

# 基于 MT-TF 模型的个人隐私定价问题研究

王瑞桃<sup>1a</sup>, 余林栋<sup>1b,2</sup>, 于恒悦<sup>1b</sup>, 刘广臣<sup>1a</sup>, 宋美<sup>1a</sup>

(1.鲁东大学 a.数学与统计科学学院; b.信息与电气工程学院, 山东 烟台 264039;  
2.东北石油大学 电气工程信息学院, 黑龙江 大庆 163318)

**摘要:** 大数据时代带给人们方便的同时,也因为信息泄露问题,引起了人们对于自身隐私数据信息的重视。本文基于市场时效性和三角模糊数提出运用隐私信息定价模型 MT-TF 量化个人隐私信息。MT-TF 模型从时间成本、财产损失和精神损失三个方面确定隐私信息初始价值,应用信息新度函数确定隐私信息的时效性价值,利用基于均衡理论的市场方法准确量化个人隐私的价值。为获得部分信息价格,该模型利用三角模糊数来获得相应的信息权重。模型实证分析结果显示,某真实个人隐私泄露案件中法院判决为13 684元,模型计算为15 802元,总体效果较好。该模型能够较好地完成个人隐私信息定价的精准计算,为我国司法、安全及合法隐私交易等领域提供有力的技术支持。

**关键词:** 隐私信息; MT-TF; 市场法; 信息时效性; 三角模糊数

**中图分类号:** O212 **文献标志码:** A **文章编号:** 1673-8020(2021) 02-0109-07

信息技术的快速发展会时刻产生大量信息。根据已公开信息,2001 年至今,已有 11.27 亿用户的隐私信息被泄露。信息时代下,如何保护信息及在信息泄露后如何挽回损失已成为重要的研究课题。如何确定个人隐私信息的价格一时间成为关注热点,目前尚未在全球领域内对信息数据定价方法形成统一认识。因此,实现隐私定价理论中定级机制合理化、应用广泛化是各国研究人员以及大数据从业者的重大难题。

一方面,对于个人而言,文献[1]认为每个人的隐私成本是一样的,只是风险不同,并通过数学规划求得长期个人信息价格。由于个人信息的不同方面敏感程度不同,易泄露程度也不同,文献[2]针对个人容易泄露的敏感信息,建立隐私价格系统模型,同时用层次分析法建立个人隐私价格定价模型;文献[3]指出一些机构和公司根据客户隐私保护需求,制定相应的保护和定价方案。

另一方面,不同类型的信息价值不同,文献[3—5]使用层次分析法和模糊综合评价法分析

隐私信息内容,通过分类最终得到隐私定价。文献[6—7]从隐私数据的细节保护角度出发,给出了各种基于  $k$ -级匿名和划分隐私保护方法;文献[8]提出一种基于帕累托最优思想的服务定价模型,综合考虑买家与供应商的利益,使用多目标粒子群优化算法得到双方满意的定价结果;文献[9]运用隐私预算对需要保护的隐私类型划级,结合 VCG 机制,为不同级别的隐私保护制定不同的价格;文献[10—11]分别根据粒子群算法、信息熵对隐私信息分类确定价格;文献[12—13]研究在社交网络下隐私信息保护成本与隐私信息泄露后的损失之间的博弈。

由于当前人们对于隐私信息的认知仍处于保守状态,关于隐私信息如何定价的研究相对较少。文献[14]指出数据保护的重要性,提出关于个人隐私信息的保护策略;文献[15]考虑在网络全球化的环境下把个人信息作为商品,提出立法实践。针对当数据泄露发生后对损失难以量化的问题,文献[16]建立了模型最大化利润,并利用模拟退火算法进行数据定价;文献[17]同样建立了一个

收稿日期: 2020-11-10; 修回日期: 2021-01-08

基金项目: 教育部产学合作协同育人项目(201901137017, 201801034031, 201802257026); 山东省高等学校教学研究与改革面上项目(M2018X066)

