

基于多层模糊综合评估的隐私保护效果评估方法

俞艺涵^{1,2}, 付钰², 吴晓平²

(1. 海军航空航天大学教练机模拟训练中心, 辽宁 葫芦岛 125001;

2. 海军工程大学信息安全系, 湖北 武汉 430033)

摘要: 针对由单一数据层面进行隐私保护效果评估无法刻画网络环境整体隐私保护效果的问题, 借助模糊综合评估方法, 从数据全生命周期凝练隐私保护效果影响因子, 通过分区块、多层次地模糊综合评估, 并依据数据隐私差异确定各因子与各区块权重, 综合计算得到网络环境整体隐私保护效果评估结果。实例分析表明, 该方法可行有效。

关键词: 隐私保护; 隐私保护效果评估; 模糊综合; 多层次

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2020082

Evaluation method of privacy protection effect based on multi-layer fuzzy comprehensive evaluation

YU Yihan^{1,2}, FU Yu², WU Xiaoping²

1. Flight Simulation Training Center, Navy Aviation University, Huludao 125001, China

2. Naval University of Engineering, Wuhan 430033, China

Abstract: Aiming at the problem that the evaluation of the privacy protection effect from a single data level cannot describe the overall privacy protection effect of the network environment, with the help of the fuzzy comprehensive evaluation method, the privacy protection effect influencing factors were condensed from the entire life cycle of the data. According to the data privacy differences, the factors and block weights were determined, and the overall calculation results of the overall privacy protection effect of the network environment were obtained. Case analysis shows that the proposed method is feasible and effective.

Key words: privacy protection, evaluation of privacy protection effect, fuzzy synthesis, multi-level

收稿日期: 2020-05-08; 修回日期: 2020-08-20

通信作者: 俞艺涵, chenike1992@163.com

基金项目: 国家自然科学基金(61672531); 网络空间安全重点专项(SQ2018YFGX210002-04)

Foundation Items: The National Natural Science Foundation of China(61672531), Cyberspace Security Special Project(SQ2018YFGX210002-04)

论文引用格式: 俞艺涵, 付钰, 吴晓平. 基于多层模糊综合评估的隐私保护效果评估方法[J]. 网络与信息安全学报, 2020, 6(6): 121-127.

YU Y H, FU Y, WU X P. Evaluation method of privacy protection effect based on multi-layer fuzzy comprehensive evaluation[J]. Chinese Journal of Network and Information Security, 2020, 6(6): 121-127.

1 引言

随着信息技术的不断发展,网络环境中的数据呈爆炸式增长。海量数据中不可避免地存在众多隐私信息,如何对开放式网络环境中的隐私信息进行保护成为近年来的研究热点^[1]。当前,针对数据生命周期各个阶段已有较多的隐私保护研究成果,如隐私保护数据发布、隐私保护数据挖掘等,但对各类隐私保护方法的实际效果缺乏规范化的评价方法。

当前对隐私保护效果的评价主要集中在数据自身层面进行,由数据经隐私保护后的隐私性和可用性作为其评价指标。一般情况下,对原始数据采取一定的隐私保护措施后,数据的隐私性将增加,而可用性将降低。以尽可能少的数据可用性损失取得尽可能多的数据隐私性增量时,隐私保护效果较好。数据隐私性可理解为隐私保护强度,隐私保护强度越大,数据隐私性越高。通常,针对不同的隐私保护方法采取相应的度量方法,举例来说,针对基于匿名的隐私保护,可通过攻击者获得隐私信息的概率来度量数据隐私性,就 k -匿名^[2]而言,攻击者获得隐私信息的概率不超过 $\frac{1}{k}$,即以参数 k 的大小度量数据隐私性;针对基于差分的隐私保护^[3],可直接利用差分隐私预算来度量数据隐私性。数据可用性可解释为数据经隐私保护后的数据质量。隐私保护处理数据将引起数据质量的损失,通过计算数据质量损失量来度量数据可用性,度量方法主要包括可辨别度量(DM)^[4]、分类度量(CM)^[5]、数据信息损失(ILoss)^[6]以及标准化确定性惩罚(NC)^[7]。在得到数据隐私性和可用性的度量结果后,可设定权重参数通过线性加权法得到最终的数据隐私保护效果评估结果。然而,脱离数据实际应用环境与隐私场景而单一地从数据自身角度对隐私保护效果进行评估缺乏客观性,网络环境中整体隐私保护效果评价结果较难整合。另外,当前缺乏对隐私概念的准确定义,一条数据是否涉及隐私受时空及数据应用环境影响较大。在对隐私保护效果进行评估时,隐私概念的模糊性将给实际数据信息与评估指标对应关系的确定造成困难。

