Postprint: Research on Personal Data Privacy Measurement in the Context of Leakage Probability

Authors: Kailiang Zhang, Zang Guoquan

Date: 2023-04-01T16:02:50+00:00

Abstract

[Purpose/Significance] Compared with scenarios of certain leakage, situations where personal data is leaked with a certain probability are more prevalent. This study measures users' personal privacy value in probabilistic leakage scenarios, proposes a novel perspective for privacy valuation, and the measurement results also hold practical significance for tiered privacy protection. [Method/Process] Based on the multi-level price list method, we measure users' financial risk return rates; leveraging users' financial risk return rates, we modify the implementation mechanism of the multi-level price list to guide users in making decisions between schemes with no privacy leakage probability and schemes with privacy leakage probability, thereby measuring users' valuation of their personal data under specific leakage probability scenarios. [Results/Conclusion] In a scenario with a 30% leakage probability, users' perceived value of personal data privacy in their social networks is approximately 89.5 RMB; concurrently, in a scenario with a 100% leakage probability, the personal data privacy values manifested through willingness to accept and willingness to pay are 124.1 and 93.8 RMB, respectively. This demonstrates that users' cognition of personal data privacy value in probabilistic leakage scenarios depends on two dimensions: the intrinsic value of personal data privacy itself and users' risk tolerance towards leakage probability.

Full Text

Preamble

Measurement of Personal Data Privacy in the Context of Leakage Probability

Zhang Kailiang[1], **Zang Guoquan**[2],[3] [1]School of Politics and Public Administration, Zhengzhou University, Zhengzhou 450001 [2]School of Information Man-

agement, Zhengzhou University, Zhengzhou 450001 [3]Research Institute of Data Science, Zhengzhou City, Zhengzhou 450001

Abstract: [Purpose/Significance] Compared with scenarios of certain leakage, situations where personal data is leaked with a certain probability are more common. This study measures users' privacy value of personal data in probabilistic leakage contexts, proposing a new perspective for privacy measurement. The results have practical significance for privacy classification and protection. [Method/Process] Based on the multiple price list method, we measure users' financial risk return rate. Using this rate, we modify the implementation mechanism of the multiple price list to guide users in making decisions between schemes with and without privacy leakage probability, thereby measuring users' valuation of their personal data under specific leakage probabilities. [Result/Conclusion] In a scenario with a 30% leakage probability, users' valuation of personal data privacy in social networks is approximately 89.5 RMB. Meanwhile, in a scenario with 100% leakage probability, the personal data privacy value expressed as willingness to accept and willingness to pay is 124.1 RMB and 93.8 RMB respectively. This indicates that users' valuation of personal data privacy under probabilistic leakage depends on both the intrinsic privacy value and their tolerance for leakage probability risk.

Keywords: personal privacy; leakage probability; privacy value; multiple price list mechanism **Classification Number:** G251 **DOI:** 10.13266/j.issn.0252-3116.2021.09.007

---

2. Related Research on Personal Data Privacy Valuation

Existing empirical research on personal data privacy valuation primarily measures the monetary value of privacy, expressed in two forms: (1) Willingness to Pay (WTP) - the monetary price users are willing to pay to protect their personal data privacy from infringement. For example, S. Egelman et al. [17] used discrete choice analysis to measure that users' WTP for location data and call records is approximately 1.5 USD. J. Kim et al. [13] used conjoint analysis to find that the Korean public is willing to spend about 6.8 USD per month on personal information protection services. (2) Willingness to Accept (WTA) - the monetary compensation users are willing to accept for sacrificing their personal data privacy. For instance, Zang Guoquan et al. [11] measured using a modified BDM mechanism that users' WTA for personal preference data and contact data is approximately 38.8 RMB and 136.3 RMB respectively. Huang Yiyun et al. [9] used the contingent valuation method to survey that users' WTA for age data and home address is approximately 6.99 RMB and 66.19 RMB respectively.