第一作者简介: 王瑞桃(1997—),女,山西朔州人,硕士研究生,研究方向为随机控制。E-mail: wangruitao62@163.com

通信作者简介: 宋美(1979—),女,山东济南人,讲师,硕士,研究方向为大数据分析数学模型。E-mail: ytsongmei@163.com

利润最大化模型进行数据定价。

本文将隐私信息定价的分析分为两部分讨论。第一,将个人所有隐私信息作为整体进行定价。考虑到隐私信息价值随着时间的推移会逐渐降低,文中根据信息失效扩散原理,分析隐私信息价值,再基于均衡理论精准计算每个人的隐私信息总价。这种基于市场法和信息时效性的全部隐私信息定价模型,本文将其简称为 MT 模型 (market timeliness total privacy price)。第二,考虑到价格承受者和信息提供者的需求,本文对个人不同类型隐私信息进行定价。因为隐私信息不同指标属性值和偏好权重均为三角模糊数,文中引入文献 [18] 中对方案有偏好的三角模糊数多属性决策方法,求取不同类型信息的权重,最终精确计算不同类型的信息价格。这种基于有偏好的三角模糊数方法建立的部分隐私信息定价模型,本文将其简称为 TF 模型 (triangular fuzzy part privacy price)。

针对隐私信息定价,本文根据实际情况将两个模型结合,将最终的定价模型称为 MT-TF 模型。

## 1 MT 模型

本模型构建的基础是市场完全竞争条件下的供求均衡理论,结合信息扩散失效分析得到基准价格;再针对不同人,考虑多个影响因子,引入个人与行业的影响因素和其他影响因子系数,最终得出准确的且有针对性的隐私定价结果。

### 1.1 信息扩散失效分析

信息作为一种特殊物质,具有很强的时效性,越早获得的信息具有越高的价值。本文根据文献 [19] 将“信息新度”定义为一个概率值,表示某一特定的信息存在到  $t$  时刻(即该信息在  $t$  时刻后失效或者消失)的概率。本文以“信息新度”作为信息扩散的量化指标,由于信息新度和信息寿命密切相关,文中通过分析信息寿命建立信息新度的数学描述。

假设信息的有效周期为  $T$ ,信息的寿命为  $T_0$ ,其中  $T_0 \in [0, \infty)$ ,即信息只在  $[0, T_0)$  内有价值。

对于某个特定的信息而言,其有效周期  $T$  是一个连续型的随机变量,给出相应分布函数

$\vartheta(t)$ :

$$\vartheta(t) = \Pr(T \leq t), t \in [0, T_0)。$$

假设  $\vartheta(t)$  是可导的,则随机变量  $T$  的概率密度函数  $\xi(t)$  满足:

$$\vartheta(t) = \int_0^t \xi(t) dt, t \in [0, T_0)。$$

因为  $\vartheta(t)$  描述信息在过去时间内的分布,信息新度描述信息在剩余时间内的分布,两者具有互补性。所以,信息新度函数  $N(t)$  的表达式为:

$$N(t) = 1 - \vartheta(t) = \Pr(T \geq t), t \in [0, T_0)。 \quad (1)$$

当  $\vartheta(0) = 0$ ,有  $N(0) = 1$ 。值得注意的是,新的信息在产生初期保证有效的概率为 100%。

根据文献 [19],信息新度的特性为:

- 1)  $N(0) = 1, \lim_{t \rightarrow \infty} N(t) = 0$ ;
- 2)  $N(t)$  是单调递减的函数;
- 3)  $N(t)$  是右连续函数。

因此,得到某信息在  $(t_1, t_2]$  区间内失效的条件概率为:

$$\Pr(t_1 < T \leq t_2) = \vartheta(t_2) - \vartheta(t_1) = N(t_1) - N(t_2),$$

$$\Pr(t_1 < T \leq t_2 | T > t_1) = \frac{\Pr(t_1 < T \leq t_2)}{\Pr(T > t_1)} = \frac{N(t_1) - N(t_2)}{N(t_1)} = 1 - \frac{N(t_2)}{N(t_1)}。$$