针对以上问题,本文结合数据实际应用环境和隐私场景对隐私保护效果的影响因素进行归纳,并考虑到隐私概念的模糊性,设计了一种基于多层模糊综合评估的隐私保护效果评估方法,该方法将隐私保护效果评估由单一数据自身层面扩展至整个数据生命周期,由多层次、分区块的隐私保护效果评估结果整合计算得到网络环境整体隐私保护评估效果,可为下一步指导隐私保护具体实施提供依据。

2 隐私保护效果影响因子

隐私保护效果评估需以明确隐私保护效果影响因子为前提,文献[8]将对隐私保护效果的评估刻画至整个隐私生命周期中,将其与以下几大因素相关联:一是隐私信息的概率分布;二是信息所有者;三是信息接收者,主要指信息接收者所拥有的背景知识、兴趣点、隐私观念等;四是信息使用时的约束条件,即数据使用在时间、空间、环境条件上的限制;五是隐私操作,即隐私保护措施。本文以此为基础,综合考虑当前网络环境下隐私数据所面临的安全风险总结得出隐私保护效果的3个主要影响因子,每个影响因子包含文献[8]所提的多个因素且存在交叉,具体如下。

(1) 数据隐私类别

数据隐私类别是指数据隐私信息中所蕴含的隐私信息的类别,其对隐私保护效果的影响具有本质性。数据隐私类别一方面决定数据在隐私层面的重要性,另一方面会影响隐私所有者、数据使用者、数据攻击者在不同隐私应用场景下对数据的隐私态度与处理方法,从而影响隐私保护效果。若将隐私保护的主体视为一种资产,则数据隐私类别是对该资产的一种量化,本文将数据隐私类别对隐私保护效果产生的影响称为资产因子。

(2) 隐私风险来源

隐私风险来源是指隐私数据在整个生命周期中可能面临隐私泄露风险的原因,其与数据隐私保护效果具有间接关联性。以数据隐私安全全局角度来看,当数据隐私面临的风险越高时,对于已实施的隐私保护措施的隐私保护效果应给予越低的评价。考虑到隐私风险间接地影响隐私保护实际效果,本文将隐私风险来源对隐私保护效果

产生的影响称为威胁因子。

(3) 隐私保护实施

隐私保护实施是指隐私数据在整个生命周期中所被采取的隐私保护措施，其对于隐私保护效果的影响具有直接性。隐私保护实施采用的隐私保护技术、隐私保护强度、隐私保护范围等因素将影响隐私信息的泄露概率，泄露概率越高，隐私信息的脆弱性越大，本文将隐私保护实施对隐私保护效果产生的影响称为脆弱性因子。

3 基于多层模糊综合评估的隐私保护效果评估流程

在实际评估过程中，上述3个隐私保护效果影响因子分别对应多个方面十个指标，因此，在采用模糊综合评估方法对隐私保护效果进行评价时，因评估因素集中的因素数量过多，不可避免地将出现一些因素的权值取值较小（接近于零）的情况，即在计算综合评价结果时丢失对应因素信息。另外，在当前网络环境中对数据隐私保护效果进行评估时，评估对象类似于一个无边界的复杂系统，很难厘清一个因素与评语集对应的隶属关系，从全局角度进行评估将丢失用户个体差异信息，使评估结果缺乏准确性。

