Team Control Number

For office use only

T1_____

T2_____

T3_____

T4 _____

87280

Problem Chosen

F

For office use only

F1_____

F2_____

F3_____

F4 _____

2018 ICM

How much is your privacy information?

Summary Sheet

For ensuring the stability of the society, we establish a system of privacy pricing.

Task1, according to the personal characteristics and information field, we choose 26 indexes to measure privacy price, and divide them into 5 categories according to the degree of correlation.

Task2, we price the privacy. The present value method is used to quantify the value of privacy. We use the Analytic Hierarchy Process and Gradient Boosting Decision Tree, through the Python language to calculate the coefficients of the parameters. Then we establish Gauss mixture model to evaluate the information related value. Finally, the simulation of 500 people's privacy information is selected to verify the feasibility of the model.

Task3, we give the pricing method, introduce the demand elasticity coefficient, and analyze that the privacy information is more flexible. Then the life cycle changes in a long time range are predicted.

Task4, with the development of the times, the risk factor will be changed. We add dynamic factors and use Matlab to fit the change function of the risk coefficient.

Task5, the value of privacy will increase with age and then stabilize at a certain level. We find out the relationship between PI, IP and PP in the privacy market through a number of literature searches.

Task6, we use small world model and dynamic game theory to do simulation and prediction, and find that the privacy of individuals and groups will be largely leaked with the development of market.

Task7, the impact of data leakage is long tail. The agency can use the actuarial model to calculate the total amount of compensation to the individual.

Task8, we put forward some suggestions to the government on the basis of the model.

Finally, we perform a sensitivity test to verify the stability of the model.

I Introduction

With the development of social informatization, the commercialization of privacy becomes an important issue. Therefore, we need to establish a privacy pricing system to maintain the stable and healthy development of the information sales market. Our tasks are as follows:

Task 1: considering the level of risk can pass personal information and data fields classifying model, considering the different characteristics of category information, find the best balance of classification accuracy and simplicity.

Task 2: the question to think about the integrity of personal information, uniqueness, scarcity, according to the value of different factors, finally comprehensive assessment of the value of personal information.

Task 3: consider the information value, risk value, and the impact of information integrity on the final information pricing.

Task 4: consider the assumptions and constraints of the pricing model, judge the relationship between information privacy and human rights, and adjust the model to make it universal in the dynamic environment.

Task 5: consider the influence of age factors on the risk-benefit ratio of the population, and consider the similarities and differences between PI and PP and IP.

Task 6: make sure each individual data sharing information leakage caused by the network effect, and to consider whether the information network effect will affect the value of the personal information system, and to have the same privacy risks related to personnel, the sale of their personal information should be restricted.

Task 7: consider the impact of large-scale PI leakage on data vendors, acquirers, and information value systems. And consider whether the responsibility for data disclosure should be responsible for information disclosure.

Task 8: organize the above questions into proposals.

II Categorize individuals into sub-groups (Task 1)

2.1 A brief introduction to personal privacy pricing

Personal privacy is a basic resource, if it is used reasonably, it will help to promote the information social development. Culnan[1] find that, if private used by business are not known to individuals, they will increase their concerns about privacy disclosure. Hagel and others' [2] research shows, some people tend to pay more attention to privacy, the reason is that they want to get compensation for information leakage. Adjei[3]thinks that, people will be willing to provide personal information to get paid if their personal profits outweigh the risks.

Thus, it can be seen that, on the premise of the sound pricing system of personal privacy most people are willing to obtain certain compensation or compensation through the sale of personal privacy. So the appropriate pricing system should be set up.

2.1.1 **The establishment and evaluation of personal privacy information archives**

The classification of personal privacy information can help research and excavate the real value contained in information, and the value of information can be further evaluated by studying the correlation degree. Therefore, we chose 26 indicators based on the principle of personal information classification and the actual situation of this subject. It includes personal attributes, financial transactions, social networks, personal assets and health care. This paper classifies personal privacy into two aspects: individual characteristics and information field.

The indicators and their classification are shown in Figure 2-1.

Figure 2-1: Classification

The reasons for the selection of personal privacy information in five aspects are analyzed as follows:

- Citizenship:They belong to the range of attribute information and are objective information that can be used to identify specific individuals.
- Social Contact: Because of the unique social attributes of human beings, it has an important research value in the field of information, which is generated in the process of life with the other spider web information.
- Finance: It reflects the situation of individual financial trade and has great value for information mining.
- Health/Medical: It has certain significance to the study of social medical treat- ment service system, and reflects the various situations of the individual body.
- Personal Assets： It reflects the personal economic status, which is valuable for public interest, such as national economic analysis, and corporate profits, and it can create economic benefits.

III Privacy Cost Model（Task2）

3.1 Brief Introduction

The cost of privacy is a comprehensive pricing of the value of privacy, and we consider three components.

The first is the expected value of income, and we divide five two - level parameters into two categories.

- Difficult to quantify: As citizenship, we empower each parameter by analytic hierarchy process[4].
- Quantify: The other 16 parameters can be quantified by using the present value method of income in accounting. Then the coefficient of each parameter is calculated by the machine learning method of the Gradient Boosting Decision Tree(GBDT)[5].

The second aspect is the risk value of privacy. Because the value of risk is not easy to measure, we use probability theory to estimate the average deviation degree of expected revenue and get the expected risk reward[6].

The third aspect is information related value. The cost of privacy depends not only on the integrity of information, but also on the relevance. For example, the value of only the name should be lower compared with the value of the name attached to the person. We set up a Gauss mixed model to study the relationship between the correlation value and the relevant value of information[7].

3.2 Model Building

3.2.1 Expected Value of Income

The expected value of income is：

$$I = \sum_{j=1}^{26} \alpha_j C_i$$

Where：

I： Expected total income value

α_j: The value coefficient of the j^{th} parameter, $\alpha_j = \omega_i^*$

C_j： The expected return value of the j^{th} parameter

Step1:The empowerment of civil identity information by AHP

The related factors of citizenship quantification are decomposed into objectives, guidelines, programs and so on. Based on this, qualitative and quantitative analysis is carried out. The analytic hierarchy process is like Figure 3–1：

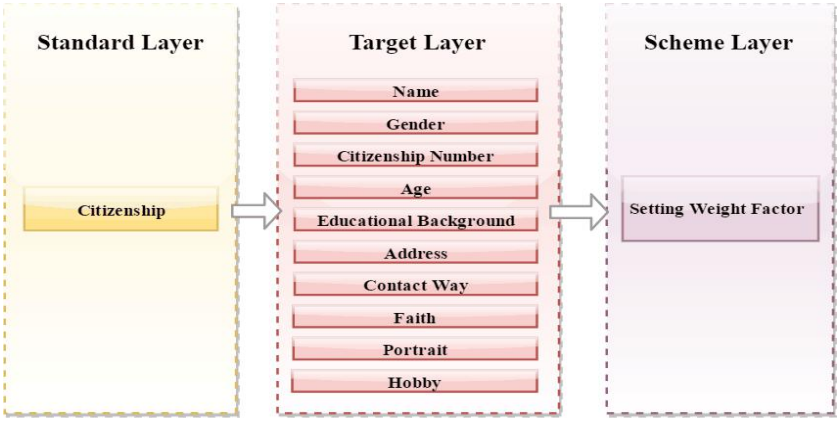


Figure 3–1: The Structure of AHP

- Construct a comparison matrix

By consulting relevant information and combining with the reality of life, we analyze the relationship and importance of 10 indicators in citizenship, and build pair wise comparison matrix.

- Hierarchical single order

The elements of each column of the judgement matrix are normalized, and the general terms of the elements are:

$$A_{ij}=A_{ij}\div\sum_{i,j=1}^{n}A_{ij}$$

Using sum and product method are used to calculate and add the judgement matrix of each column after the normalization. The W_i is obtained.

$$W_i=\sum_{i,j=1}^{n}B_{ij}$$

The element of the W_i is the sort weight value of the relative importance of the same level factor to a certain factor of the upper level factor. For $W=(W_1,W_2,...,W_n)^T$,In the process of normalization, the approximate solution of the eigenvector is obtained. Calculating the maximum eigenvalue of the judgment matrix λ_{max}

$$\lambda_{max}=\sum_{i=1}^{n}\frac{(AW)_i}{nW_i}$$

After that, we can make sure that the rank ordering is consistent with the consistency check. The so-called consistency check is the allowable range of A.