接下来给出存在至  $t$  时刻的信息中,瞬时失效的信息占已存在至  $t$  时刻的信息的比值  $u_t$ :

$$u_t = \lim_{\Delta t \rightarrow 0^+} \frac{N(t) - N(t + \Delta t)}{\Delta t} \cdot \frac{1}{N(t)} = - \frac{N'(t)}{N(t)}。 \quad (2)$$

改写式(2),得

$$-u_t dt = \frac{dN(t)}{N(t)} = d[\ln N(t)],$$

为方便计算,进行变量替换

$$-u_x dx = \frac{dN(x)}{N(x)} = d[\ln N(x)],$$

对上式两边在  $[t, t+k]$  积分,得:

$$- \int_t^{t+k} u_x dx = \int_t^{t+k} d[\ln N(x)] = \ln \frac{N(t+k)}{N(t)},$$

当  $t=0, k=t$ ,则信息新度的分布函数为:

$$N(t) = e^{-\int_0^t u_x dx}, \quad (3)$$

信息新度的密度函数为:

$$f(t) = [F(t)]' = -[N(t)]' = -u_t e^{-\int_0^t u_x dx}。 \quad (4)$$

在实际操作中,可以将信息新度  $N(t)$  随时间  $t$  的变化过程看作是信息总价值衰减的过程,

某确定信息的价值可以用该时刻已经扩散的信息量占扩散过程中能够达到的最大信息量的比率表示。根据文献[19]提到的信息和节点同势,节点量的变化等价于信息量的变化,本文参考文献[20],用信息节点总量衰减的过程表示信息总价值衰减的过程。

定义信息饱和函数:

$$Q(t) = q(t) + m(t),$$

其中:  $q(t)$  为信息扩散作用下的饱和增加函数,其表达式为:

$$q(t) = \frac{4\pi}{3}(\tau t)^3 (\tau > 0, t > 0),$$

$\tau$  为时间强度参数,  $m(t)$  为信息扩散阻力作用下的饱和衰减函数,表达式为:

$$m(t) = \lambda t^\beta (\lambda < 0, \beta \geq 3, t \geq 0),$$

这里,  $\beta$  为形状参数,  $\lambda$  为强度参数。在进行扩散实验时,信息饱和函数的推广表达式<sup>[20]</sup>为:

$$Q(t) = \tau t^3 + \lambda t^\beta. \quad (5)$$

当  $Q(t) = 0$  时,说明信息衰减的速度等于信息增加的速度,即信息停止扩散。

用  $\Omega(t)$  表示信息的价值,几何意义是信息扩散的节点量随时间  $t$  变化的函数  $f(t)$  与  $x$  轴和  $t$  时刻的时间纵轴围成的面积,则  $N(t)$  与  $\Omega(t)$  的关系可以表示为:

$$N(t) = 1 - \frac{\Omega(t)}{\Omega(t_e)}, \quad (6)$$

其中,  $t_e$  是信息扩散的结束点。通过式(3)和式(6)可计算得到信息时效的结点。

## 1.2 隐私信息总价值的模型计算

基于均衡价格理论,隐私的初始成本是隐私价格的基础,主要包括下面三部分:

1) 时间成本。时间成本主要包括两个方面:一是销售电话或者其他骚扰电话带来的时间消耗;二是需要处理因隐私泄露带来的额外事务所消耗的时间。

2) 财产损失。财产损失主要是受害者因隐私泄露所造成的金钱和其他方面的物质损失。

3) 精神损失。隐私泄露带来的精神损失往往无法度量,但一般比财产损失与时间成本都要高。精神损失大体上可以量化为心理治疗的费用、工作与晋升机会的丧失造成的损失和误工费用的总和。

综合考虑,隐私的初始成本可以表示为:

$$P_0 = \sum_{i=1}^3 p_i, \quad (7)$$

其中:  $P_0$  为隐私的初始成本,  $p_i, i = 1, 2, 3$  分别表示时间成本、财产损失和精神损失。

结合信息的时效性,构建个体资产定价的基本模型,其一般表达式为:

$$P = A \cdot B \cdot C \cdot D \cdot E \cdot P_0 \cdot \int_0^{t_e} N(t) dt, \quad (8)$$