综上，在利用模糊综合评估方法对隐私保护效果进行评估时，为使评估结果尽可能地科学、准确，本文利用模糊综合评估的多级分层模型展开具体评估工作，具体步骤如下。

Step1 确定评估数据的最小评估单元 d 。将待评估数据 D 按最小评估单元划分，记为 $D = \{d_1, d_2, \dots, d_n\}$ 。

Step2 构建评估指标集。分别以威胁因子、资产因子和脆弱性因子组成隐私保护效果评估因素集，记为 $U = \{U_T, U_I, U_P\}$ 。将评估因素集进一步划分，具体如下。

一级评估因素 $U = \{U_T, U_I, U_P\}$ 。

二级评估因素包括威胁因子评估指标中的个人网络行为、不可信服务商、各类攻击敌手；资产因子评估指标中的人口统计学信息、财务信息、通信信息和偏好信息；脆弱性因子评估指标中的用户端、网络服务商和网络第三方，记为 $U' = \{U_{T_1}, U_{T_2}, U_{T_3}, U_{I_1}, U_{I_2}, U_{I_3}, U_{I_4}, U_{P_1}, U_{P_2}, U_{P_3}\} = \{U'_1,$

$U'_2, U'_3, U'_4, U'_5, U'_6, U'_7, U'_8, U'_9, U'_{10}\}$ ，并根据具体评估需求，确定 U' 中各因素的具体指标，形成隐私保护效果评估指标集。

Step3 将隐私保护效果评估评语集分为5个级别，分别为优、良、中、弱、差，记为 $V = \{V_1, V_2, V_3, V_4, V_5\}$ 。

Step4 在所有 d_i 中，根据二级评估因素对应指标构造对应的模糊评估矩阵 $R_1, R_2, \dots, R_{10}$ ，其中，隶属度的确定方法需根据指标性质进行选择，如用户隐私观念、服务商隐私政策等模糊性较高的指标通常由专家经验法得到其隶属度；资产因子中的指标可由模糊统计法得到其隶属度；一些可计算的指标则根据具体计算结果与评语集的对应关系得到其隶属度，如当采取 k -匿名进行隐私保护时，首先，计算其隐私泄露概率确定隐私保护实施指标与评语集的隶属度关系；然后，由二元决策法得到其对应的权值向量 $A_1, A_2, \dots, A_{10}$ ，并根据算式计算出二级因素评估结果。

$$\begin{cases} B_1 = A_1 \odot R_1 \\ B_2 = A_2 \odot R_2 \\ \vdots \\ B_{10} = A_{10} \odot R_{10} \end{cases} \quad (1)$$

根据二级因素评估结果，将其作为行向量构造对应一级因素模糊评估矩阵。

$$R_T = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} \quad R_I = \begin{pmatrix} B_4 \\ B_5 \\ B_6 \\ B_7 \end{pmatrix} \quad R_P = \begin{pmatrix} B_8 \\ B_9 \\ B_{10} \end{pmatrix} \quad (2)$$

由二元决策法得到其对应的权值向量 A_T, A_I, A_P ，并根据算式计算出一级因素评估结果。

$$\begin{cases} B_T = A_T \odot R_T \\ B_I = A_I \odot R_I \\ B_P = A_P \odot R_P \end{cases} \quad (3)$$