- Pairwise comparison matrix consistency test

The consistency index of the pairwise comparison matrix is Consistency Index (C.I.),

$$CI=\frac{\lambda_{max}-n}{n-1}$$

To measure the size of CI, a random consistency index RI is introduced. The method is: construct 500 pairwise comparison matrices randomly, and get the consistency index

$$RI=\frac{CI_1+CI_2+...+CI_{500}}{500}$$

The ratio of the consistency index CI. to the same order mean random consistency index R.I. is called the random consistency ratio Consistency Ratio(CR)。

$$CR=\frac{CI}{RI}$$

The CR<0.1 is obtained through the level analysis software, and the consistency check is passed.

- Determining the weight factor

  The eigenvectors are obtained by calculation,which is the weight of the 10 indexes.

  $\alpha_{j\in[1,10]}$=W=(0.056,0.019,0.155,0.025,0.075,0.134,0.053,0.095,0.186,0.20)

Step2：Earnings present value method quantifies other information

Privacy information seems to be difficult to quantify, but as long as the method is appropriate, we can quantify it to a certain extent. We regard privacy information as an intangible asset, a lot of international ways of quantifying intangible assets. Here we use the present value method[8].

The present value method of income is a kind of asset evaluation method to estimate the value of the assets evaluated by estimating the future expected revenue of the assessed assets and turning them into the present value. We determine the value of privacy information by calculating the future revenue of the privacy value. For example, Table 3-1 shows that the future earnings of businesses refer to the profits gained by merchants' knowledge of private information, or the expenditure reduced by government agencies and public organizations.

Table 3–1: Information value quantization table

Orginal Parameter		Quantized Parameters
Personal Assets	Total Assets	Average Deposit
	Income	Disposable Income
	Expenditure	Disposable Expenditure
	Intellectual Property	Education Expenditure
Social Contact	Social Way	Social Consumption Profit
	Friends	Social Web Site Profit
	Social Signal	Profit of Communication
Finance	Financial Credit	Average Credit
	Trading Information	Average Stock Investment
	Transaction on Amount	Average Transaction Amount
	Debt	Average Liabilities
Health/Medical	Physical Health Status	Health Expenses
	Medical Insurance	Insurance Cost
	Medical History	Medical Records
	Medical Expense	Medical Expenditure
	Genetic Map	Medical Research Funds

Step3：GBDT calculation of expected value coefficient

Gradient Boosting Decision Tree is a combination algorithm for machine learning[9],By iterating forward distribution algorithm, continue with the new weights, each round of iteration to get a strong learner and loss function, the next iteration of the goal is to find is a weak learning regression tree model, make the next round to minimize the loss, it is to find a decision tree, let sample loss to become smaller. This is the most significant algorithm in machine learning at present. Figure 3 - 2 shows the principle of its calculation[10].

The last formula of the above steps is the information gain after the division of the decision tree is created, and the best splitting point is the position with the highest information gain. In order to find the maximum gain position of the feature, we need to traverse all the characteristics, find the highest gain for every feature, calculate all the features gain, the specific process is like Figure 3 – 3, and then use greedy algorithm to repeat the process.

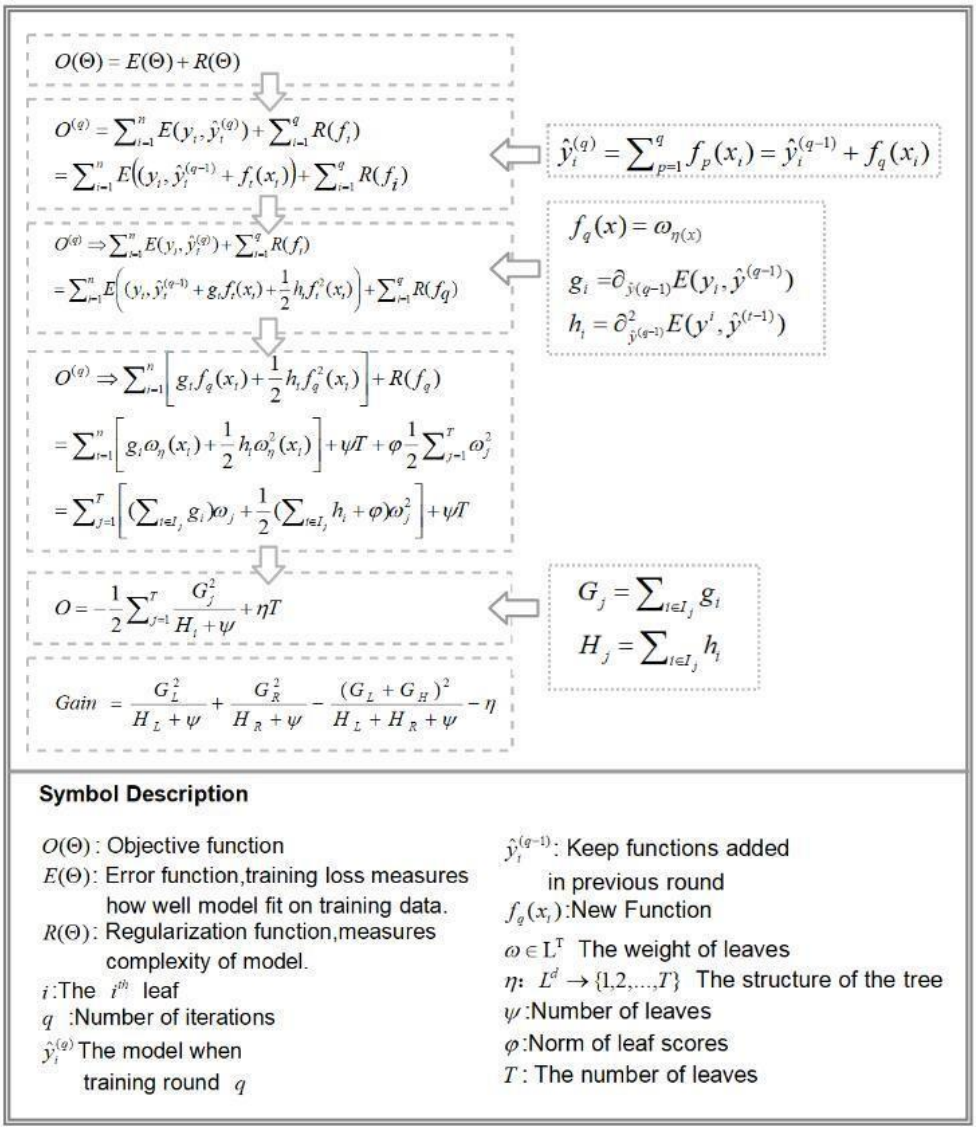


Figure 3–2: GBDT Principle

We use Python to program according to the above principles, then we can increase the accuracy and find the most appropriate coefficient by adjusting the steps, iterations, the maximum depth of decision tree, the minimum sample size and so on.Finally, we can get the expected value of each parameter value coefficient:

$$\omega^* = -\frac{G_j}{H_j + \psi}$$

3.2.4 Venture Value

Step1. The standard deviation is calculated.

Generally, we can use standard deviation to measure the degree of general response risk, but we have multidimensional index. In order to facilitate comparison with other parameters, it is more ideal to use standard deviation rate.

$$Cof = \frac{\sigma}{E}$$

Where

Cof： coefficient of variation, E： expected value, σ： Standard deviation

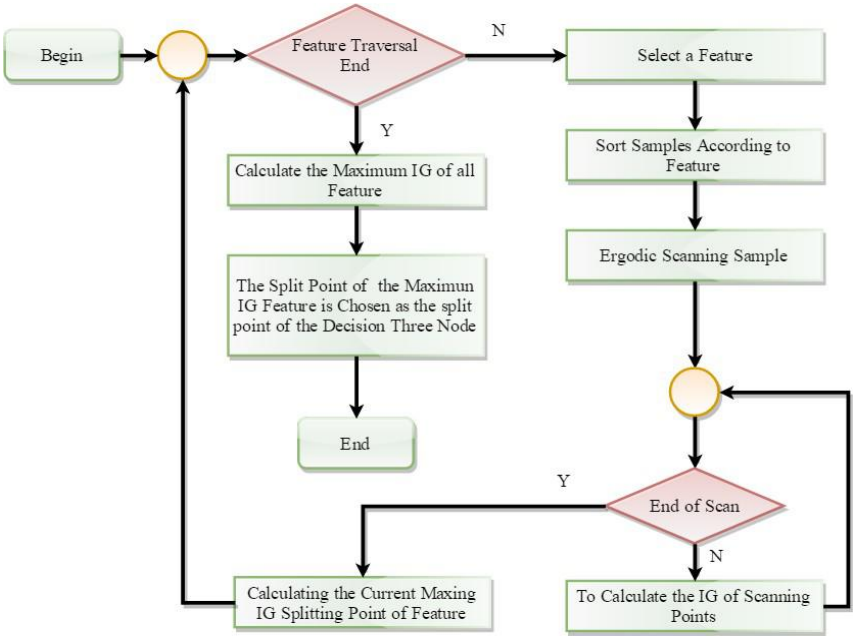


Figure 3–3:Information gain traversal

Step2. Determine the risk factor.