其中:  $\bar{P} = P_0 \cdot \int_0^{t_e} N(t) dt$  表示该个体所在行业的隐私基准价格;  $A$  表示个人形象系数,  $B$  表示个体的资产状况系数,  $C$  表示个体所在行业的比较系数,  $D$  表示个体的发展前景系数,  $E$  表示不确定因素。各个系数的具体说明如下:

个人形象系数: 形象分数包括容貌气质、个人信誉和工作资历,用来表征个人的直观形象,该影响因子的最终计算公式为  $A = A^* / \bar{A}$ , 其中,  $A^*$  为专家对个人的打分,  $\bar{A}$  为该项的大众基准分数;

资产状况系数: 通过个人的净资产与行业平均资产的比值,表征个人资产的数量水平;

行业比较系数: 通过个体工资与个体所在行业的平均工资的比值,表示个体在行业内的层次;

发展前景系数: 发展前景用来表示个人在未来可能创造的价值,包括个人所在行业的发展前景、学历和业绩,其中,个人所在行业的发展前景的计算方法为个人所在行业的平均工资与全国平均工资的比值;

不确定因素: 表示对个人产生重大影响的特殊因素,如果没有特殊事件(如特殊疾病患者、特殊行业工作人员或涉密人员)对自身产生重大影响,该项通常取 1。

通过对基准价格的修正,将  $A, B, C, D$  和  $E$  代入式(8),可以得到最终的隐私价格。

## 2 TF 模型

### 2.1 模型建立

隐私信息的定价是一个多目标决策问题,由于种种条件限制,包括价格承受者的偏好与客观价格之间的偏差,模型需要在隐私的总价格上再做调整,制定出一个合理的最终定价,使定价不仅能体现隐私的市场价值,也可以最大程度上满足定价承受者的偏好。

由于属性权重无法精确求知,隐私信息的市场价值和价格承受者的偏好值是三角模糊数的多属性决策问题。本节采用文献 [21] 中的决策方法,同时考虑市场价值与价格承受者偏好的总偏差最小化,建立规划模型,得到不同类型隐私信息权重。模型框架如图 1 所示。

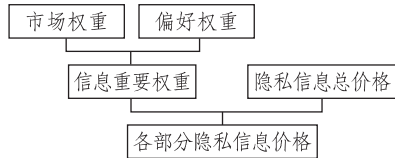


图 1 部分隐私定价模型框图

Fig.1 Block diagram of partial privacy pricing model

对于不同属性隐私信息定价问题,算法具体执行过程如下:

1) 设  $X = \{x_1, x_2, \dots, x_n\}$  为指标集,代表不同类型的个人隐私;  $Y = \{y_1, y_2, \dots, y_m\}$  为属性集,代表个人隐私的各个属性;  $\omega = (\omega_1, \omega_2, \dots, \omega_m)$  为属性的权重,其中  $\omega_j \geq 0, \sum_{j=1}^m \omega_j = 1$ 。根据大数据分析给出不同类型信息  $x_i$  对于各属性  $y_i$  的属性值  $a_{ij} = (a_{ij}^L, a_{ij}^M, a_{ij}^U)$ , 则模糊决策矩阵  $A = (a_{ij})_{m \times n}$ 。市场价格承受者对类型  $x_i$  的隐私有一定的主观偏好,其主观偏好值  $v_j = (v_j^L, v_j^M, v_j^U)$ ,  $0 \leq v_j^L \leq v_j^M \leq v_j^U \leq 1$ 。

