Advances in Intelligent Systems Research, volume 147

International Conference on Network, Communication, Computer Engineering (NCCE 2018)

Big Data Privacy Protection Model Based on Multi-level Trusted System

Nan Zhang [a)], Zehua Liu [b)] and Hongfeng Han [c)]

International school, Beijing University of Posts and Telecommunications, Beijing 100876, China.

[a)] 2015212939@bupt.edu.cn
[b)] 2015213474@bupt.edu.cn
[c)] Corresponding author: hanhongfeng@bupt.edu.cn

Abstract. This paper introduces and inherit the multi-level trusted system model that solves the Trojan virus by encrypting the privacy of user data and achieve the principle: "not to read the high priority hierarchy, not to write the hierarchy with low priority". Thus, ensuring that the low-priority data privacy leak does not affect the disclosure of high-priority data privacy. This paper inherits the multi-level trustworthy system model of Trojan horse and divides seven different risk levels. The priority level 1 ~ 7 represent the low to high value of user data privacy and realize seven kinds of encryption with different execution efficiency Algorithm, the higher the priority, the greater the value of user data privacy, at the expense of efficiency under the premise of choosing a more encrypted encryption algorithm to ensure data security. For enterprises, the price point is determined by the unit equipment users to decide the length of time. The higher the risk sub-group algorithm, the longer the encryption time. The model assumes that users prefer the lower priority encryption algorithm to ensure efficiency. This paper proposes a privacy cost model for each of the seven risk subgroups. Among them, the higher the privacy cost, the higher the priority of the risk sub-group, the higher the price the user needs to pay to ensure the privacy of the data. Furthermore, by introducing the existing pricing model of economics and the human traffic model proposed by this paper and fluctuating with the market demand, this paper improves the price of unit products when the market demand is low. On the other hand, when the market demand increases, the profit of the enterprise will be guaranteed under the guidance of the government by reducing the price per unit of product. Then, this paper introduces the dynamic factors of consumers' mood and age to optimize. At the same time, seven algorithms are selected from symmetric and asymmetric encryption algorithms to define the enterprise costs at different levels. Therefore, the proposed model solves the continuous influence caused by cascading events and ensures that the disclosure of low-level data privacy of users does not affect the high-level data privacy, thus greatly improving the safety of the private information of user.

Key words: Multi-level Trusted System Model; Risk Sub-group and Risk Level; Price Point Model; Privacy Value; Privacy Cost Model; Pricing Model; Symmetric and Asymmetric Encryption Algorithms.

INTRODUCTION

As people's living standards improve, people rely more and more on electronic communications and social media. However, some people seem don't like to share too much private information. They feel that sharing too much can easily lead to a loss. "Privacy monetization" has become a magic weapon in today's data economy. Information can be sold at a price when that is legally acquired and the transaction is warranted at the same time. This also shows that it is very important to establish and improve the law of data management. Massive data gathering and mining, its commercial value is self-evident, however, the balance between commercial value and personal privacy protection, which is the problem everyone will face.

MODEL ASSUMPTION

Assume that most users have low levels of privacy encryption requirements, the price is affected by market supply and demand, government policymakers want companies to reduce their cost per unit of ownership over time, government policymakers want companies to raise their cost per unit when users are small., and Assume the computer executes the encryption algorithm, the time can be quantified.

PRICE POINT MODEL

Seven-Level Model - Different Areas

This paper divided the different data into seven areas [1~4]. The proportion of different personal information data in the existing database is shown in table 1.

TABLE 1. Seven-level model

Level 7	Financial information
Level 6	Consumer information
Level 5	Family information
Level 4	Whereabouts information
Level 3	Social Relations
Level 2	Personal information
Level 1	Contact information

Contact information includes address, phone number, email, etc., Personal information includes name, gender, portrait, ID card, etc., Social Relations includes Curriculum Vitae, Social Background, etc., Whereabouts information includes pages visited, countries visited, etc., Family information includes whether married, family members, etc., Consumer information includes purchases, the place to buy, and more, and Financial information includes the account, the password, and so on.