By the analytic hierarchy process described above, the weight of each parameter is considered as a risk factor, indicating the risk of privacy exposure.

The risk coefficient is determined by the analytic hierarchy process. After introducing the risk factor, we can calculate the expected rate of return on the risk of privacy:

$$R_i = \sum_{j=1}^{26} r_j Cof$$

Where

R_j： The rate of risk reward for the j^{th} parameter

r_j： The risk coefficient of the j^{th} parameter

Step3. Calculating the amount of the risk of privacy, that is the value of the risk.

$$K = \sum_{j=1}^{26} I_j R_j = \sum_{j=1}^{26} \frac{\alpha_j r_j C_j \sigma}{E}$$

Where

K： Value of privacy risk, I_j： Expected return value of the j^{th} parameter

3.2.4 Information related value

For the personal privacy information provided, the value of the data should be further excavated. According to practical experience, the more sensitive information is, the more sensitive rules can be mined, so the value of information is bigger, so we define "information related value" to measure the degree of information correlation.

According to the study, the relationship between the two is "S" positive correlation, so the Gauss mixed model is established[11], the Gauss probability density function is used to quantify accurately.

The relationship between information related value and information correlation is consistent with the following Gauss function curve trend such as figure **Figure 3–4.**

$$D(\delta) = ae^{-\frac{(\delta-\mu)^2}{2\sigma^2}}$$

The μ and σ are the parameters, and the maximum likelihood method can be used[12]to estimate the parameters

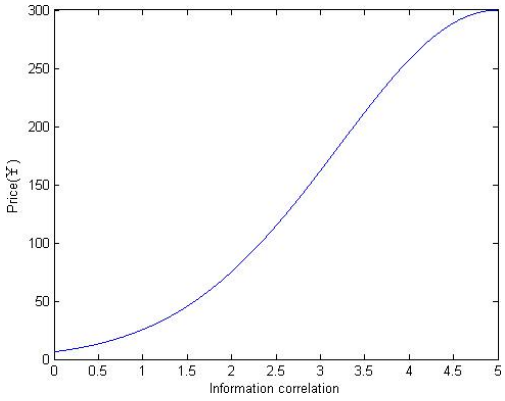


Figure 3–4:Relation diagram of information related value and correlation

The MATLAB program is used to program the correlation value of the information provided by the individual, and the information related value can be calculated. Personal data provided that the two person, the expected value and risk return value is the same, privacy cost pricing is not representative of the two men is the same, it is necessary to study the degree of correlation according to specific content provided by the two data, estimate the value of the relevant information, so as to calculate the final cost of privacy pricing.

The information related value is calculated on the basis of the expected value of income, and the concrete steps are as follows:

$$D = ae^{-\frac{(\delta-\mu)^2}{2\sigma^2}}$$

Where：

D： Information related value, I： Expected value of income, δ： Correlation

3.2.4 Total value of the cost of privacy

$$CP = I + K + D == \sum_{j=1}^{26} \alpha_j C_j \bullet \left(1 + \sum_{j=1}^{26} r_j Cof + ae^{-\frac{(\delta-\mu)^2}{2\sigma^2}} \right)$$

Where：

CP： Total price of privacy

I： Expected value of income

K： value of risk

D： Information related value

α_j： Expected return coefficient $\alpha_j = \omega_j^*$

Cof： coefficient of variation

R_i： Risk rate

r_i： risk coefficient

σ： standard deviation

E :expectation

δ： correlation

3.3 Model solution

3.3.1 Data compilation

Based on the parameters that are quantified, we have collected data for 1997-2016 years in China's census and consumption.

Data source:National Bureau of Statistics of China（http://data.stats.gov.cn/）

3.3.2 Data processing

After a simple arrangement of the data, further processing of the data is needed.

Data cleaning: check whether the data is correct. The outliers are modified by fuzzy matching to fill the missing values.

Data standardization: using the Z-score method to standardize the data, the research shows that the method has the highest compatibility in many criteria standardization[13].

3.3.3 Calculation coefficient

Through the above method, we get the risk coefficient, value coefficient and correlation degree of each dimension through the analytic hierarchy process and GBDT by Python programming. The specific values are Table 3–2.

3.3.4 Privacy cost pricing

The expected value of earnings is based on the data of China's nearly five years (2011-2016) and is further calculated by the above coefficients.

Table 3–2 Value coefficient and risk factor

Secondary	Tertiary	Income Coefficient	Risk Coefficient
Citizenship	**Name**	**0.056**	**0.020**
	Gender	0.019	0.020
	Citizenship Number	0.155	0.163
	Age	0.025	0.026
	Educational Background	0.075	0.059
	Address	0.134	0.173
	Contact Way	0.053	0.059
	Faith	0.095	0.083
	Portrait	0.186	0.197
	Hobby	0.200	0.200
Personal Assets	**Total Assets**	**0.045**	**0.581**
	Income	**0.014**	**0.163**
	Expenditure	**0.052**	**0.163**
	Intellectual Property	**0.084**	**0.092**
Social Contact	**Social Way**	**0.076**	**0.405**
	Friends	**0.008**	**0.149**
	Social Signal	**0.023**	**0.234**
Finance	**Financial Credit**	**0.025**	**0.211**
	Trading Information	**0.049**	**0.186**
	Transaction on Amount	**0.008**	**0.127**
	Debt	**0.007**	**0.687**
Health/Medical	**Physical Health Status**	**0.041**	**0.233**
	Medical Insurance	**0.058**	**0.060**
	Medical History	**0.017**	**0.091**
	Medical Expense	**0.053**	**0.233**
	Genetic Map	**0.074**	**0.383**

Table 3-3 Privacy cost price

	Citizenship										Personal				Social			Finance				Health/Medical					¥			
ID	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x19	x20	x21	x15	x16	x17	x18	x22	x23	x24	x25	x26	*I*	*K*	*D*	*CP*
1	0	1	1	1	1	0	0	1	1	1	1	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0	926.04	472.93	170.18	1569.16
2	1	1	1	1	1	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	365.19	734.33	173.54	1273.06
3	0	1	1	1	1	0	0	1	0	0	1	0	1	1	0	1	0	1	0	0	1	1	1	1	1	1	1506.07	528.41	198.76	2233.23
4	1	1	1	1	1	1	0	0	1	0	0	1	0	0	0	1	1	1	0	1	0	1	1	1	1	1	234.76	701.66	39.85	976.28
5	0	1	1	1	1	0	1	1	0	0	0	1	1	1	1	0	1	1	1	0	1	1	0	1	1	0	730.35	770.34	61.78	1562.46
.																														
.																														
.																														
224	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	1	1	0	0	1	0	1	611.83	55.10	88.19	755.12
225	1	1	1	1	1	1	0	1	0	0	0	1	1	1	0	1	1	0	0	1	0	1	1	1	0	0	598.51	403.65	80.72	1082.88
226	1	1	1	1	1	1	1	0	1	1	0	0	0	1	0	1	0	1	1	0	0	0	1	0	0	0	660.82	440.44	111.43	1212.69
227	0	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	0	1	1	0	650.62	692.13	79.51	1422.27
.																														
.																														
.																														
497	1	1	1	1	0	1	1	1	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0	1	0	815.32	419.52	138.94	1373.78
498	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	1	1	1	0	0	1	0	1488.56	343.07	231.13	2062.76
499	0	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	0	1	1	0	0	0	1	0	1	0	170.35	512.77	92.07	775.18
500	1	0	0	0	0	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	394.81	885.23	33.59	1313.63
																											Average	811.34	454.54	116.25	1381.42
																											Standard Deviati	371.82	235.88	43.94	449.61

1 of them represent this privacy information, and 0 are not provided.

Then, we simulated the transactions of 500 people's personal privacy data, such as Table 3 – 3, and calculated the privacy cost pricing based on these data, and made a brief analysis.