2) 按照效益型、成本型的属性处理方法,将模糊决策矩阵  $A$  转化为规范化后的决策矩阵  $R = (r_{ij})_{nm}$ , 其中,  $r_{ij} = (r_{ij}^L, r_{ij}^M, r_{ij}^U)$  是三角模糊数,表示  $x_i$  关于属性  $y_i$  的属性值,计算法则如下:

$$r_{ij} = \frac{a_{ij}}{\|a_i\|},$$

$$\frac{1}{r_{ij}} = \frac{(1/a_{ij})}{\|1/a_i\|},$$

$$\|a_i\| = \sqrt{\sum_{j=1}^n a_{ij}^2},$$

$$\|1/a_i\| = \sqrt{\sum_{j=1}^n (1/a_{ij})^2}.$$

根据相似度的定义,  $a$  与  $b$  的相似程度  $s(a, b)$  为:

$$s(a, b) = \frac{a^L b^L + a^M b^M + a^U b^U}{\max \{ (a^L)^2 + (a^M)^2 + (a^U)^2, (b^L)^2 + (b^M)^2 + (b^U)^2 \}}.$$

用函数  $F(\omega)$  表示  $x_i$  类型隐私的客观价值  $R(r_{ij})$  与价格接受者偏好之间的总相似度。引入偏差函数  $F(\omega)$ , 将多目标决策问题转化为以下单目标

决策问题:

$$\begin{cases} \max F(\omega) = \sum_i^n \sum_j^m s(r_{ij}, v_j) \omega_i, \\ \text{s.t. } \omega = (\omega_1, \omega_2, \dots, \omega_n) \in H, \\ \omega_i \geq 0, \sum_i^n \omega_i = 1, \end{cases} \quad (9)$$

其中,  $H$  为已知隐私属性的权重的集合。

### 2.2 隐私信息属性的划分

基于各信息的领域特征以及其社会影响力, 本文将个人信息划分为六类: 金融信息、身份信息、生物特征信息、通讯信息、健康信息、地理位置信息。同时, 每一类信息的价值都包含在以下四个属性中:

新闻价值: 指个人信息在社会新闻中引发的社会关注程度;

风险指数: 指个人信息被泄露的可能性, 衡量该类信息的被保密程度;

商业价值: 指个人信息被应用到商业活动中产生的经济价值;

社会影响: 指个人信息对公众的思想、感觉、态度或行为所产生的影响和效果。

本文将基于以上的信息类别和信息属性, 利用均衡价格理论, 并加入信息的时效性来确定信息的价值。同时, 考虑价格承受者和实际价值的多方面问题, 利用多目标决策完成信息定价。

### 3 定价模型的实证分析

隐私定价模型具有很高的实际意义, 法律辅助作用是其中非常重要的部分。当隐私泄露的受害者向司法机关对侵害者提起诉讼时, 隐私定价模型可以作为一个可靠依据, 帮助法官计算赔偿金额。

本文以“中国网络暴力第一案”为例介绍定价模型。王某是案件当事人, 2008 年初, 因感情原因, 其身份、地址、照片等隐私信息被泄漏, 使其承受巨大的舆论压力, 并失去工作; 之后, 王某对泄露其信息的网络公司提起诉讼, 获赔 13 684 元<sup>[21]</sup>。

#### 3.1 数据来源

王某, 大专学历, 北京市盛世长城广告公司普通职员, 工作 6 年。2008 年泄密事件案发生时, 王某月收入为 2974 元, 总资产约为 27.2 万元<sup>[22]</sup>。

根据文献 [23] 的数据显示, 在 2008 年, 中国商

务服务业就业人员年平均工资为32 422元,年平均资产为35 494元,全国就业人员年平均工资为32 244元。根据文献 [24]的数据,平均每人每天会接到1个骚扰电话,平均通话时长 29 s。根据《中华人民共和国 2017 年上半年网络诈骗数据研究报告》显示,平均每起案件的受害者经济损失约为11 641.7元。国家二级心理咨询师诊疗费用 6 元/min,一般治疗方案为 10 次,每次 3 h。

### 3.2 计算成本

初始成本(单位:元)中三项分别为:时间成本  $p_1 = \frac{35\,494 \times 29 \times 30}{8 \times 3600 \times 365 \times 5/7} = 4.11$ ,财产损失  $p_2 = 11\,641.7/6 = 1\,940.3$ ,精神损失  $p_3 = 10\,800$ 。