根据一级因素评估结果，构造 d_i 的模糊评估矩阵。

$$R_{d_i} = \begin{pmatrix} B_T \\ B_I \\ B_P \end{pmatrix} \quad (4)$$

根据隐私保护效果评估权值向量 A_d 计算 d_i 的评估结果。

$$B_{d_i} = A_d \odot R_{d_i} \quad (5)$$

A_d 的确定由评估时的具体情况和数据隐私差异决定。 B_{d_i} 的计算过程如图 1 所示。

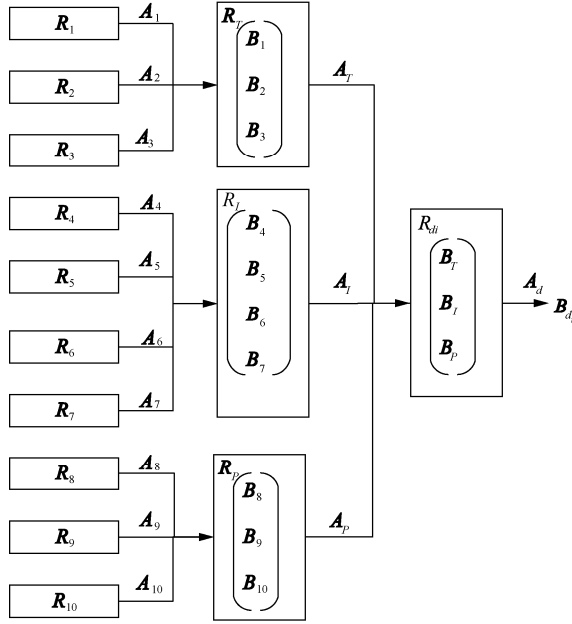


图 1 最小评估单元评估过程层次
Figure 1 The minimum evaluation unit evaluates the process hierarchy

Step5 构建待评估数据 D 的模糊评估矩阵。

$$R_D = \begin{pmatrix} B_{d_1} \\ B_{d_2} \\ \vdots \\ B_{d_n} \end{pmatrix} \quad (6)$$

根据文献[9]中方法计算 d_i 的隐私等级并进行归一化处理, 由此作为 D 的评估权值向量 A_d 的生成依据, 赋予隐私等级高的单元更高的权值。由此将评估过程中涉及权值计算的过程均依据数据隐私差异进行, 将数据隐私差异对隐私保护效果评估的影响综合到评估全过程中。计算 D 的最终评估结果如下。

$$B_D = A_D \odot R_D \quad (7)$$

得到待评估数据集 D 的最终评估结果 B_D 后, 可根据最大隶属度原则对当前隐私保护效果做出判断。同时, B_D 保持着良好的可扩展性。从空间维度上来看, D 作为网络环境中的一部分数据, 可由 B_D 作为底层因素构建上一层的模糊判断矩阵, 根据上述步骤按层次进一步对大数据环境下

的隐私保护效果进行评估, 如图 2 所示。从时间维度上来看, 对同一待评估数据集在不同时刻的评估结果可进行差值计算, 并通过计算结果得到这一时间段内隐私保护效果的变化情况。另外, 通过多层次的隐私保护效果评估可以发现影响隐私保护效果的主要因素, 由此可为针对性的隐私保护资源分配提供便利。

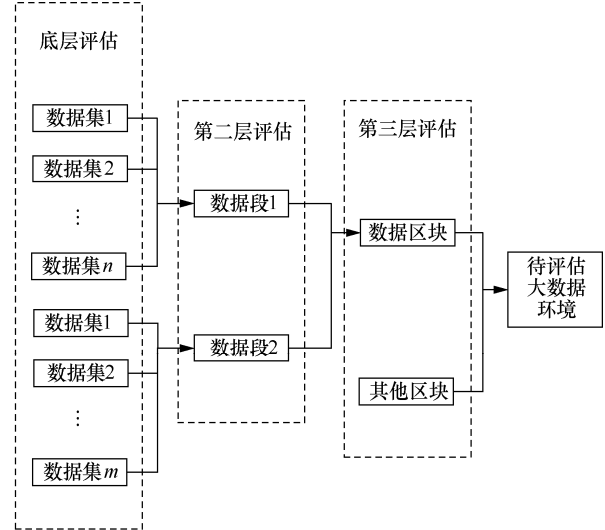


图 2 空间维度上的评估扩展示意
Figure 2 Schematic diagram of evaluation extension on spatial dimension

4 实例分析

公开临床医疗数据有利于各类医学问题的研究, 但在公开的临床医疗数据集中通常会涉及患者隐私信息。目前存在众多临床医疗数据集, 其保护患者隐私的方法各不相同。本节以所提基于多层模糊综合评估的隐私保护效果评估方法对真实医疗数据集 MIMIC-III^[10] 的隐私保护效果进行评估。