We calculate each individual's expected return value, risk value and relative value, have been found according to the total privacy, personal privacy information provided $600-3000 between the floating in.

If a person provides all his privacy information is all 1, then we calculate the total privacy for $5578 (per year).

IV Pricing system & Supply and demand relationship （Task3）

When we sell personal privacy information as commodities, we are bound to be affected by market fluctuations, including micro supply and demand, macro policy adjustment and economic balance.

4.1 Pricing system

Individuals can earn certain profits by selling their own privacy information. Businesses or government agencies want to purchase these information, and they will pay personal money, plus some profits besides privacy costs.

$$B = CP + S$$

Where:

B： Pricing of privacy information CP： The total price of privacy

S： Personal sale of hidden profits

According to the current market profits, in addition to luxury accessories and profiteering, the profit margin of a commodity is roughly 10%-30%. As a person's private seller, how much profit can be made for its own price, but the profit margin can not exceed 30%.

The profit margin is $\phi\left(\phi \le 0.3\right)$, we can get the following:

$$B = CP + CP \bullet \phi = CP(1+\phi)$$
$$= (1+\phi)\sum_{j=1}^{26} \alpha_j C_j \bullet \left(1 + \sum_{j=1}^{26} r_j Cof + ae^{-\frac{(\delta-\mu)^2}{2\sigma^2}}\right)$$

4.2 Demand elasticity of privacy value

When privacy information becomes a commodity in the market, it will receive the influence of market fluctuations caused by various factors.This kind of influence will cause the change of demand and price of privacy information. In order to respond to the degree and relationship of change, we introduce the concept of demand price elasticity in economics to explain the demand elasticity of privacy value[14].

$$E_d = \frac{\partial Q}{\partial B} \bullet \frac{B}{Q}$$

Where：

E_d： Elasticity coefficient of demand， Q： Quantity demanded，

B： Pricing of privacy information

In general, the elastic coefficient is negative, and in order to be simple, we see it as a positive number. It is also divided into five cases of Table4-1.

Table 4–1 Elasticity coefficient of demand

$E_d = 0$	Nonelastic,the price will change in any case, and the demand will not change.
$0 < E_d < 1$	Lack of elasticity，The rate of change in demand is less than the rate of price change.
$E_d = 1$	Single elasticity，at this time, the rate of price change is equal to the rate of change in demand.
$E_d > 1$	Resilient,A ratio of change in price to a ratio of change in demand.
$E_d = \infty$	Perfect elasticity，The price is fixed and there is an unlimited demand.

According to the concept of life cycle, we divide the market of this kind of information into three stages, as shown by Figure 4–1

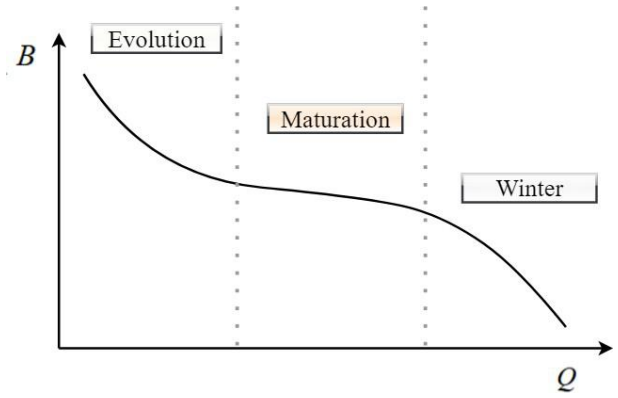


Figure 4–1 Life Cycle

- **Development period**: In this period, the relationship between the value of privacy and the amount of demand is more flexible. When the price of privacy rises, most people will be willing to sell, and when the price falls, people will choose to keep their privacy.
- **Mature period**: With the deepening of the concept of privacy information becoming a commodity, institutions or organizations have fixed demand for privacy information, and the coefficient of elasticity will be smaller, or even a fixed price.
- **Decline period**: When most of the people's privacy is sold for years, the market for privacy information is close to saturation, and there may be a price - free phenomenon.

V Assumptions and Dynamic improvement（Task 4）

5.1 Assumptions and constraints of personal information pricing model

1. It is assumed that people with the same conditions have the same willingness to protect privacy. In the real world, due to the particularity of each person, the degree of importance attached to personal information is different. Its distribution is roughly normal distribution. We assume that each person's attention to their personal information is only related to the individual characteristics and information related fields.

2. We don't consider the personal information of people in a special field. Special areas include political areas, military areas, and so on. In some special areas, the disclosure of

personal information will result in greater consequences and can't be quantified according to general standards.

3. We only consider the data of model system above. In order to quantify the value of personal data, we only select the characteristic factors that have a great impact on the value of personal information and evaluate the value of personal information according to the characteristic factors.

5.2 Whether information privacy should be regarded as human rights

In our view, information privacy should not be protected fully as human rights.

On the one hand, a part of personal information is closely related to the dignity of the person and reflects the personality factors of the individual. From the phenomenon of the commercialization of personal information, we can find that personal information reflects both personality and property interests[15].Therefore, it is reasonable to protect the property right.

To sum up, we think that the personal information should be divided into two aspects, that is, the personality elements of personal information and the property elements of personal information. In the collection and utilization of personal information, we should pay attention to the desensitization of personal information and make use of the information data without damaging the legitimate rights and interests of the information owners.

5.3 Dynamic analysis of cost of privacy

5.3.1 The change of personal values with time

With the development of time, the spread of privacy protection ideas is expanding. People are increasingly familiar with the consequences of privacy disclosure and pay more attention to privacy. Therefore, people's expectations for the return of privacy are getting higher and higher. According to the analysis of the privacy cost model, the risk value is changed due to the change of people's thought.

5.3.2 The dynamic analysis of the model

With the change of time， he value of risk increases gradually, we add dynamic factors on the basis of the original model. The risk factors of 26 indicators (r_j) are changed from constants to functions , which are changing by time(t).

Taking the total asset as an example, its trend of change is shown in Figure 5-1.

The data is fitted by the MATLAB fitting toolbox. We get a higher fitting degree R-square, which is 0.9779. The fitting effect is very good. So, the function relationship between the risk coefficient (r_{18}) and the time (t) of the 18th indexes is as follows:

$$r_{18} = -0.6404e^{(-0.07066t+0.6887)}$$

By fitting the risk coefficients of the other 25 indexes in the same way, we can calculate the function relation of the risk coefficients of each index. Combined with the formula of risk value:

$$K = \sum_{j=1}^{26} I_j R_j = \sum_{j=1}^{26} \frac{\alpha_j r_j C_j \sigma}{E}$$

We will add the functional relationship between the dynamic factors (r_j)and the time (t) into the upper formula, then it can calculate the change of risk value over time to realize the dynamic evaluation of the model.

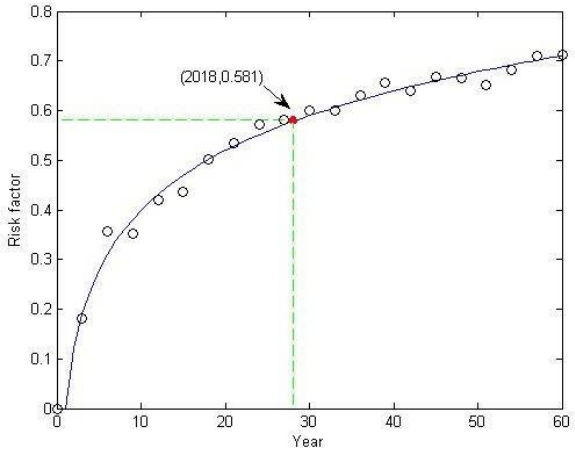


Figure 5-1: The change trend of total assets

VI Generational differences and conceptual comparisons（Task5）

6.1 Generational Differences

With the different ideas of the times and the growth of age, there are more or less differences in the cognition of the risk and income of privacy information. In this regard, we collect relevant data, and do an analysis.

Data sources:

(1) Data report on China's Internet consumption ecology in 2015- 2017.
http://doc.mbalib.com/view/e720ad8eb4ebac9f83d56ec790a6dd1c.html

(2) Census of China's national bureau of statistics.
http://www.stats.gov.cn/tjsj/pcsj/

(3) Chinese data Yearbook
http://data.stats.gov.cn/easyquery.htm?cn=C01

We get the consumption data of all ages in personal assets, social networking, financial transactions, health care, and then analyze and regress to get Figure 6 - 1, from 19 to 63 years old, the average growth of people's activities and consumption expenditure and 19 years old.