According to the survey and analysis, we ranked the above seven areas according to the degree of privacy, and sorted the results as shown in Table 1. In other words, the layer 7 of financial information is generally considered to be safe and confidential.

Seven-Level Model - Different Risks

Business management will encounter many risks, Such as: strategic risk, investment risk, investment decision risk, investment implementation risk, investment exit risk, policy risk, international operation risk, overseas investment risk, international engineering contract risk, market risk distribution, market development risk, securities market risk, Customer Risk, Customer Credit Risk, Customer Relationship Maintenance Risk, Customer Business Model Risk, Brand and Reputation Risk, Brand Strategy Risk, Branding and Maintenance Risk, Reputation Risk ... Some of these risks can be combined into different levels.

We construct a privacy cost model. The higher the user selects the risk sub-group, the higher the value of the private information. For the enterprise, the higher the cost of privacy, so we lead to the privacy cost model. We consider dividing the two levels of the Level 7 Trusted System Model into three areas.

TABLE 2. The two levels of the level 7 trusted system model

	social media	financial transactions	health/medical records
Payment Information	Ali payment	Bank card	Medicare card
Family members information	Facebook	Visa card	Apple watch`s pedometer data

Trade-offs between data protection: There is a negative correlation between data protection level and execution efficiency, so tradeoff exists. The higher the level of data protection, the higher the data privacy value of the user,

which requires a higher encryption algorithm, thereby reducing the execution efficiency. According to the model assumption, most users tend to lower-level encryption algorithms for higher efficiency.

Level 7 Trusted Model has a certain priority. The higher the value of the user's private information, the higher its risk sub-group level, and the higher its priority. For example, there is a star in order to prevent strangers from disturbing. He set his phone number to the highest level of sensitivity and the highest level of encryption. In the star's opinion, his phone number information is much more valuable than his name information, and the value of a name alone is lower than value of a name with the person's picture attached.

The Privacy Cost Model constructs structure of the cost per layer. Cost increases with the level of increase, the cost difference between layers also gradually increased. The first layer of the image is as follows, and the images at other levels are similar.

TABLE 3. Design a pricing structure for PI

Level	Layer1	Layer2	Layer3	Layer4	Layer5	Layer6	Layer7
Cost	100	125	175	275	400	550	800

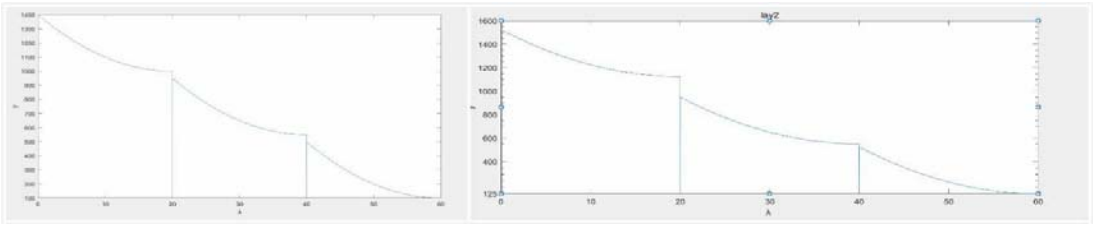

FIGURE 1. Layer1 with cost 100 and layer2 with cost 125

Pricing Model

(1) The market regulation of the profit maximization quantitative analysis

For any single product, the cost C consists of the fixed cost FC and the variable cost VC, that is,

C=FC+VC

Among them, a certain constant FC. VC = vex, v is the average unit cost of change, and x is the product's sales volume. According to Hypothesis 3, v is known.

As for the income R, is divided into two kinds of consideration of tax rates and do not consider the tax rate.

If you do not consider the tax rate: R = PX

If you consider the tax rate: R = PX (1-tr)

Where P is the price, x is the sales volume, try is the tax rate and is a known quantity.