由式(7)知,隐私的初始成本为:

$$P_0 = \sum_{i=1}^3 p_i = 12\,744.41。 \quad (10)$$

根据 2008 年的网站数据,结合式(4)得到信息新度的密度函数为

$$f(t) = 1.470e^{-((t-10.49)/6.068)^2} - 0.1229e^{-((t-3.666)/2.689)^2},$$

结合式(5)可以得到信息饱和函数:  $Q(t) = 0.0459t^3 - 0.02835t^{3.147}$ ,进一步得出,当  $t_e = 26.5$ (单位:月)时,信息扩散结束。从信息产生到信息结束扩散的这段时间内,信息新度的总密度为  $\Omega_e = \int_0^{26.5} f(t) dt = 15.7335$ ,代入式(6),得到:

$$N(t) = 1 - \frac{1.470e^{-((t-10.49)/6.068)^2}}{15.7335} - \frac{0.1229e^{-((t-3.666)/2.689)^2}}{15.7335}。 \quad (11)$$

下面根据案例中的数据,计算各项影响因子系数。

#### 1) 个人形象系数 A

个人形象分数共包含三方面:容貌气质、个人信誉和工作资历,每个方面的基准分数为 5 分。王某在该行业工作 6 年,跳槽 2 次,学历水平偏低,考虑到王某在广告公司工作,评价人员在容貌气质和工作资历方面给他的打分是 3 分和 5 分;同

时,王某在短时间内跳槽频繁,且在婚姻方面有负面事件,评价人员在个人信誉方面的打分是 1 分。则形象指数为  $A = \frac{3 + 5 + 1}{15} = 0.6$ 。

#### 2) 资产状况系数 B

用个人资产和同行业平均资产比值表示资产状况系数,其资产状况指数为

$$B = \frac{27.2}{7.0092} = 3.8806。$$

#### 3) 行业比较系数 C

用个人工资和全国平均工资的比值表示行业比较系数,其行业比较指数为

$$C = \frac{32\,422}{32\,244} = 1.0055。$$

#### 4) 发展前景系数 D

发展前景包括个人所在行业的发展前景、个人的学历和个人的业绩。同行业平均工资和全国平均工资比值表示行业发展前景系数,商务服务业的发展前景指数为  $D_1 = \frac{35\,494}{32\,244} = 1.1008$ ;王某学历较低,学历影响系数  $D_2 = 0.8$ ;个人业绩指数记为  $D_3 = 1$ 。所以,发展前景系数为

$$D = D_1 \cdot D_2 \cdot D_3 = 0.8806。$$

#### 5) 不确定因素 E

个人没有特殊事件,  $E = 1$ 。

综合以上影响因子系数,结合式(8)、(10)和(11),得隐私信息的价值

$$P = A \cdot B \cdot C \cdot D \cdot E \cdot P_0 \int_0^{t_e} n(t) dt = 65\,575.42, \text{单位为元。}$$

### 3.3 各类隐私价格

本文通过社会调查结果来确定各类隐私价值的权重,如表 1 所示,各类隐私的市场比重为模糊数集。根据王某对法院提出的诉求,可推断其对各种隐私信息的偏好权重如表 2 所示,该结果为模糊数集。

表 1 各类隐私在各项属性下的市场权重范围

Tab.1 The market weight range of all kinds of privacy under each attribute

属性名称	新闻价值	风险指数	商业价值	社会影响
金融信息	(0.35,0.4,0.45)	(0.8,0.85,0.9)	(0.75,0.8,0.85)	(0.7,0.75,0.8)
身份信息	(0.2,0.25,0.3)	(0.1,0.15,0.2)	(0.3,0.35,0.4)	(0.5,0.55,0.6)
生物特征信息	(0.3,0.35,0.4)	(0.15,0.2,0.25)	(0.3,0.35,0.4)	(0.25,0.3,0.35)
通讯信息	(0.7,0.75,0.8)	(0.7,0.75,0.8)	(0.7,0.75,0.8)	(0.6,0.65,0.7)

