Exploring the Privacy Black Hole

Junli Zhang [a], Yan Wang, Rui Wu

Xi'an Eurasia University, China

[a]156588351@qq.com

Abstract. With the popularization of electronic communication and social media, people-related information data has increased massively, which brings about a series of "privacy data risks" such as information data leakage. Conducting the personal behavior data as the index, the application domain as the criterion layer, and the personal privacy price as the goal, the personal privacy pricing model is established based on the analytic hierarchy process. Then we get the revised model of private pricing by combing the privacy valuation function and the market supply and demand relationship. Taking brand, quality, strategic factors into consideration, we establish the group privacy pricing system, national privacy pricing model.

Keywords: privacy data risks; privacy pricing model.; people-related.

1. Background Introduction and Question Restatement

With the popularization of electronic communication and social media, people-related information data has increased massively, while information data has been growing, and a series of "privacy data risks" such as private information sharing and information data leakage have occurred. Privacy data information is becoming more and more important, and its application will become a hot topic in the future.

You can edit or delete the content provided in this template and replace it with the text and figures of your article. Please be sure you do not accidentally leave any of this text as part of your paper!

1.1 The Current Situation of Privacy Leaks

Personal privacy will generate different values in different fields, for example, educational information is not valuable in medical field, but it is extremely important in financial field, this creates the situation that different fields with different privacy protection costs. According to the data and cases between 2004/7/1 and 2016/6/30, we summarize the size of lost due to the privacy leaks in different fields, as following figure shows:
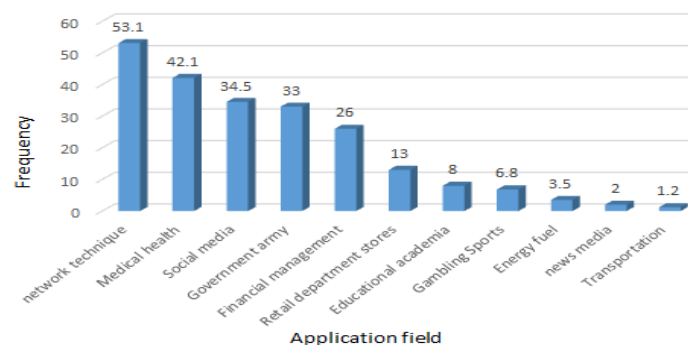

Figure 1. Scale comparison of privacy disclosure in different fields (unit: billion)

1.2 Privacy Cost Model

Due to the different uses of the basic data in different fields, the value of the private data increases, and the demand of privacy increases.
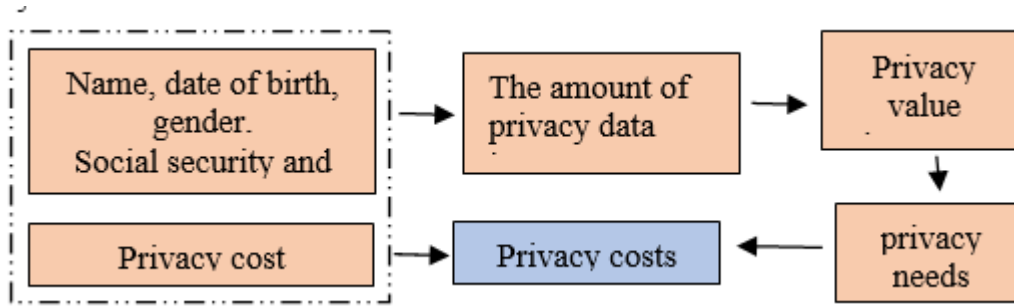
Figure 2. The relationship between basic information and cost

From the figure above, we can see that the addition of the basic information can increase the privacy costs.

Set the total privacy cost as W, and the privacy needed protection including personal basic information and relevant information in other fields. Thus, the total cost for privacy protection is:

$$W = \sum_{i=1} \alpha_i A_i + \sum_{i=1} \beta_i B_i + \cdots + \sum_{i=1} \kappa_i K_i + \sum_{i=1} \lambda_i I_i \tag{1}$$

A_i is the indicator of basic personal information, B_i is the indicator of network technique, C_i is the indicator of medical treatment and health...... K_i is the indicator of news media, I_i is the indicator of transportation. $\alpha, \beta, ..., \kappa$ and λ are the influence factors of $A_i, B_i, ..., K_i$ and L_i.