The profit π is equal to the income or minus the cost C. Therefore,

H = PX-vex-FC or π = PX (1-tr) -vex-FC

The total profit equals the sum of the profits of each product, which is expressed as

$$\Pi = \sum_{i=1}^{n} \Pi_i = \sum_{i=1}^{n} (P_i X_i - v_i X_i - FC_i)$$

Or

$$\Pi = \sum_{i=1}^{n} \Pi_i = \sum_{i=1}^{n} (P_i X_i (1 - t_{ri}) - v_i X_i - FC_i)$$

Among them, Pi said the price of the itch product. From a mathematical point of view, the problem of maximizing the total profit of a bank is the planning problem

$$\max_{P_1 P_2 \ldots P_n} \Pi \quad \text{or} \quad \max_{X_1 X_2 \ldots X_n} \Pi \tag{1}$$

The meaning of G1 is that determining the price (or sales) of each product makes the total profit the most. According to Hypotheses 1 and Hypotheses 2, the expression of total profit π is a multivariate function of price P (or sales volume x).

(2) The market under the regulation of the total sales volume to maximize quantitative analysis Total sales equal to the sum of each product sales weighted,

$$X = \sum_{i=1}^{n} k_i X_i$$

Where Xi represents the sales volume of a single product. Ki said the itch product sales weight. According to hypothesis 4, Ki is a known parameter.

$$\max_{P_1 P_2 \ldots P_n} \sum_{i=1}^{n} k_i X_i \tag{2}$$

The meaning of G2 is that determining the price of each product maximizes the total weighted sales. Also, based on the assumptions 1 and 2, the expression of total sales is also a multivariate function of price P.

(3) Government guidance under the price of profit maximization quantitative analysis

Guidance price, that is, the price of these products to be limited, cannot be priced in accordance with G1. In this case, the bank requires each product to meet its own expected profit margin. On this basis, to maximize profits. At this point, the above model will become a constrained programming problem.

Constraints can be expressed mathematically. The specific form is as follows:

$$P_i X_i - V_i X_i - FC_i \geqslant \Pi_i \quad \text{or} \quad P_i X_i (1 - t_r) - V_i X_i - FC_i \geqslant \Pi_i$$

Ki is the expected profit margin of the itch product, that is, the profit cannot be lower than the known πif. If must be non-negative. The left formula does not consider the tax rate, the formula on the right to consider the tax rate.

Banks operate in accordance with the guidance prices set by the supervisory authority, so to speak, there is a minimum requirement that the product cannot lose money. In this case, it is sufficient to take πif = O in the constraint conditions written above.

Put the constraints into the planning problem and get it

$$\max_{\substack{\{P_1 P_2 \ldots P_n\} \\ st. P_i X_i - V_i X_i - FC_i \geq \Pi_i \\ i=1,2,\ldots,m}} \Pi \quad \text{or} \quad \max_{\substack{\{P_1 P_2 \ldots P_n\} \\ st. P_i X_i (1-t_r) - V_i X_i - FC_i \geq \Pi_i \\ i=1,2,\ldots,m}} \Pi \tag{3}$$

Among them, St. Represents constraints.

G3 shows that in the case that m products (n = 1, 2, m) of n products are guided prices and m price-limited m products achieve the expected profits, the other n-m the price of the product maximizes the total profit of the bank

(4) Government guidance under the total sales volume to maximize the quantitative analysis

For a product that directs the price, the bank simply asks for it to meet the expected profit margin and, under such constraints, formulates the prices for each of the other products to maximize the total sales. The pricing model is expressed mathematically as a planning problem below,

$$\max_{\substack{\{P_1 P_2 \ldots P_n\} \\ st. P_i X_i - V_i X_i - FC_i \geq \Pi_i \\ i=1,2,\ldots,m}} \sum_{i=1}^{n} k_i X_i \tag{4}$$

The letters and symbols in G4 have the same meaning as G3. [1]

Custom formula: y=(x-x0)2+y0 (0≤x≤x0)
(x-x1)2+y1 (x0<x<x1)
(x-x2)2+y2(x1<x<x2)