According to Table 1 , existing personal data privacy valuation methods can be divided into two categories: questionnaire surveys and auction experiments. Questionnaire surveys include conjoint analysis, discrete choice analysis, and contingent valuation method. Conjoint analysis and discrete choice analysis

both construct product profiles based on product attributes and levels, using orthogonal methods to form a certain number of product sets, but their evaluation methods differ. Conjoint analysis uses scoring and ranking methods, while discrete choice analysis asks users to actually choose which product or service they prefer. The contingent valuation method uses hypothetical scenarios to directly ask users about their WTP or WTA for personal data privacy. Auction experiments include n-level price auction (where n-1 highest bidders win and pay the nth highest price), Becker-DeGroot-Marschak (BDM) auction (where the bidder competes with a random function; if the bidder's price is higher than the random function's return price, the bidder wins and pays the random function's return price), and random N-level price auction (where bidder i competes with N-1 other participants, and a bid from N bidders is randomly selected as the market clearing price; bidders with bids greater than or equal to this price win and pay that price). Among these, n-level price auction and random N-level price auction are suitable for auctions of multiple identical items, while BDM auction is suitable for single-item auctions.

A comprehensive review of the above empirical studies reveals that existing research assumes certain privacy leakage and does not consider situations where leakage occurs only with a certain probability. In reality, privacy leakage follows a random process, and privacy decisions depend on both the value of personal content privacy and the leakage probability. Therefore, this study aims to measure users' valuation of their personal data privacy under probabilistic leakage conditions, referred to as the Value of Privacy under Leakage Probability (VPLP).

---

3. Measurement Model

Risk is defined as the probability of an undesired event occurring [23]. Therefore, risk comprises two elements: the undesired event and its probability of occurrence. Personal data privacy leakage is a risk event because data leakage is an undesired event for data subjects, and leakage itself is a probabilistic event. In personal data storage systems, personal data may or may not be leaked, depending on security protection measures. The stricter the protection measures, the lower the leakage probability; conversely, the higher the leakage probability. This study uses "privacy leakage risk" to explicitly express "the probability of privacy leakage events occurring."

This research adopts the Multiple Price List (MPL) mechanism, designing an easily understandable financial probabilistic risk multiple price list to measure users' Return on Financial Risk (RFR). Based on users' measured RFR, we then measure their valuation of personal data privacy under probabilistic leakage conditions.

3.1 Financial Risk Return Rate

The MPL mechanism is a common method in experimental economics for measuring user risk attitudes, first proposed by H.P. Binswanger [24]. It presents users with a series of different payoff combinations, each containing a safe option (with no probability of loss) and a risky option (with a probability of loss). The safe option's fixed payoff decreases row by row, while all risky options have the same expected payoff. By asking users at which level of decreasing safe payoff they would abandon the safe option to pursue the potentially higher risky payoff, we can indirectly infer their risk attitude through their decision-making.

Using the MPL mechanism, we design an RFR multiple price list as shown in Table 2 . Table 2 contains two options: Option A is the safe option with payoffs decreasing by a yuan each row. In row i, Option A's payoff is Y(A) = x - i • a, where i [0, k]. Option B is the risky option (with probability p of losing c yuan), with each row's expected payoff being E(B) = y • (1-p) + (y-c) • p. When a user switches from Option A to Option B at row i, this indicates the user is willing to abandon the safe payoff for the risky payoff. At this point, the user's RFR is:

$$RFR = \frac{E(B_i) - Y(A_i)}{Y(A_i)}, \quad i \in [0, k]$$

3.2 Leakage Probability Privacy Value Measurement

By modifying Option B in the RFR multiple price list (Table 2) from "with probability p of losing c yuan" to "with probability p of personal privacy leakage," we can simulate a realistic leakage probability scenario, creating the VPLP multiple price list (Table 3). Privacy value represents the privacy subject's estimate of potential losses when privacy objects are violated, including both personal and property aspects (with the property conversion of personal losses becoming increasingly common due to economic and social development). This can be expressed in monetary terms to represent users' valuation of such privacy.