1.3 Medical Field

In the medical field, the privacy cost is influenced by basic personal information and medical health records, therefore, in this field, the cost of privacy is:

$$W_C = \sum_{i=1}^{12} \alpha_i A_i + \cdots + \sum_{i=1}^{9} \chi_i C_i \tag{2}$$

Among the medical field, there are 12 choices in basic personal information, and 9 choices in medical health records, including medical history information, height, weight, medical examination reports, blood pressure, sleep duration, walking distance per week, riding distance per week and fitness time. After analyzing the privacy data in medical field, we find the following results as figure shows:
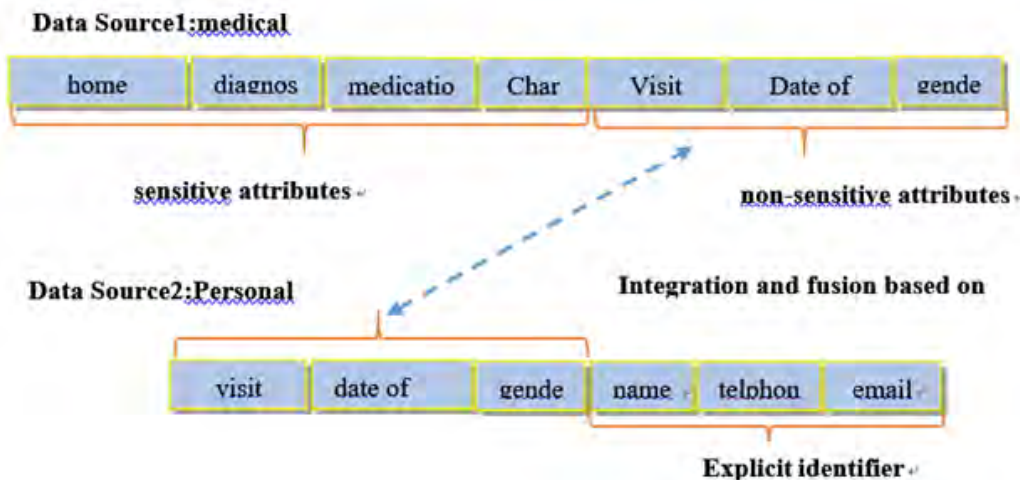

Figure 3. Application of medical data

1.4 Social Media Field

In the social medial field, the privacy cost is influenced by basic personal information and the relevant information on social media, therefore, in this field, the cost of privacy is:

$$W_F = \sum_{i=1}^{12} \alpha_i A_i + \sum_{i=1}^{5} \phi_i F_i \qquad (3)$$

A_i is the indicator of basic personal information, choosing 12 of all to analyze F_i is the indicator for the relevant personal information on social media, choosing 5 to analyze, such as call record, phone bill, chat software, contact way on software and online time.

1.5 Financial Trading Field

In the financial trading field, the privacy cost is influenced by basic personal information and the relevant information in financial trading, therefore, in this field, the cost of privacy is:

$$W_D = \sum_{i=1}^{12} \alpha_i A_i + \sum_{i=1}^{12} \delta_i D_i \qquad (4)$$

Among this field, financial trading field information has 12 types, including demand deposit information, fixed asset information, intangible asset information, the number of bank account, total value of asset, loan record, credit card record, bank card transaction record, the number of online shopping and the cost of online shopping.

