Smit, "The privacy trade-off for mobile app downloads: The roles of app value, intrusiveness, and privacy concerns," *Decision Support Systems*, vol. 106, pp. 44–52, 2018.

[26] Y. Wu, G. Feng, N. Wang, and H. Liang, "Game of information security investment: Impact of attack types and network vulnerability," *Expert Systems with Applications*, vol. 42, no. 15-16, pp. 6132–6146, 2015.

[27] M. Xiang, X. Wang, R. Jia, and L. Wang, "The evaluation index and empirical study on risk of privacy informationdisclosure of mobile app users," *Library and Information Service*, vol. 64, no. 18, pp. 34–44, 2018.

[28] Y. Z. Xu, J. L. Zhang, Y. Hua, and L. Y. Wang, "Dynamic credit risk evaluation method for e-commerce sellers based on a hybrid artificial intelligence model," *Sustainability*, vol. 11, no. 19, pp. 5521, 2019.

[29] Y. Xu and Y. Ji, "Android application risk assessment method based on permission," *Computer Applications and Software*, vol. 37, no. 4, pp. 69–74+100, 2020.
[30] Y. Xu and Y. Ji, "An android application risk assessment strategy based on permission characteristics," *Computer Applications and Software*, vol. 37, no. 4, pp. 69–74,100, 2020.
[31] M. Yang, R. Jiang, T. Gao, W. Xie, and J. Wang, "Research on cloud computing security risk assessment based on information entropy and markov chain," *International Journal Network Security*, vol. 20, no. 4, pp. 664–673, 2018.
[32] L. Ying, "Research on personal privacy protection in big data environment," *Computers and Networks*, vol. 46, no. 2, pp. 68–70, 2020.
[33] G. Zhu, M. N. Feng, Y. Chen, and J. Y. Yang, "Research on fuzzy evaluation of privacy risk for social network in big data environment," *Information Science*, vol. 34, no. 9, pp. 19, 2016.
[34] H. Zhu, C. X. J. Ou, W. J. A. M. van den Heuvel, and H. Liu, "Privacy calculus and its utility for personalization services in e-commerce: An analysis of consumer decision-making," *Information & Management*, vol. 54, no. 4, pp. 427–437, 2017.

## Biography

**Ming Yang** is an associate professor at the school of information, Yunnan University of Finance and Economics, China. He received his Ph.D. in system analysis and integration from the school of software at Yunnan University. His main research interests include information management and data mining.

**Li Jia** is an associate professor at the school of information, Yunnan University of Finance and Economics, China. His main research fields are network communication and security control, data mining technology.

**Tilei Gao** is a lecturer at the school of information, Yunnan University of Finance and Economics. He is also a Ph.D candidate in system analysis and integration at the school of software at Yunnan University. His main research interests include software engineering and information management.

**Tao Zhang** received the M.A. degrees in Information Management and System from School of Information, Yunnan University of Finance and Economics, China, in 2012. At present, he is a lecturer at the School of Information, Yunnan University of Finance and Economics. Also he is a Ph.D. candidate. His research interests include Information Security, Information Management and Information System, Recommendation system.

**Wanyu Xie** is a lecturer at Kunming Metallurgy College. She received Her master's degree in computer science and technology from the school of information science and engineering at Yunnan University. Her main research interests is information management.