Therefore, in the VPLP multiple price list, Option B can be monetarily expressed as "You can receive y yuan, but with probability p of losing *vplp* yuan." Table 3 shows the VPLP multiple price list design.

Based on Table 3's design, users' privacy decisions are transformed into decisions between safe and risky options in a financial context. In fact, Table 2 represents a direct financial scenario (all financial elements expressed directly in monetary form), while Table 3 represents an indirect financial scenario (VPLP expressed indirectly through user value cognition). Both belong to the financial domain. When Tables 2 and 3 use the same loss probability, users' financial risk return rates should be identical in both contexts.

This leads to the VPLP measurement formula:

$$VPLP = y - (RFR + 1) \cdot (x - i \cdot a), \quad i \in [0, k]$$

where i is the row number where the user switches from Option A to Option B in Table 3, and RFR is the user's financial probability risk return rate calculated based on Table 2.

---

4. Measurement Experiment

4.1 Personal Data Samples

This experiment obtains personal basic data and personal privacy data samples through questionnaires. The questionnaire should contain three parts:

4.1.1 Demographic Data: Including user name, gender, age, education level, and other personal identification and semi-identification data, used for analyzing leakage probability privacy values based on demographic characteristics.

4.1.2 Personal Privacy Data: The model is suitable for any domain of personal privacy data, whether macro-level privacy data entries or micro-level privacy data items. Since the experiment may require users to disclose personal privacy data, the surveyed personal privacy data must be readily available on-site. Based on this requirement, this experiment selects users' WeChat data (including screenshots of WeChat account pages, recently posted Moments pages, chat main interface, etc.) and Weibo data (including screenshots of Weibo account pages, recently posted Weibo pages, recent Weibo comment interfaces, etc.) as user personal privacy data samples.

4.1.3 Direct Cognitive Data on Personal Privacy Value: This aims to obtain users' direct reported values of their personal privacy value, including the price users are willing to pay to protect their personal privacy data (WTP) and the compensation price users are willing to accept when disclosing personal privacy data (WTA).

4.2 Experimental Steps

4.2.1 Step 1: Measure RFR: First, the experimenter distributes an RFR multiple price list to users, containing K+1 rows, each with safe Option A and risky Option B. Users must make a choice between the two options in each row. Second, record the initial row number where the user's decision switches from Option A to Option B, and calculate the user's RFR. Finally, use a random function to return a row number i between 0 and k, and pay the user experimental compensation based on their choice in that row. If the user chose Option A in that row, pay (x - $i \cdot a$) yuan. If they chose Option B, use a [0,1] random function to determine Option B's outcome. If the random function value is greater than p, pay y yuan; otherwise, pay (y - c) yuan. (The experimenter

informs users in the experimental agreement that, due to cost considerations, both Step 1 and Step 2 pay actual amounts at a 50:1 ratio.)

In the RFR experiment, using a random function to determine the row number and paying compensation based on the user's choice in that row ensures the sincerity of user decisions. This method originates from the Becker-DeGroot-Marschak (BDM) auction mechanism [25]. If users' actual choices differ from their true preferences, and the random function returns a row number i within their "lying interval," users must bear the expected payoff difference between Options A and B. Only when users' actual choices align with their true preferences do they avoid additional risk. Therefore, this method can measure users' true RFR. The same principle applies to Step 2.

4.2.2 Step 2: Measure VPLP: First, the experimenter distributes a VPLP multiple price list to users, containing K+1 rows, each with safe Option A and risky Option B. Users must make a choice between the two options in each row. Second, record the initial row number where the user's decision switches from Option A to Option B, and calculate their VPLP using the user's RFR. Finally, use a random function to return a row number i between 0 and k, and pay the user experimental compensation based on their choice in that row. If the user chose Option A in that row, pay (x - $i \cdot a$) yuan. If they chose Option B, use a [0,1] random function to determine Option B's outcome. If the random function value is greater than p, pay y yuan without disclosing the user's personal privacy data; otherwise, pay y yuan but disclose their personal privacy data to all users within the experimental group (this both simulates real privacy leakage scenarios and controls the scope of privacy leakage to avoid actual privacy breaches).