Step1 随机抽取 MIMIC-III 数据集的 100 条记录进行评估, 以每 10 条记录作为最小评估单元, 记为 $D = \{d_1, d_2, \dots, d_{10}\}$ 。

Step2 针对临床医疗数据集隐私保护效果评估需求, 生成如图 3 所示隐私保护效果评估指标集。

Step3 将临床医疗数据集隐私保护效果评估评语集分为优、良、中、弱、差 5 个等级, 记为 $V = \{V_1, V_2, V_3, V_4, V_5\}$, 其具体定义如表 1 所示。

Step4 计算 $D = \{d_1, d_2, \dots, d_{12}\}$ 的模糊评估结果。以 d_1 为例, 通过专家经验法和模糊统计法

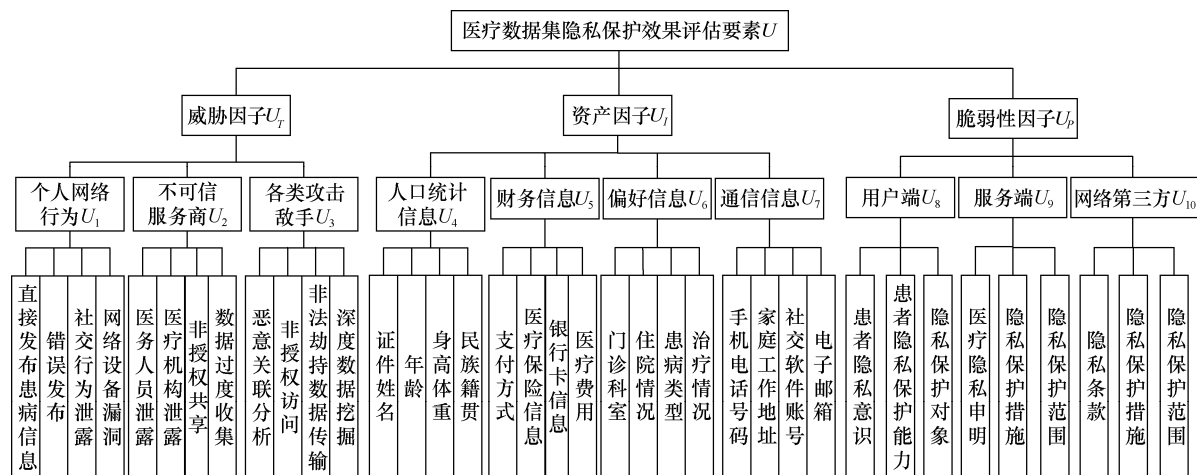


图3 临床医疗数据集隐私保护效果评估指标集

Figure 3 Evaluation index set of privacy protection effect of clinical medical data set

表1 临床医疗数据集隐私保护效果评估评语集定义

Table 1 Definition of comment set for evaluation of privacy protection effect of clinical medical data set

等级	定义	数值范围
优	隐私保护效果极好，患者隐私信息极难泄露，泄露风险可忽略不计	0.90~1.00
良	隐私保护效果较好，患者隐私信息较难泄露，泄露风险较低	0.75~0.90
中	隐私保护效果一般，患者隐私信息泄露可能性中等，泄露风险一般	0.55~0.75
弱	隐私保护效果较差，患者隐私信息泄露容易泄露，泄露风险较高	0.30~0.55
差	隐私保护效果极差，隐私保护缺失或不起效，患者隐私泄露风险极高	0~0.30