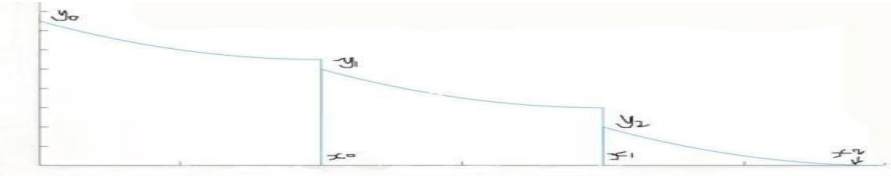

FIGURE 2. The figure of custom formula

Optimization of the Model

Following is the relation among PI, PP and IP:

Different points:

Private information refers to the proprietary market knowledge owned by individual market participants, where experience is the most invaluable individual knowledge of market participants.

Generally believed that the individual knowledge includes three categories:

Knowledge of the individual's own characteristics. Such as personal physical condition and working ability;

Knowledge of personal behavior. Such as effort, enthusiasm and so on;

Knowledge of individual's understanding and understanding of the state of the environment. Mainly refers to the individual's grasp of market information and awareness;

Personal property refers to personal belongings. Including citizens 'legitimate income, housing, savings, supplies, artifacts, books and materials, trees, livestock and law that allow all citizens' means of production and other legal property.

Intellectual property is the power generated by intellectual-based creative activities. it includes:

Invention patents, trademarks and industrial designs and other aspects of industrial property.

Natural sciences, social sciences and works of literature, music, drama, painting, sculpture, photography and cinematography.

The similar points:

The legitimacy: personal information, personal property, intellectual property is subject to legal protection.

Personal: they are related to individuals.

The secrecy: unknown or do not want people to know.

Subjectivity: Confirming whether a particular private matter is private depends mainly on the subjective feelings of the oblige. Even if the oblige has known some information of the oblige, but the oblige believes that the information remains confidential, it must Respect the wishes of the rights holders.

Typical generational differences such as age differences and income differences.

Perceptions of PI and data privacy at different ages are different. For example, when a government official disagrees with the government, he uses some means to keep his social information confidential, but young college students do not care whether their political opinions or social messages are being circulated or not. Relatively young people, who take much less responsibility, naturally they wouldn't 't too much care about the privacy of how much impact on their own. On the contrary, older people, who have more social responsibility, are naturally more concerned about the proportion of private risk-benefit impact on themselves.

There are also income differences. More and more data collection and analysis may increase class differentiation in perception of PI and data privacy risk-reward ratio. Rich people want to protect their privacy, and they are also in the best position to take advantage of privacy protections and privacy management services that cater to their needs. From the actual situation, the rich can more easily accept the monthly payment of a privacy protection fee. Judge Kosciuszko may be willing and able to pay 200 dollars a month to protect his privacy, but ordinary consumers may find the sale uneconomical. The negative impact felt by the lower population on big data is probably the greatest. The poor have never had any privacy, after all, "castle" and "high wall" are for the rich. However, even in today's era, the poor are among the first to lose basic privacy protection.

As generations age, the model will change. The growth of age can affect the change of the model. For example, people of different ages have different opinions when facing the same student grade database. With a score of 70 out of 100, a 50-year-old looking at the grade is not the same as a high school student who is about to face a college entrance examination at 18, looking at the grade. Older people look at this result without any pressure, because he is too old to face the pressure of exams. This high school student will immediately face the risk of failing to pass the

grade, so he has some sense of oppression. So, in the model design process, we should take into account changes in age.

In this paper, Assumptions and constraints should address issues such as government regulations (e.g. price regulations, specific data protections such as certain records that may not be subject to the economic system) and cultural and political issues.

The pricing model introduces the assumptions and constraints is the number of people come in a unit of time, equivalent to the rate, that is, the value of the flow (the value λ) and a maximum flow (the value a) That is the maximum saturation.