4.3 Pre-Experiment

Based on the target user quantity for the formal experiment and following sample size empirical rules [26], we calculate the pre-experiment user quantity. This pre-experiment only conducts VPLP measurement to determine relevant parameters. Once determined, these parameters are used for both RFR and VPLP formal experiments.

4.3.1 Parameter x Determination: Parameter x should be greater than or equal to parameter y to ensure all users' decisions involve a switch from Option A to Option B, increasing experimental result credibility. Otherwise, risk-averse users with low personal privacy value might choose Option B in all rows, making it impossible to measure financial risk return rate and thus unable to measure leakage probability privacy value. The simplest approach is to set parameter x equal to parameter y.

4.3.2 Parameter y Determination: According to prospect theory [27], when two options have equal expected payoffs but one involves negative user payoff, users' aversion to this risk (negative payoff) reduces their willingness to choose that option. Therefore, parameter y should be greater than the estimated VPLP to avoid situations where users' monetary gains from choosing Option B are in-

sufficient to compensate for personal privacy disclosure losses, thereby reducing their willingness to choose Option B. In previous research using BDM auction mechanisms, users' WTA for online privacy data was measured to be between 38.8-237.9 RMB [11]. Therefore, based on collected personal privacy samples, this experiment sets the preliminary parameter y at 150 RMB. Pre-experiment results can verify whether this parameter setting is reasonable. If most users never choose Option B or only choose it in very few rows in the pre-experiment, the y value is too low and should be increased. Conversely, it indicates the y value is reasonable. Based on pre-experiment results, when $y = 150$, 96.7% of users chose Option B at some point, with an average of 13 rows where Option B was selected, confirming that setting y at 150 RMB is reasonable.

4.3.3 Parameters *a* and *k* Determination: Parameter a only reflects the decreasing rate of Option A's payoff and does not significantly affect experimental results. However, a cannot be too large or too small. If a is too large, the payoff difference between adjacent rows is excessive, potentially preventing accurate observation of user decision behavior. If a is too small, the multiple price list has too many rows, potentially affecting user decision experience. Based on the initial y value, this study sets parameter a at 5. Parameter k determines the number of rows in the multiple price list. Only when Option A's expected payoff in the last row is much smaller than Option B's can a switch from Option A to Option B occur. Therefore, based on parameter a, this study sets parameter k at 19 (total rows $= k + 1 = 20$ rows), making Option A's payoff in the last row 55 RMB.

4.3.4 Probability *p* Determination: Setting probability p should consider real-world privacy leakage conditions. The ideal approach would be to calculate the actual privacy leakage probability by statistically analyzing total personal privacy data volume and leaked personal privacy data volume nationwide. However, after extensive searches, no relevant authoritative statistical data was found. As a compromise, this study uses pre-experiments to observe user decision results (which are credible judgments based on users' own experiences and perceptions) to determine p value. Four groups of pre-experiments were conducted with leakage probability p set at 20%, 30%, 40%, and 50% respectively (probabilities that are too high or low may be unrealistic), with other parameters identical across groups. Based on pre-experiment result distributions, when $p = 30\%$, the distribution of users' switching row numbers in the VPLP multiple price list approximates a normal distribution, making $p = 30\%$ a reasonable choice.

4.4 Formal Experiment

All formal experiments are conducted through self-developed software implemented using Unity tools and C# language. The same user must complete both the RFR measurement formal experiment and the VPLP measurement formal experiment. A software example is shown in Figure 1 [Figure 1: see original paper].