According to the analysis of the second problem, consumption expenditure can be regarded as expected value, so our model needs to make some changes.

If pertinence is stronger, or we want to accurately calculate the value of privacy information, we can increase the value of each kind of privacy value according to the growth of every dimension.

If only for the overall calculation, and without careful consideration of the increase in the situation, the average growth value can be used.

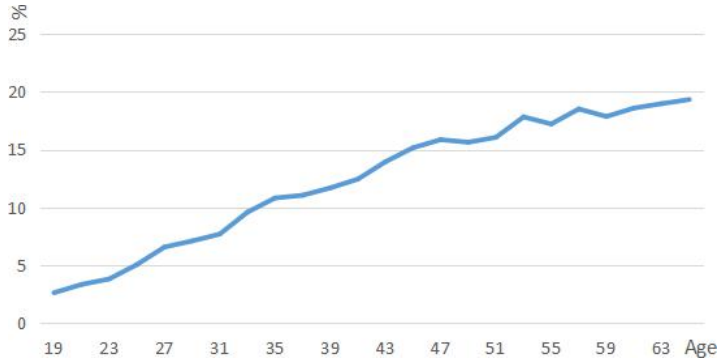


Figure6–1 Generational Differences of consumption level

6.2 Compare with PI，PP and IP

·Private information[16]： Non - public information ,Commercialization is not entirely beneficial to individuals, and the sale of personal information will bring a series of immeasurable risks, which may be invisible and long-term.

- PP：Personal property is divided into tangible and intangible assets [17].

The sale of PP by changing the commodity itself to a certain value is not determined by the relationship with the owner.

- IP：Rights generated by creative activities based on intelligence[18].

The sale of IP will not cause the risk of disclosure of privacy information.

The conceptual relationships of PI, PP, and IP are shown in Figure 6 - 2.

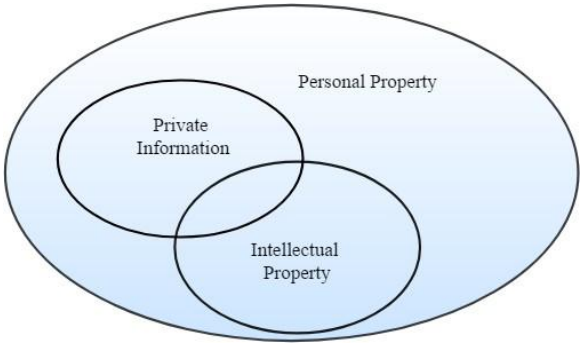


Figure 6–2 The conceptual relationships of PI, PP, and IP

VII Network effects of data sharing（Task 6）

We use Matlab to build a small world model for information leakage, and find that individuals who do not sell privacy are increasingly implicated with the development of privacy trading market. Then through the dynamic game theory, it is predicted that the information of the group will be leaked seriously because of the individual behavior, and if it is properly constrained, the interests of the group can be maximized.

7.1 The analysis and hypothesis of the problem.

1. Using the "economic man" hypothesis in economics, it is assumed that all the residents in the community are economic man, that is, each step should only consider their own interests.
2. Using a modified small-world model to simulate group information leaks, members who sell the information leak personal information about members of his or her neighborhood and have a p chance to reveal information about other members.
3. Assuming that any member's sale of personal information is immediately known to other members, this is a reasonable assumption in small groups.
4. Suppose each member has the same impact on information disclosure.
5. Suppose that the information breakers leak information to themselves and to the associated members have the same impact.
6. Information disclosure does not consider human rights constraints.

7.2 Model

7.2.1 A small world model of personal privacy disclosure.

Establish a small world model to describe the degree of privacy disclosure in a group, In the model, the proportion of people willing to sell personal information for profit is n , The proportion of people who are unwilling to sell personal information in return for benefits is $1-n$, individuals who sell their privacy will lose the privacy of two adjacent individuals in the model, the probability of having n causes loss of privacy of non-adjacent individuals. Use Matlab software to simulate.

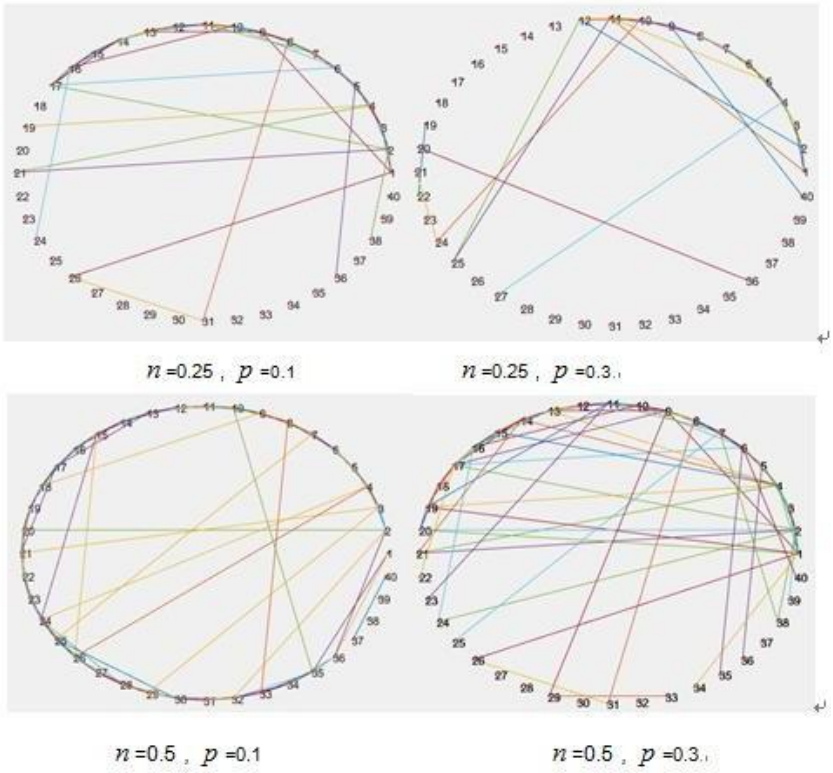


Figure 7– 1 :Simulation of small world model.

Assuming that the group has a population of 40, adjust the range of *n* and *p* ,and simulate the loss of privacy.

The individuals who are connected by the attachment are individuals who are at risk of privacy disclosure. According to the Matlab simulation analysis, it is obvious that with the increasing of the parameters P and N, the privacy disclosure in the group becomes more and more serious.

7.2.2 Analysis of prisoner's dilemma based on dynamic game theory.

In a group, some people are willing to sell their own privacy, others think that privacy is more important, we assume that the sale of privacy is an open, in groups, there are four kinds of condition.

Table 7-1 The game of interest within the group

		Members of B	
Game matrix		Sale privacy	No sale
Members of A	Sale privacy	Both A and B get money and lose their privacy.	B get money, **Both A and B lose their privacy.**
	No sale	A get money, **Both A and B lose their privacy.**	NO one get money and lose their privacy.

According to "economic man" hypothesis, each group member only consider their interests and action, is willing to sell a member of the privacy will sell their privacy, under this precondition, even if other members do not want to sell privacy, their privacy will be violated, as a result, even if part of the members are reluctant to sell privacy for interests, in seeking to stop the cases, will sell the privacy in order to get more benefits.

Based on the above simulation analysis, we get the conclusion. In the outside world without intervention, a group with the same privacy risks into sub-game refining Nash equilibrium, cannot get benefit optimization, in order to jump out of the prisoner's dilemma, to maximize the interests of groups and groups to limit the privacy of personal selling behavior is necessary.

VIII Impact of data disclosure on the value of privacy(Task 7)

When information security issues arise in the organizations that purchase personal privacy information, there will be a large amount of privacy information leakage. If the privacy is stolen by criminals, it can pose a threat to personal safety and social stability.

8.1 The impact of data leakage on the value of privacy

We discuss the impact of data leakage based on the components of privacy value.

1. When a large number of data is leaked, personal privacy information is improperly distributed in the market, then the bussiniess' willingness to purchase personal information will be reduced, and the expected return value of the individual will also decline.

2. Because of people's awareness of the risk of personal information leakage is increased, and their willingness to sell information will decline.

3. The relevant value will decrease with the decrease of expected return value.

In general, as Figure 8-1 shows, data leakage can cause the decline of personal privacy value, and the willingness of people to sell their privacy value will also decline, which will lead to the economic crisis of privacy trading market.