健康信息	(0.45,0.5,0.55)	(0.55,0.6,0.65)	(0.75,0.8,0.85)	(0.35,0.4,0.45)
地理位置信息	(0.2,0.25,0.3)	(0.3,0.35,0.4)	(0.5,0.55,0.6)	(0.35,0.4,0.45)

表 2 价格承受者对不同属性的期望权重范围

Tab.2 The expected attribute values of the price acceptor for different indicators

信息类型	期望权重	信息类型	期望权重
金融信息	(0.30,0.35,0.40)	通讯信息	(0.60,0.65,0.70)
身份信息	(0.75,0.80,0.85)	健康信息	(0.45,0.50,0.55)
生物特征信息	(0.60,0.65,0.70)	地理位置信息	(0.85,0.90,0.95)

将表 1、2 中数据代入式(9),利用 Lingo 软件求解该优化模型,得出最终各类信息的重要性权重结果,并计算各部分的隐私定价,如表 3 所示。

表 3 各类信息的重要性权重和最终定价

Tab.3 The weight of information and the final price

信息类型	权重	定价/元
金融信息	0.239	15 671
身份信息	0.107	7016
生物特征信息	0.103	6 754
通讯信息	0.214	14 032
健康信息	0.202	13 244
地理位置信息	0.134	8786

在该案中,王某被泄露身份信息与地理位置信息,故应得赔偿款 15 802 元,案例中法院判决王某获赔 13 684 元,两者结果吻合度较高。

在模型计算的过程中,部分评价系数(如个人形象指数等)具有一定的主观性,收集的数据、区域等偏差可能会使结果产生一定误差。而在案件审理时,司法机关会充分考虑地方传统、社会影响与人伦道德等因素,会与模型计算形成很好的互补,所以该模型对于法律案件的审理有很好的辅助作用。

## 4 结语

本文从三个方面较为全面地确定了隐私信息初始价值。不同于其他方法,本文考虑了隐私信息的时效性,将初始价值精确为隐私信息的时效性价值;同时,本文又利用市场法由时效性价值进一步精确得出个人隐私总价值。考虑到现实很多情况需要得到个人信息每部分的价格,本文将偏好权重与市场权重相结合,得到符合实际情况的部分隐私信息的售卖价格。

该模型在实际案例中的应用,体现其合理性高且人性化强的特点。由于隐私信息是一个存在已久的问题,其价值考量还必须考虑当地的传统

文化、思想习俗以及人伦道德等因素,具体应用时也必须进行多方面的参数修正,所以 MT-TF 模型的普适性将是未来完善的主要方向。

## 参考文献:

- [1] 孙云龙.隐私成本的模型研究[J].数学建模及其应用,2018,7(4):64-74.
- [2] 任翠萍,张俊丽.面向大数据基于信息熵的隐私成本定价系统[J].赤峰学院学报(自然科学版),2018,34(7):62-63.
- [3] RUST R T, KANNAN P K, PENG N. The customer economics of internet privacy[J]. Journal of the Academy of Marketing Science, 2002, 30(4): 455-464.
- [4] 马汝秋,刘伟男,杨俊.基于层次分析的隐私模糊综合评价模型[J].中国高新区,2018(13):60.
- [5] 李兆碧.基于层次分析法的大数据时代隐私价格模型的建立与分析[J].科技与创新,2018(17):110-111.
- [6] MACHANAVAJJHALA A, KIFER D, GEHRKE J, et al. L-diversity: privacy beyond k-anonymity[J]. ACM Transactions on Knowledge Discovery from Data, 2006: 24-35.
- [7] LI N, LI T, VENKATASUBRAMANIAN S. t-Closeness: privacy beyond k-anonymity and l-diversity[C]// Proceeding of the 23rd International Conference on Data Engineering, 2007: 106-115.
- [8] 史玉良,邵雅丽,周中民,等.基于帕累托最优的隐私保护服务定价模型[J].计算机学报,2016,39(6):1267-1280.
- [9] 史武超,吴振强,刘海.一种基于 VCG 机制的差分式隐私服务定价机制[J].计算机技术与发展,2017,27(6):119-123.
- [10] 申琦.风险与成本的权衡:社交网络中的“隐私悖论”:以上海市大学生的微信移动社交应用(APP)为例[J].新闻与传播研究,2017,24(8):55-69.
- [11] 彭慧波,周亚建.基于隐私度量的数据定价模型[J].软件,2019,40(1):57-62.
- [12] 黄启发,朱建明,宋彪,等.社交网络用户隐私保护的博弈模型[J].计算机科学,2014,41(10):184-190.