The software design prevents abnormal experimental behavior. First, when users select Option B in all rows of the multiple price list, the system automatically rejects their experimental submission. Second, when making decisions, users only need to click Option B in a specific row; the system automatically selects Option A for all rows above and Option B for all rows below. These settings save experimental time, improve efficiency, and prevent anomalous data.

---

5. Sample Statistics and Data Analysis

5.1 Sample Statistics

This experiment was conducted in August 2020. Ten experimenters, including the author and master's students in Information Science from Zhengzhou University, administered the experiment. The pre-experiment used on-site experiments, while the formal experiment used a combination of online groups and email. Due to the involvement of risk tolerance and decision-making across two tables, respondents needed certain knowledge levels. Therefore, this experiment selected university students and graduate students from Zhengzhou University as participants. Before participating, users carefully read the experimental procedures. Users who agreed to participate signed an experimental agreement and received 5 RMB compensation.

A total of 275 people participated in this experiment. Descriptive statistics of the user sample are shown in Table 4 , and experimental results are shown in Table 5 .

(1) Users' RFR: Users' RFR is calculated based on the safe fixed payoff of Option A and the expected payoff of Option B when users' choices in the RFR multiple price list begin to switch from safe Option A to risky Option B. According to the RFR calculation formula, if RFR > 0, it indicates that users only choose risky Option B when its expected payoff exceeds Option A's fixed payoff, showing that users are unwilling to accept expected payoff differences where Option B's expected payoff is lower than Option A's fixed payoff—meaning these users are risk-averse. Similarly, if RFR < 0, users are risk-seeking; if RFR $= 0$, users are risk-neutral. Therefore, RFR measurement results show that 12% of users are risk-neutral (RFR $= 0$), making decisions based solely on expected payoff. 82.5% of users are risk-averse (RFR > 0), with 3.3% being extremely risk-averse (RFR $= 1.18$, consistently choosing Option A from the initial row). 5.5% of users are risk-seeking (RFR < 0), though no user consistently chose Option B from the initial row. The reason no user consistently chose Option B from the initial row is that in the initial row, Option B's maximum payoff does not exceed Option A's safe payoff, making it unnecessary for users to bear Option B's risk. Consistently choosing Option B from the initial row would be irrational (the experimental software design prevents this situation).

(2) Users' VPLP: Experimental results show that users with identical RFR

may have different VPLP values. Users with lower RFR may exhibit higher VPLP, and users with higher RFR may exhibit lower VPLP. Similarly, users with identical WTA or WTP may have different VPLP values. This comprehensively demonstrates that users' VPLP depends on both personal privacy value itself (under certain privacy leakage conditions) and leakage probability factors.

5.2 Correlation Analysis Between VPLP and Demographic Characteristics

Through K-S normality testing, VPLP values follow a normal distribution (K-S test, $p = 0.018$). Therefore, independent samples t-test can be used for correlation analysis between VPLP and demographic characteristics.

5.2.1 Gender: Male users' mean VPLP is 81.1, while female users' is 97.5. Using user gender as the grouping variable for independent samples t-test on VPLP values, results show that female users' VPLP is significantly higher than males' (independent samples t-test, $p < 0.001$). Research in user risk attitude shows that females' risk aversion is significantly higher than males' [28]. In privacy measurement, existing studies show that females' valuation of their personal privacy is significantly higher than males' [7, 19]. Since VPLP represents users' comprehensive cognition of both leakage probability and personal privacy value, it is reasonable that female users' VPLP values are higher than males'.

Age: Since the overall sample's age span is not large, age intervals cannot be used for grouping. This study only selects some representative users (with certain age gaps) for testing. For example, 27-28-year-old students' VPLP is significantly higher than 18-19-year-old students (independent samples t-test, $p = 0.014$). Existing research shows that users' risk aversion gradually increases with age [28]. Future research should expand sample size to analyze the impact of user age on VPLP.

5.2.2 Education Level: The mean VPLP for undergraduate, master's, and doctoral students is 83.3, 91.6, and 96.8 respectively. Using education level as the grouping variable for pairwise independent samples t-tests, results are shown in Table 6 .