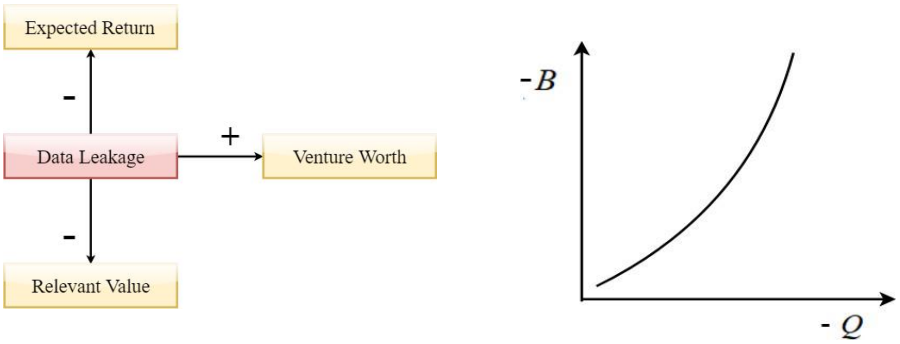


Figure 8–1：The impact of data leakage on the value of privacy

8.2 Time effect of data leakage

As time goes on, the impact of data leaks can change, too. We consider the influence of two trends, one is the long tail effect [19], the other is the bullwhip effect[20]. As shown in Figure 8-2.

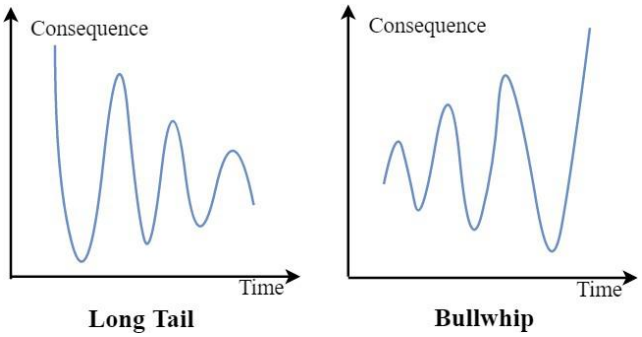


Figure 8–2: Long Tail and Bullwhip Effect

- Long tail effect： If the public opinion of the leak is guided by the demand side of personal information, such as intermediary companies and government agencies, then there will be a long tail effect.
- Bullwhip effect： If the message of the leaked data is spread rapidly across the network, the overall public opinion is not controlled by any organization, then there will be the bullwhip effect.

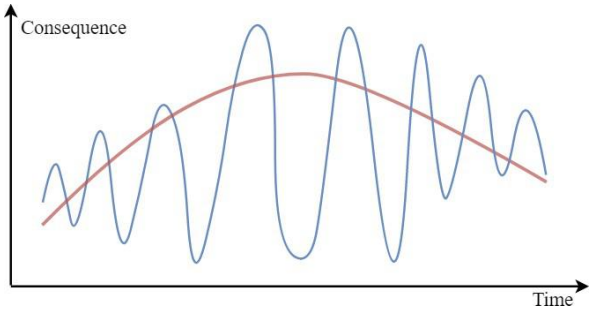


Figure 8–3 : The impact of data leakage changes over time

Although these two cases will appear, the bullwhip effect can be controlled by the other.With impact events increasing, effects of data leakage has become smaller. Therefore, the overall impact of data leakage, such as Figure 8-3. It will rise first, then drop, and there will be fluctuations in it.

8.3 Agency compensation model

When the data leakage occurs, the agency needs to compensate everyone for a certain loss. In addition to the value of the data itself, the agency also needs to compensate for other losses caused by data lost. We use the actuarial model of the insurance[21] industry to calculate the amount of compensation. The establishment method of actuarial model is generally parametric modeling, and its process is shown in Figure 8-4.

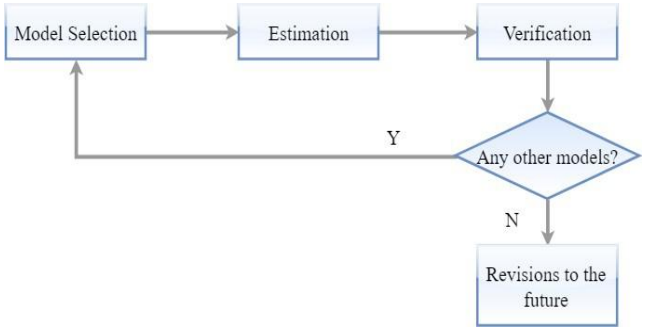


Figure 8–4 : Flow chart of parameter construction of actuarial model

Finally, the amount of compensation we have calculated is as follows:

$$U = I + \Pi$$

Where：

U :The amount of the final compensation to the individual

I :The value of the information itself ,and it is also the original expected value.

Π ： Other losses should be compensated for by the actuarial model.

IX Sensitivity analysis

In order to test the stability of the model, we consider that different people pay different attention to personal privacy information and have different attitudes and tolerance to risk disclosure. Therefore, the weight assignment of value coefficient and risk coefficient is different from person to person. In order to test the universality of the coefficients determined by the gradient lifting tree and the analytic hierarchy process, we have designed a questionnaire on personal privacy information attitude (see Appendix II). In this paper, 50 people of different ages were selected as the respondents, we investigated their weighting of 26 indicators and obtained new value coefficients and risk factors. Through the analysis of the data, 3 groups of invalid data were excluded, and 47 new value coefficients and risk factors were obtained. We put these coefficients into the model and repriced the privacy costs of the 500 people generated by the simulation. The new pricing situation is shown in Figure 9-1.(Because of the large amount of data, only partial data is displayed.)

¥	1	2	3	. . .	28	29	30	. . .	45	46	47	Average	Original
1	2418.38	514.48	138.68		1678.21	631.23	1553.20		2167.91	2290.65	2212.43	1529.31	1569.16
2	2170.00	710.77	370.17		333.31	1985.96	2475.36		586.87	1245.99	761.85	1244.81	1273.06
3	976.14	1262.71	868.07	. . .	1723.50	267.38	1148.80	. . .	708.05	1527.78	269.61	1002.89	2233.23
4	2440.34	3156.78	2170.18		4308.75	668.45	2872.00		1770.11	3819.44	674.03	2507.21	976.28
5	213.61	426.25	1668.34		2125.89	1829.91	1596.25		2064.98	1033.28	1474.04	1408.55	1562.46
.		.				.				.		.	.
.		.		. . .		.		. . .		.		.	.
.		.				.				.		.	.
224	759.22	982.11	675.17		1340.50	207.96	893.51		550.70	1188.27	209.70	780.02	755.12
225	1243.98	1360.05	1409.85	. . .	1515.01	1026.92	473.99	. . .	1032.11	762.64	1404.80	1167.50	1082.88
226	324.77	242.11	487.81		2271.10	1868.35	1827.77		2399.22	2410.90	263.25	1398.47	1212.69
227	444.85	1115.50	2152.79		173.90	1906.60	1597.18		2504.45	2508.15	1912.18	1409.52	1422.27
.		.				.				.		.	.
.		.		. . .		.		. . .		.		.	.
.		.				.				.		.	.
497	1129.58	1170.88	1486.88		542.34	1063.58	1956.28		733.50	1933.00	1167.24	1444.82	1373.78
498	2345.73	2334.04	820.51	. . .	1408.93	1449.89	1795.29	. . .	3119.91	1882.15	2049.26	1794.48	2062.76
499	884.87	589.07	780.10		1128.67	1023.80	301.64		1496.41	534.34	470.12	818.34	775.18
500	1474.78	981.78	1300.16		1881.12	1706.33	502.73		2494.02	890.56	783.53	1363.90	1313.63

Figure 9-1: Local diagram of new pricing results

The average value of the new pricing (Average) is not significantly different from the original model (Original). We use SPSS software to carry out the independent-samples T-test on these two sets of data, and get the test results such as Figure 9- 2.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
data	Equal variances assumed	.083	.775	-.115	24	.910	-19.79385	172.83605	-376.50991	336.92222
	Equal variances not assumed			-.115	23.998	.910	-19.79385	172.83605	-376.51157	336.92388

Figure 9-2: The result of independent-samples T-test

It can be seen from the test results that, in the test of homogeneity of variance, **p=0.775>0.05**, we can accept the original hypothesis, it can be considered that the variance is equal, the independent-samples T-test should be used.

The results of independent-samples T-test shows, **p=0.910<0.05,** we can't accept the original hypothesis. Therefore, there is no significant difference between the above two sets of data, so the stability of the model is good.