At the same time, we have met data protection regulations such as prices and specific data protection. One price and other data protection regulations, we use the relationship between supply and demand, we do so:

In order to cater to the needs of national policies, when the demand exceeds the supply, we allow the enterprises to reduce the selling price of a single item, but the selling price is still above the cost to ensure that the enterprise will not Loss of operation. However, as the flow of people increases, we also guarantee that enterprises cannot operate profits. However, there is an upper limit on the growth of people's flow. It is the value of a, which is the maximum number of customers a company can tolerate in a unit time. When the number of people is too small (λ is small), that is, the demand is less than the supply, and we allow the firm to raise the selling price of a single item to maintain its basic operating profit.

We also meet specific data protection regulations. Customers can change the level of encryption of their private information according to their individual needs. A star sets his phone number to the highest level of sensitivity and the highest level of encryption.

In the era of big data, we should pay more attention to the protection of personal privacy. Internet companies cannot forget the corresponding social responsibilities and legal obligations. They should know that ignoring laws and regulations and lack of social responsibility will eventually lead to cocooning. Today, privacy has become a commodity that can be bought and sold. Therefore, privacy should be regarded as constitutional rights and even basic human rights.

Consider introducing a dynamic element to your model by introducing the variations over time in human decision-making given changing personal beliefs about the worth of their own data. We have introduced a dynamic element into the model. This dynamic element is mood. We set the mood element (the value m), the mood will change trading data over time, such as online shopping and history search. We divide the mood m into three values. The three values represent:

TABLE 3. The assignment of the dynamic elements

Description	Assignment
When the mood is good.	m=+1
when the mood is not bad	m=0
when the mood is sad	m=-1

The value m is an independent variable over time, the proportion of transaction data (the value T) is the dependent variable.

TABLE 4. The distribution of the transaction data

T=50%	(m=+1)	T=35%	(m=0)	T=15%	(m=-1)

Encryption Algorithm

This paper provides customers with 7 different algorithms to protect their privacy information, and customers can choose different algorithms [5] based on the level of security they need for privacy information. And the higher the security of the algorithm, customers may have to pay more expensive fees.

TABLE 5. The different encryption algorithms

Level1	Caesar password
Level2	"Virginia" password
Level3	DES Algorithm
Level4	AES Algorithm
Level5	Hash function
Level6	RSA Algorithm

CONCLUSION

This paper provides customers with seven security levels of encryption, allowing customers to choose different levels of privacy based on different encryption methods. Certainly, the more secure the encryption method is, customers will pay more expensive fees. And then this paper sets up three models, they are pricing model, price point model and privacy cost model. Pricing model draws the effect of per-hour flow of people on privacy costs by considering the impact of human traffic per unit of time on market volatility. In addition, this paper also considered the influence of mood and age on the model and optimized the model. When setting up the model, this paper also considers some of the constraints and limitations on the model, including national legislation. In this seven-layer encryption algorithm, this paper takes symmetric encryption and asymmetric encryption respectively to avoid some bad conditions, including the user's privacy information being stolen. This paper hopes that this model can serve the people well and make the privacy information of people be safeguarded as the premise, to maximize people's privacy information to serve the society and benefit the society. When private information can be traded, we want this model to be of value to the government.

REFERENCES

1. WuJianwei. Design and Implementation of Commercial Bank Product Pricing Model[C]. Zhejiang University.
2. Ellison N B. Social network sites: Definition, history, and scholarship [J]. Journal of computer-mediated Communication, 2007, 13(1): 210-230.
3. Kwan H, Lee C, Park H, et al. What is Twitter, a social network or a news media? [C]//Proceedings of the 19th international conference on World Wide Web. ACM, 2010: 591-600.
4. Wasserman S, Faust K. Social network analysis: Methods and applications [M]. Cambridge university press, 1994.
5. David G I, Wells D L, Kim J B. A database encryption system with sub keys [J]. ACM Transactions on Database Systems (TODS), 1981, 6(2): 312-328.