According to Table 6, doctoral students' VPLP is higher than master's students, and master's students' VPLP is higher than undergraduates, but the differences are not significant except that doctoral students' VPLP is significantly higher than undergraduates. This may be related to the user sample selected, as all participants were university students with limited age spans between education levels and age overlaps (e.g., some master's students being younger than undergraduates, some doctoral students being younger than master's students). Existing research shows that education level significantly affects users' valuation of their personal privacy [9, 20]. Future research should expand sample size to analyze the impact of education level on VPLP.

5.3 Correlation Analysis Between VPLP and Questionnaire Data

Correlation analysis between VPLP, WTP, and WTA: The mean values of VPLP, WTP, and WTA are 89.5, 93.8, and 124.1 respectively. Using paired samples t-test, VPLP is correlated with WTP and WTA respectively. Test results show that WTA is significantly higher than WTP (paired samples t-test, $p < 0.001$), and WTP is significantly higher than VPLP (paired samples t-test, $p = 0.022$).

These results indicate that both WTA and WTP are significantly higher than VPLP because WTA and WTP represent personal privacy value cognition under 100% leakage probability, while VPLP represents personal privacy value cognition under 30% leakage probability. With other conditions unchanged, when privacy leakage probability decreases from 100% to 30%, users' valuation of their personal privacy also decreases. Therefore, in personal data storage systems, strict privacy protection measures should be implemented to reduce privacy leakage probability and alleviate users' privacy concerns. Additionally, the significant difference between WTA and WTP demonstrates the widespread existence of the "privacy paradox" phenomenon—users are concerned about personal privacy but are unwilling to make greater efforts to protect it.

This study measures subjects' average VPLP at approximately 89.5 RMB, with undergraduate, master's, and doctoral students averaging approximately 83.3 RMB, 91.6 RMB, and 96.8 RMB respectively. Compared with certain leakage scenarios, users' valuation of personal data privacy is significantly lower under probabilistic leakage scenarios. In the VPLP experiment, 2.5% of users consistently refused to choose Option B, possibly due to extremely high risk aversion to privacy leakage probability, excessively high personal privacy value cognition, or a combination of both.

The leakage probability privacy value measurement model designed in this study has certain universality for measuring different types of personal privacy values. Different enterprises and institutions can select different types of personal privacy data for measurement based on their usage purposes. For example, online marketing companies may focus more on user behavior data for analyzing shopping preferences, while courts may focus more on personal financial data that tends to generate disputes. Note that different types of personal privacy data have different values, and different types of personal privacy data also have different leakage probabilities. When using this method, experimental parameters should be set reasonably based on personal privacy data type.

Advantages of this study: (1) **Implicitness**: The experiment indirectly infers leakage probability privacy values from user behaviors rather than direct user reports. Research shows that stated preferences often differ from observed behaviors (e.g., the "privacy paradox") [29]. Inferring preferences from observed user choices, even in relatively artificial laboratory settings, is closer to users' true cognition than direct reports in hypothetical scenarios. Moreover, direct privacy value measurement (such as WTA, WTP) forces users to consciously se-

lect answers, which is unreliable because users may struggle to accurately assess privacy-related risks and losses. Studies [30-32] show that indirect methods outperform direct surveys for privacy value measurement. (2) **Simultaneous consideration of personal privacy content value and leakage probability**: The significant difference between this study's leakage probability privacy value measurement model and existing research is that the former presents probabilistic privacy leakage risk, while the latter shows certain privacy leakage threats. Since privacy leakage follows a random process, both personal content privacy value and leakage probability are crucial for privacy decisions. In reality, people must decide how much to invest in protecting their information from uncertain random threats. Privacy preferences elicited under certain threat scenarios may not fully align with behaviors under random threat scenarios because the latter involves people's risk aversion levels. Therefore, compared with certain privacy leakage, the random risk nature of privacy disclosure is more realistic, objective, and research-valuable.