X Analysis of the advantages and weaknesses of the model

- Advantages：

1. We consider many basic and derived parameters and establish a personal information pricing model that reflects the impact of all aspects.
2. The impact of the emergency pricing model is considered.
3. Model is a simple but flexible and reliable. When a certain factor changes greatly, the model can be universally adjusted by simple adjustment.
4. The method used in the pricing model is novel, R-squared is high, and the error rate is small.

- Weaknesses:

1. Some parameters need to be adjusted according to the actual situation.
2. It is impossible for all people can accept the commercialization of privacy.
3. The areas involved are not exhaustive.

Privacy Pricing Policy Recommendations

Dear decision maker:

In the information society, the analysis and processing of information has brought great social and economic value. Information becomes more and more important.

At present, the personal information of citizens has been circulated in the market, but personal information is an important part of information cluster. The laws, policies and pricing strategies related to the commercialization of personal information still exist in a large gap in related fields, resulting in huge risks. In order to standardize the market order of information, reduce the risk of citizen information leakage brought by information commercialization. Through the comprehensive consideration of various factors, we have set up a pricing system for the commercialization of personal information.

Our team considers the value of information from both the information owner and the information user. We obtain cost measurement through information, assess the risk and value of information sharing, and predict the value of information. Furthermore, we consider the influence of dynamic factors and sudden events on information value and establish the information commercialization pricing system.

Privacy information pricing

We consider the following factors to price the model.

prospective earnings： Consider privacy as an intangible asset, quantifying the value of personal information through the quantitative method of intangible assets. It can also be understood that when individuals sell their privacy as a commodity, they can make a big profit for their privacy buyers.

Risk assessment value： When people sell their privacy, the privacy they sell brings certain risks. As compensation, the information buyer should compensate the seller for risk. Because each person's information and information value is different, the risk value is different.

Relevance value： In fact, the privacy value of the system is often higher than that of individual privacy prices. Other things being equal,if privacy is a system, the value of privacy is higher.

Other factors： In addition to the above three factors, we also consider other factors to dynamically compensate the model.

- **Generational differences：** Age differences lead to different ideas. This leads to greater cognitive differences in the value of privacy.
- **Age difference ：** The younger age differences can also be reflected in the concept, and the information value of the individual will change as the age changes. Individuals' perceptions of risk and benefits also change with age.
- **Unpredictable risk：** Some risks are latent and cannot be predicted in today's

circumstances.

Effect of pricing

Applicable fields： We make pricing proposals for PI in three areas

- social media
- financial transactions
- health/medical records

Model effect：

According to the online survey and simulation, the same crowd divided on the value of information cognition has a little different, but the overall value of the float is less than $10, the pricing model of basic can meet the expected value of people.

And more than 50% of people are willing to sell their privacy information at this price, but only if the information can be reasonably used, and it will not pose a threat to personal safety.

National macro-control

In order to avoid unexpected privacy information market conditions, lead to data leakage or other social instability happens, the government must want to undertake certain macroeconomic regulation and control, limit privacy price within a certain range.

- **Promulgates the laws and regulations of economic regulation and regulation**: establish a perfect system to adapt to the development of market economy, and calm the relationship between supply and demand.
- **Disciplinary punishment for illegal behaviors**: avoid using other people's privacy information for illegal activities.
- **Restricted trading of special information**: information of certain special groups, such as government personnel, should be restricted according to circumstances.

Accident prevention

If something unexpected happens, for example, a large amount of privacy is leaking, we need to develop a complete solution.

- **Agency compensation**: on the basis of personal information value, make certain increment compensation.
- **Guide public opinion**: the government and media guide public opinion without infringing on individual rights and interests, and prevent riots or panic, causing social instability.
- **Control the market**: try to avoid leaking data into the market to prevent economic recession or imbalance.
- **Deal with it as soon as possible**: timely and properly handle the unexpected situation.

Sincerely

Team 87280

References:

[1] Culnan M J.'How did they get my name':an exploratory investigation of consumer attitudes toward secondary information use[J].MIS Quarterly,1993,17(3):341-363.
[2] Hagel J III,Rayport J F. The coming battle for customer information[J]. Harvard Business Review,1997,75(1):53-65
[3] Adjei J K. Monetization of personal identity information: technological and regulatory framework[C].India: International Conference on Information Science&Security,2015
[4] Al-Harbi A S. Application of the AHP in project management[J]. International Journal of Project Management, 2001, 19(1):19-27.
[5] Dai yixin, sun rongling. Realization and quantification of intangible asset value [J]. Chinese soft science,2000(07):74-77.
[6] Zhu guo. Application of EXCEL in financial management -- calculation of investment risk value [J]. Journal of chifeng college (natural edition), 2009, 25(4):130-131.
[7] Sun Guangling, Tang Xianglong. Hierarchical semi supervised learning algorithm based on Gauss mixture model [J]. Computer research and development,2004(01):159-164.
[8] Li Hongxun. Asset evaluation and management: China Forestry Publishing House, 2000
[9] Friedman J H. Stochastic Gradient Boosting[J]. Computational Statistics & Data Analysis, 2002, 38(4):367-378.
[10] Introduction to Boosted Trees, Tianqi Chen, 2014
From http://homes.cs.washington.edu/~tqchen/pdf/BoostedTree.pdf]
[11] Sun Guangling, Tang Xianglong. Research and development of semi supervised learning algorithm [J]. layered computer based on Gauss mixture model,2004(01):159-164
[12] Zhang Rongquan, Du Yuming, Yang Jianyu. A LFM signal maximum likelihood estimation model and a fast algorithm for parameter estimation [J]. Journal of radio wave science,2005 Journal of radio wave science (05):101-105.
[13] Xu Yunhui, Li Zhongfei. Dynamic portfolio selection based on income sequence related dynamic mean variance model [J]. Theory and practice of system engineering, 2008(08): 125-
[14]Diego S. Price elasticity of demand[J]. Betascript Publishing, 2009, 3(4):1717-1718.
[15] PengYun. Research on property property of personal information from the angle of Anglo American Property Law. Legal system and society:ten-day periodical, 2011(11): 249
[16] A. Beimel and Y. Stahl, Robust information-theoretic private information retrieval, in Proceedings of the 3rd International Conference on Security in Communication Networks (SCN'02), pp. 326–341, 2003. Cite is from DGH 2012, op. cit.
[17] Personal property". Sir Robert Harry Inglis Palgrave. Dictionary of political economy, Volume 3. 1908. p. 96
[18] Personal property". Sir Robert Harry Inglis Palgrave. Dictionary of political economy, Volume 3. 1908. p. 96
[19] Tang Haijun. Theory of long tail theory of economics[J]. Modern management science. 2009(1):62-64.
[20] Tang Haijun. Theory of long tail theory of economics[J]. Modern management science. 2009(1):62-64.
[21] Xiao Yan. Actuarial model [M]. Renmin University of China press, 2013.

Appendix I

Python

```
#####################################GBDT#################################
import numpy as np
from sklearn.ensemble import GradientBoostingRegressor
gbdt=GradientBoostingRegressor(loss='ls', learning_rate=0.1,
n_estimators=100,
subsample=1, min samples split=2, min samples leaf=1, max depth=3
, init=None, random_state=None, max_features=None, alpha=0.9, verbose=0,
max_leaf_nodes=None, warm_start=False
)
train_feat=np.genfromtxt("f_train.txt",dtype=np.float32)
train_id=np.genfromtxt("f.txt",dtype=np.float32)
test_feat=np.genfromtxt("f_test.txt",dtype=np.float32)
test_id=np.genfromtxt("ff.txt",dtype=np.float32)

gbdt.fit(train_feat,train_id)
pred=gbdt.predict(test_feat)
print(gbdt.feature_importances_)
total_err=0

for i in range(test_feat.shape[0]):
    print(pred[i],test_id[i])
    err=(pred[i]-test_id[i])/test_id[i]
    total_err+=err*err

print(total_err/pred.shape[0])
##Result##
'''
====================== RESTART: F:\math\gbdt\usegbdt.py
======================
 [0.05769512  0.00610115 0.05629678 0.01203783 0.02682956 0.08157958
  0.03342411  0.09468546 0.03692312 0.05884569 0.03874673 0.06441484
  0.05666908  0.01886848 0.04130826 0.08557419]
 2.0430143820015707  2.04364
 1.7260317333863573  1.72623
 1.4539094183773056  1.45404
1.187036709759271 1.18706
0.936417771272397 0.93639
2.324334324614137e-08
'''
```