1.6 The Composition of PI Prices

Due to the cost of privacy and PI as the part of privacy, when PI is treated as a commodity, the price of it is influenced by the following private information, shown as the figure:
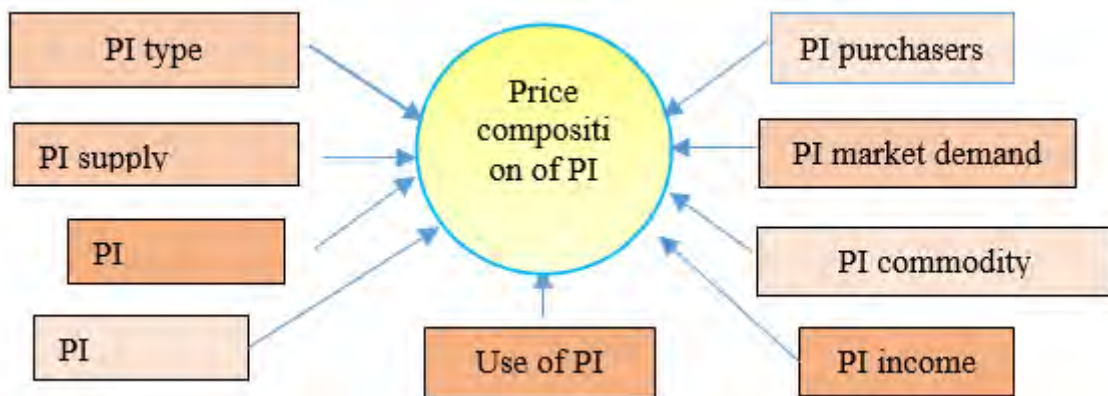

Figure 4. PI commodity price system information

As a commodity, it is impated by many factors and under different fields with different influences. Making a comparison between the value of a name and the value of a photo, the photo is more valuable for the reason that we can collect the information of gender through his or her outfit and characteristics, and infer the location by the background. However, it is much easier to find out someone by searching his or her name but not by photo.

2. The Pricing of Private Good

2.1 The Privacy Pricing Model Based on Analytic Hierarchy Process

Taking the supply and demand into consideration, we estabilish a pricing model which conducting an analytic hierarchy process and scored by experts, then we obtain a layered structure as follow:
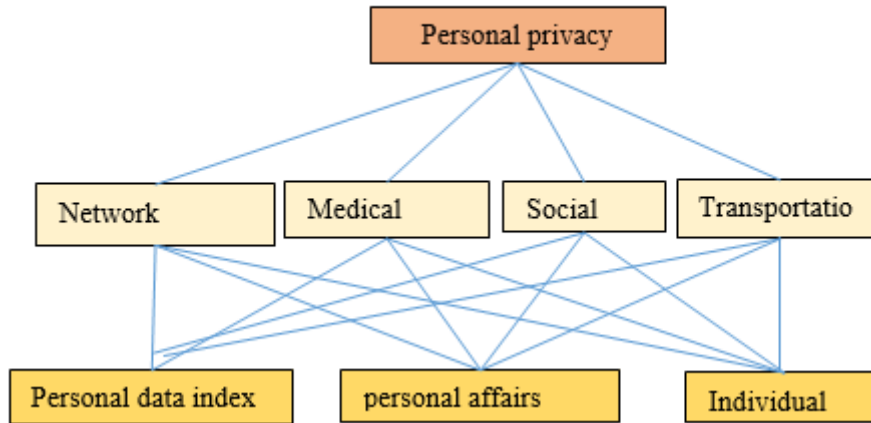

Figure 5. Hierarchical analysis of personal privacy pricing

Adopting analytic hierarchy process and making pairwise comparison matrix, which also through the consistency test, then we obtain the individual weight vector in every field, further we build the pricing model:

$$P = \sum K_{i,j}\left(\sum \alpha_i A_i + \cdots + \sum \gamma_i M_i\right) \tag{5}$$

A_i represents the indicator of basic personal information, M_i represents the indicator of relevant fields.

2.2 Revised Privacy Pricing Model

There are two prime factors which could impact the pricing of privacy: privacy parameter and how participant i pricing the value of privacy.