Limitations: (1) **Privacy leakage probability value**: Due to the lack of authoritative statistical data on privacy leakage, this study sets the privacy leakage probability at 30% based on pre-experiment result distributions, which may involve experimental error. (2) **User sample**: For convenience, this study selected only university students from Zhengzhou University, lacking geographical breadth and demographic diversity. Future research should expand geographical coverage to avoid macro factors like regional economy and culture, and expand social coverage to analyze the impact of demographic characteristics (especially education level, age, occupation, etc.) on VPLP.

References

[1] Xinhua Net. Huazhu's 500 million user information suspected of leakage, police have launched investigation [EB/OL]. [2020-11-12]. http://www.xinhuanet.com/fortune/2018-08/29/c_{1123343927}.htm. [2] Sohu Net. LinkedIn hacked, about 159 million users' data stolen [EB/OL]. [2020-11-12]. https://m.sohu.com/a/292645813_{557054}. [3] HANN I H, HUI K L, LEE T, et al. Online information privacy: measuring the cost-benefit tradeoff [C]//Proceedings of the international conference on information systems 2002. Illinois: AISeLibrary, 2002. [4] ACQUISTI A, GROSSKLAGS J. An online survey experiment on ambiguity and privacy [J]. Communications & strategies, 2012, 1(88): 19-39. [5] POTOGLIOU D, PATIL S, GIJON C, et al. The value of personal information online: results from three stated preference discrete choice experiments in the UK [C]//Proceedings of the 21st European conference on information systems. Netherlands: ECIS Press, 2013. [6] LIM S, WOO J R, LEE J, et al. Consumer valuation of personal information in the age of big data [J]. Journal of the Association for Information Science and Technology, 2018, 69(1): 60-71. [7] DENG Shengli, ZHAO Haiping. Personal information value assessment and individual differences in information leakage contexts: an empirical study based

on discrete choice models [J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(3): 266-276. [8] GROSSKLAGS J, ACQUISTI A, HEINZ H J. When 25 cents is too much: an experiment on willingness-to-sell and willingness-to-protect personal information [C]//Proceedings of the 6th annual workshop on the economics of information security. Cambridge: WEIS Press, 2007. [9] HUANG Yiyun, LU Tong, YAN Qiang. Evaluation of personal information value on e-commerce websites [J]. Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition), 2017(5): 33-41. [10] HUBERMAN B A, ADAR E, FINE L R. Valuing privacy [J]. IEEE security & privacy, 2005, 3(5): 22-25. [11] ZANG Guoquan, ZHANG Kailiang, YAN Li. Research on personal data value measurement: based on modified BDM mechanism [J]. Library and Information Service, 2020, 64(7): 103-109. [12] KRASNOVA H, HILDEBRAND T, GUNTHER O. Investigating the value of privacy in online social networks: conjoint analysis [C]//Proceedings of the international conference on information systems. Phoenix: ICIS Press, 2009. [13] KIM J, NAM C, KIM S. The economic value of personal information and policy implication [C]//Proceedings of the 26th European regional ITS conference. Los Angeles: ITS Press, 2015. [14] PU Y, GROSSKLAGS J. Towards a model on the factors influencing social app users' valuation of interdependent privacy [J]. Processing on privacy enhancing technologies, 2016(2): 61-81. [15] TSAI J Y, EGELMAN S, CRANOR L, et al. The effect of online privacy information on purchasing behavior: an experimental study [J]. Information systems research, 2011, 22(2): 254-268. [16] BERESFORD A R, KUEBLER D, PREIBUSCH S. Unwillingness to pay for privacy: a field experiment [J]. Economics letters, 2012, 117(1): 25-27. [17] EGELMAN S, FELT A P, WAGNER D. Choice architecture and smartphone privacy: there's a price for that [M]//The economics of information security and privacy. Heidelberg: Springer, 2013. [18] KRASNOVA H, ELING N, ABRAMOVA O, et al. Dangers of "Facebook login" for mobile apps: is there a price tag for social information? [C]//Thirty fifth international conference on information systems. Auckland: ICIS Press, 2014. [19] KIM J E, YEO J. Valuation of consumers' personal information: a south Korean example [J]. Journal of family and economic issues, 2010, 31(3): 297-306. [20] OTSUKI M, SONEHARA N. Estimating the value of personal information with SNS utility [C]//Proceedings of the eighth international conference on availability, reliability and security. Los Alamitos: IEEE Computer Society Press, 2013. [21] DANEZIS G, LEWIS S, ANDERSON R. How much is location privacy worth? [C]//Proceedings of the fourth workshop on the economics of information security. Cambridge: WEIS Press, 2005. [22] SPIEKERMANN S, KORUNOVSKA J, BAUER C. Psychology of ownership and asset defense: why people value their personal information beyond privacy [C]//Proceedings of the international conference on information systems. New York: ACM Press, 2012. [23] Baidu Chinese. Risk [EB/OL]. [2020-11-13]. https://hanyu.baidu.com/zici. [24] BINSWANGER H P. Attitudes toward risk: experimental measurement in rural India [J]. American journal of agricultural economics, 1980, 62(3): 395-407. [25] BECKER G M, DEGROOT M H, MARSCHAK J. Measuring utility by a single-response se-