- [13] 郭龙飞. 社交网络用户隐私关注动态影响因素及行为规律研究 [D]. 北京: 北京邮电大学, 2013.
- [14] PURTOVA N. The law of everything: broad concept of personal data and future of EU data protection law [J]. *Law, Innovation and Technology*, 2018: 40–81.
- [15] KANG J. Information privacy in cyberspace transactions [J]. *Stanford Law Review*. 1998, 50(4): 1193–1294.
- [16] YANG J, XING C. Personal data market optimization pricing model based on privacy level [J]. *Information (Switzerland)*, 2019, 10(4): 123.
- [17] 董祥千, 郭兵, 沈艳, 等. 基于利润最大化的数据资产价值评估模型 [J]. *大数据*, 2020, 6(3): 13–20.
- [18] 徐泽水. 对方案有偏好的三角模糊数型多属性决策方法研究 [J]. *系统工程与电子技术*, 2002(8): 9–12.
- [19] 刘志斌, 赵金楼, 于剑男. 网络信息扩散时效分析及其建模 [J]. *情报杂志*, 2012, 31(4): 122–126.
- [20] ZHAO J, LIU Z, YU J. Saturation analysis of network information diffusion and modeling [J]. *Advances in Information Sciences and Service Sciences*, 2011, 3(6): 237–244.
- [21] 法律快车. 2018年最新侵犯个人隐私权案例 [EB/OL]. (2018-07-26) [2019-05-18]. <https://www.lawtime.cn/info/minfa/qlrsyinsi/201808093396770.html>.
- [22] 姜岩事件 [EB/OL]. (2011-12-29) [2019-05-18]. <https://baike.so.com/doc/6695128-6909037.html>.
- [23] 中华人民共和国国家统计局. 统计数据 [DB/OL]. (2018-12-31) [2019-05-18]. <http://www.stats.gov.cn/>.
- [24] 南方日报. 2016年中国骚扰电话形势分析报告 [EB/OL]. (2016-09-28) [2019-05-18]. [http://tech.southcn.com/t/2016-10/09/content\\_157130323.htm](http://tech.southcn.com/t/2016-10/09/content_157130323.htm).

## Personal Privacy Pricing Based on MT-TF Model

WANG Ruitao<sup>1a</sup>, YU Lindong<sup>1b,2</sup>, YU Hengyue<sup>1b</sup>, LIU Guangchen<sup>1a</sup>, SONG Mei<sup>1a</sup>

(1.a.School of Mathematics and Statistics Science; b.School of Information and Electrical Engineering, Ludong University, Yantai 264039, China;

2.School of Electrical Engineering & Information, Northeast Petroleum University, Daqing 163318, China)

**Abstract:** The era of big data not only brings convenience to people, but also causes people to pay more attention to their private data information due to information leakage. In this paper, a privacy information pricing model, i.e. MT-TF model, was proposed based on the market timeliness and triangular fuzzy number to quantify individual privacy information. The model determined the initial value of privacy information from three aspects including time cost, property loss and mental loss. It used the information newness degree function to determine the timeliness value of privacy information and used the market method based on equilibrium theory to accurately quantify the value of personal privacy. In order to obtain partial information price, the model used triangular fuzzy number to acquire corresponding information weight. The test results of the model example show that the court judgment in a real personal privacy disclosure case is 13 684 yuan, the model calculation is 15 802 yuan, and the overall effect is good. The MT-TF model can better complete the accurate calculation of personal privacy information pricing, and provide strong technical support for judicial, security and legal privacy transactions in China.

**Keywords:** privacy data; MT-TF; market approach; information timeliness; triangular fuzzy number

(责任编辑 顾建忠)