Assume that every participant i, the price of privacy $L_i(\varepsilon_i)$ is proportional to privacy parameter u_i, that is:

$$L_i(\varepsilon_i) = \varepsilon_i \times u_i \tag{6}$$

$u_i \geq 0$ is participant i's definition of privacy value. According to the private information, we can obtain the privacy price through our model, combining the personal willingness and market supply and demand, we build the model as following:

$$P = \sum K_{i,j}\left(\sum \alpha_i A_i + \sum \gamma_i M_i\right)\left(1 + \frac{1}{\varepsilon_i}\right)(1 + Q) \tag{7}$$

Assuming that other variables remain unchanged, when the personal willingness of selling his or her privacy increases, the price decreases; keep other variables remain unchanged, if the demand for personal privacy information increases, the cost of privacy increases.
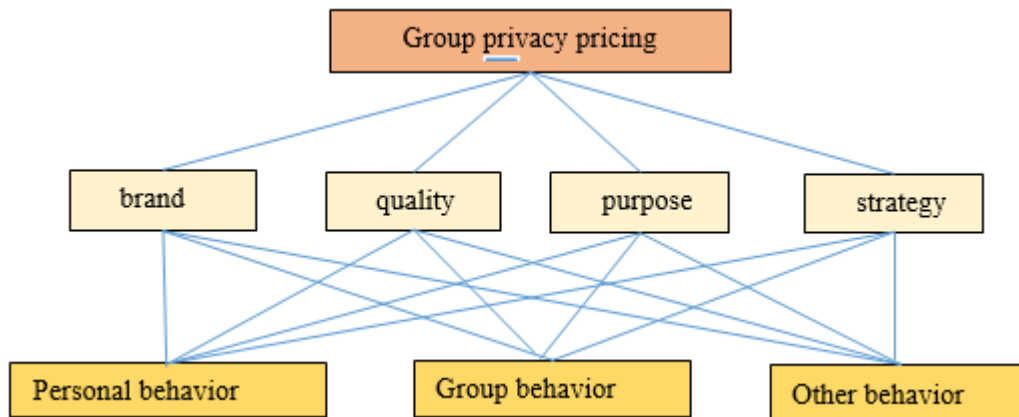
2.3 Group Privacy Pricing System

Figure 6. Hierarchical analysis of group privacy pricing

When someone determines to share his or her own privacy will influence others within a group, individual information can link other people in social, career, finance or other relationships. Group privacy pricing effect structure is shown as follow

According to the analytic hierarchy process, we can obtain the group PI price:

$$P_2 = \sum \eta_i \left(\sum \mu_i P + \sum \varsigma_i N_i \right) \qquad (8)$$

η_i is the indicator of the impact of other behaviors, P is the indicator of the individual, μ_i is the indicator of individual behaviors, N_i is the indicator of group behaviors, ς_i is the factor that impact group baheviors.

2.4 National Privacy Pricing Model

The pricing of national PI needs to take individual information pricing, group information pricing and other influenced factors into consideration, the model is:

$$P_3 = \sum \eta_i \left(\sum \mu_i P + \sum \varsigma_i N_i \right) + \sum K_{i,j} \left(\sum \alpha_i A_i + \sum \gamma_i M_i \right) \left(1 + \frac{1}{\varepsilon} \right) (1 + Q) + \xi \qquad (9)$$

When people have the right to sell the data, the result of the model will be changed. When the model is conducted by the outside factors, the cost of PI will be influenced, so that the price of PI will change as well, the result is that the market demand changes.

When the privacy leaks, it will bring risks into application field, and different effect, shown as follow: When people sell their own privacy, they can get return, maybe risk as well. Assume that users are rational, we can predict the return and risk through selling privacy. The key factor that influence his determine on selling is that whether the return is high enough to cover the risk, if the return is higher enough, then the user would to sell, conversely, the user would not to sell.

References

[1]. Wang Min. Research on the hierarchical protection of personal privacy in the big data age [D]. Wuhan University, 2016.

[2]. Wu Xiao Tong. Research on privacy protection and key technology in large data environment [D]. Nanjing University, 2017.

[3]. Zhang Xiaojian, Meng Xiaofeng. Data publishing and analysis oriented differential privacy protection [J]. Journal of computer, 2014,37 (04): 927-949.

[4]. Shen Qi. The trade-off between risk and cost: privacy paradox in social network -- Taking WeChat mobile social networking application (APP) in Shanghai as an example, [J]. journalism and communication research, 2017,24 (08): 55-69+127.

[5]. Dong Xin. Research on key technologies for data security and privacy protection in cloud computing [D]. Shanghai Jiao Tong University, 2015.

[6]. Li Yuan. Research on the protection of personal information in the big data age [D]. Southwest University of Political Science and Law, 2016.