quential method [J]. Behavioral science, 1964, 9(3): 226-232. [26] LAWRENCE N W. Social research methods: qualitative and quantitative approaches [M]. Translated by HAO Dahai. Beijing: China Renmin University Press, 2000. [27] KAHNEMAN D, TVERSKY K A. Prospect theory: an analysis of decision under risk [J]. Econometrica, 1979, 47(2): 263-291. [28] CHEN Rong, WANG Yifeng, QIU Zihua. Implied risk aversion: measurement, influencing factors, and information content [J]. Journal of Xiamen University, 2016(1): 116-127. [29] ACQUISTI A, TAYLOR R, WAGMAN L. The economics of privacy [J]. Journal of economic literature, 2016, 54(2): 442-492. [30] GRAEFF T R, HARMON S. Collecting and using personal data: consumers' awareness and concerns [J]. Journal of consumer marketing, 2002, 19(4): 302-318. [31] LEWIS K, KAUFMAN J, CHRISTAKIS N. The taste for privacy: an analysis of college student privacy settings in an online social network [J]. Journal of computer-mediated communication, 2008, 14(1): 79-100. [32] PREIBUSCH S. Guide to measuring privacy concern: review of survey and observational instruments [J]. International journal of human-computer studies, 2013, 71(12): 1133-1143.

Author Contributions: ZHANG Kailiang: data collection and analysis, chart drawing, drafting; ZANG Guoquan: research concept, study design, final revision.

Abstract: [Purpose/significance] Compared with the situation in which the probability of personal data leakage is 100 percent, the situation in which personal data is leaked with a certain probability is more common. Thus, this paper aims to measure users' personal data privacy value under the certain probability of privacy leakage, which puts forward a new perspective of privacy measurement and the measurement results are of practical significance to privacy classification protection. [Method/process] Based on the multiple price list, the user's return on financial risk is measured. Modified the implementation mechanism of multiple price list to elicit users' decisions between the risk-free scheme and the scheme with the probability of privacy leakage. Based on the above steps, value of privacy under leakage probability of users can be measured. [Result/conclusion] When the probability of privacy leakage is 30%, users' average VPLP in the social networks is about RMB 89.5; at the same time, when the probability of privacy leakage is 100%, users' average "willing to accept" and "willing to pay" of personal data in the social networks is about RMB 124.1 and RMB 93.8. Users' VPLP depends on the value of personal privacy itself and probability of privacy leakage.

Keywords: personal privacy; leakage probability; privacy value; multiple price list mechanism

Note: Figure translations are in progress. See original paper for figures.