Matlab

```
%%%%%%%%%%%%%%%%%%%%%%%%%Small world model%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function matrix = SW()
tic
a=1:1:40;
c=randperm(numel(a));
N=a(c(1:20));m=2;
p=0.1;
matrix=sparse([],[],[],40,40,0);
for i=m+1:N(1:20)-m
for j=i-m:i+m
matrix(i,j)=1;
end
end
for i=1:m
for j=1:i+m
matrix(i,j)=1;
end
end
for i=N(1:20)-m+1:N(1:20)
for j=i-m:N(1:10)
matrix(i,j)=1;
end
end
for i=1:m
for j=N(1:20)-m+i:N(1:20)
matrix(i,j)=1;matrix(j,i)=1;
end
end
for i=1:N(1:20)-m-1
for j=i+1:i+m
r=rand(1);
if r<=p
unconect=find(matrix(i,:)==0);
M=length(unconect);
r1=ceil(M*rand(1));
matrix(i,unconect(r1))=1;
matrix(unconect(r1),i)=1;
end
end
end
for i=N(1:10)-m+1:N(1:10)-1
```

```
for j=[i+1:N(1:20) 1:i- N(1:20)+m]
r=rand(1);
if r<=p
unconect=find(matrix(i,:)==0);
r1=ceil(length(unconect)*rand(1));
matrix(i,unconect(r1))=1;
matrix(unconect(r1),i)=1;
end
end
end
for i=N(1:20)
for j=1:m
r=rand(1);
if r<=p
unconect=find(matrix(i,:)==0);
r1=ceil(length(unconect)*rand(1));
matrix(i,unconect(r1))=1;
matrix(unconect(r1),i)=1;
matrix(i,j)=0;matrix(j,i)=0;
end
end
end
for m=1:N(1:20)
matrix(m,m)=0;
end

toc
end
 function tu_plot(rel,control)
r_size=size(rel);

if nargin<2
    control=0;
end
if r_size(1)~=r_size(2)
    disp('Wrong Input! The input must be a square matrix!');
    return;
end
len=r_size(1);
rho=50;
r=2/1.05^len;
theta=0:(2*pi/len):2*pi*(1-1/len);
[pointx,pointy]=pol2cart(theta',rho);
theta=0:pi/36:2*pi;
```

```
[tempx,tempy]=pol2cart(theta',r);
point=[pointx,pointy];
hold on
for i=1:len
temp=[tempx,tempy]+[point(i,1)*ones(length(tempx),1),point(i,2)*ones(lengt
h(tempx),1)];
    plot(temp(:,1),temp(:,2),'r');
    text(point(i,1)-0.3,point(i,2),num2str(i));
end
for i=1:len
    for j=1:len
        if rel(i,j)
           link_plot(point(i,:),point(j,:),r,control);
 end
    end
end
set(gca,'XLim',[-rho-r,rho+r],'YLim',[-rho-r,rho+r]);
axis off
function link_plot(point1,point2,r,control)
temp=point2-point1;
if (~temp(1))&&(~temp(2))
    return;
end
theta=cart2pol(temp(1),temp(2));
[point1_x,point1_y]=pol2cart(theta,r);
point_1=[point1_x,point1_y]+point1;
[point2_x,point2_y]=pol2cart(theta+(2*(theta<pi)-1)*pi,r);
point_2=[point2_x,point2_y]+point2;
if control
    arrow(point_1,point_2);
else
    plot([point_1(1),point_2(1)],[point_1(2),point_2(2)]);
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Risk factor %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
x=0:60;
y=0.71*log(x)/log(60)
figure('color',[1 1 1]);
plot(x,y)
hold on;
```

```
plot([28,28],[0,0.581],'g--')
hold on;
plot([0,28],[0.581,0.581],'g--')
xlabel('Year')
ylabel('Risk factor')
hold on;
gl=[0:3:60]
Y=interp1(x,y,gl);
dian=[0
0.182,0.356,0.351,0.420,0.437,0.501,0.535,0.571,0.581,0.599,0.601,0.630,0.
656,0.640,0.668,0.665,0.651,0.683,0.710,0.712];
 scatter(gl,dian,'k')


%%%%%%%%%%%%%%%%%%%%%%%%%%%% The correlation %%%%%%%%%%%%%%%%%%%%%%%%%%%%

clc,clear
a=xlsread('data.xlsx','Sheet1','A4:Z504'); %The data is in Appendix II
name=xlsread('data.xlsx','Sheet1','A4:A504');
CitizenshipNumber=xlsread('data.xlsx','Sheet1','C4:C504');
Address=xlsread('data.xlsx','Sheet1','F4:F504');
ContactWay=xlsread('data.xlsx','Sheet1','G4:G504');
if(name==0&Citizenship==0&Number==0&Address==0&ContactWay==0)
    Y=0;
else
    gl=xlsread('data.xlsx','Sheet1','AA4:AA504');
x=0:0.00001:5;
y=300*gaussmf(x,[1.8 5]);
figure('color',[1 1 1]);
plot(x,y)
Y=interp1(x,y,gl)
end
xlabel('Information correlation')
ylabel('Price(£¤)')
```

Appendix II

An Investigation Into Attitudes To Personal Privacy.

1. Your gender：

A. Male

B. Female

2. Your age：

A．≤11years

B.12 years—18 years

C.19 years—35 years

D. 36 years—65 years

E.≥65 years

3. Your career：

A. Government officials

B Technical personnel

C. Junior officers

D. Business services

E. Business and service personnel.

F. Production, transportation equipment operators and related personnel.

G.solder

4. You attach great importance to personal privacy information.：

A. attach great importance to

B. attach importance to

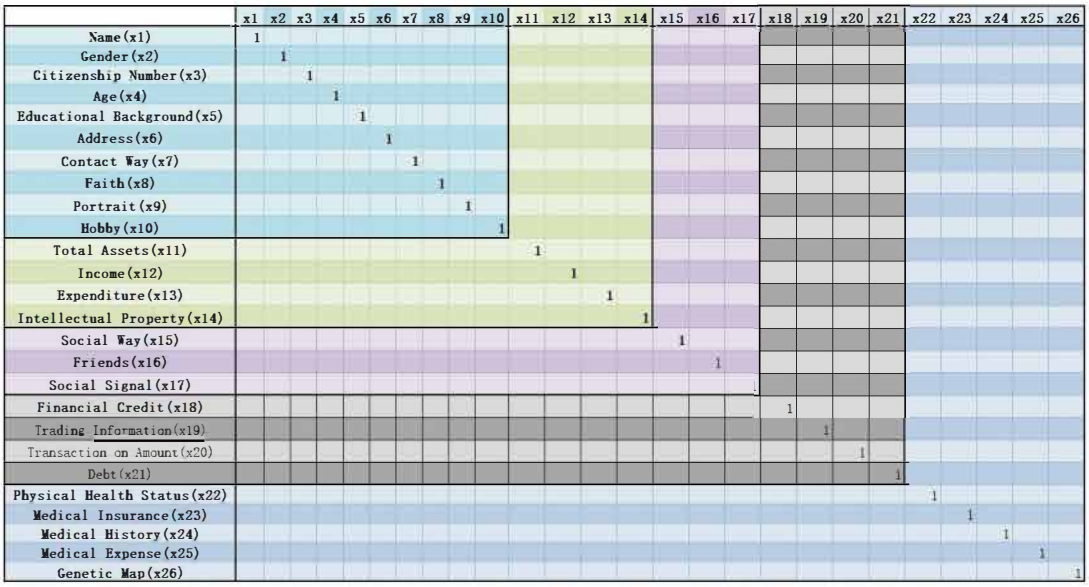
C. general emphasis

D. do not take the

E. it doesn't matter

5. Please fill in the following table with the following information you think:

For example, if you think the home address information is three times as important as the name information, please fill in the "3" in the corresponding form; If you think the importance of gender information is 1/8 times that of interest and hobbies, please fill in "1/8" in the corresponding form.


x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26
Name(x1)
Gender(x2)
Citizenship Number(x3)
Age(x4)
Educational Background(x5)
Address(x6)
Contact Way(x7)
Faith(x8)
Portrait(x9)
Hobby(x10)
Total Assets(x11)
Income(x12)
Expenditure(x13)
Intellectual Property(x14)
Social Way(x15)
Friends(x16)
Social Signal(x17)
Financial Credit(x18)
Trading Information(x19)
Transaction on Amount(x20)
Debt(x21)
Physical Health Status(x22)
Medical Insurance(x23)
Medical History(x24)
Medical Expense(x25)
Genetic Map(x26)