得到其对应的模糊评估矩阵，其中资产要素下的模糊评估矩阵通过统计记录中出现对应信息的频率与评语集等级对应的数值对照生成。

$$\begin{aligned}
 R_1 &= \begin{bmatrix} 0.98 & 0.01 & 0.01 & 0 & 0 \\ 0.95 & 0.03 & 0.01 & 0.01 & 0 \\ 0.96 & 0.03 & 0.01 & 0 & 0 \\ 0.96 & 0.04 & 0 & 0 & 0 \end{bmatrix} \\
 R_2 &= \begin{bmatrix} 0.97 & 0.02 & 0.01 & 0 & 0 \\ 0.98 & 0.02 & 0 & 0 & 0 \\ 0 & 0 & 0.11 & 0.25 & 0.64 \\ 0.38 & 0.24 & 0.26 & 0.11 & 0.01 \end{bmatrix} \\
 R_3 &= \begin{bmatrix} 0.07 & 0.06 & 0.21 & 0.10 & 0.56 \\ 0.14 & 0.09 & 0.17 & 0.12 & 0.48 \\ 0.09 & 0.09 & 0.11 & 0.09 & 0.62 \\ 0.12 & 0.04 & 0.06 & 0.07 & 0.71 \end{bmatrix} \\
 R_4 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\
 R_5 &= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 R_6 &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\
 R_7 &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

$$R_8 = \begin{bmatrix} 0.22 & 0.21 & 0.13 & 0.19 & 0.25 \\ 0 & 0 & 0 & 0.16 & 0.84 \\ 0 & 0.12 & 0.23 & 0.19 & 0.46 \end{bmatrix}$$

$$R_9 = \begin{bmatrix} 0 & 0 & 0.01 & 0.21 & 0.78 \\ 0.36 & 0.27 & 0.33 & 0.04 & 0 \\ 0.82 & 0.05 & 0.03 & 0.04 & 0.06 \end{bmatrix}$$

$$R_{10} = \begin{bmatrix} 0.98 & 0.02 & 0 & 0 & 0 \\ 0.36 & 0.27 & 0.33 & 0.04 & 0 \\ 0.82 & 0.05 & 0.03 & 0.04 & 0.06 \end{bmatrix}$$

由二元对比决策法得到模糊评估矩阵对应的

权值向量。

$$\begin{aligned} A_1 &= [0.65, 0.15, 0.1, 0.1] & A_2 &= [0.25, 0.25, 0.3, 0.2] \\ A_3 &= [0.65, 0.15, 0.1, 0.1] & A_4 &= [0.5, 0.1, 0.2, 0.2] \\ A_5 &= [0.15, 0.3, 0.3, 0.25] & A_6 &= [0.15, 0.15, 0.35, 0.35] \\ A_7 &= [0.3, 0.3, 0.2, 0.2] & A_8 &= [0.35, 0.4, 0.25] \\ A_9 &= [0.2, 0.4, 0.4] & A_{10} &= [0.35, 0.35, 0.3] \end{aligned}$$

选择平均综合多因素的模糊算子 $M(\bullet, +)$ ，根据式(1)计算二级因素评估结果，将其作为行向量构造对应一级因素模糊评估矩阵并归一化。

$$\begin{aligned} R_r &= \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = \begin{bmatrix} 0.9715, 0.018, 0.009, 0.0015, 0 \\ 0.5635, 0.058, 0.0875, 0.097, 0.194 \\ 0.0875, 0.0655, 0.179, 0.099, 0.569 \end{bmatrix} \\ R_l &= \begin{pmatrix} B_4 \\ B_5 \\ B_6 \\ B_7 \end{pmatrix} = \begin{bmatrix} 0.5, 0, 0.1, 0.4, 0 \\ 0, 0.15, 0.3, 0.55, 0 \\ 0, 0, 0.5, 0.5 \\ 0, 0.5, 0.5, 0, 0 \end{bmatrix} \\ R_p &= \begin{pmatrix} B_8 \\ B_9 \\ B_{10} \end{pmatrix} = \begin{bmatrix} 0.077, 0.1035, 0.103, 0.178, 0.5385 \\ 0.472, 0.128, 0.146, 0.074, 0.18 \\ 0.715, 0.1165, 0.1245, 0.026, 0.018 \end{bmatrix} \end{aligned}$$

再依据数据隐私差异由二元决策法得到其对应的权值向量。

$$\begin{aligned} A_r &= [0.1, 0.3, 0.6] \\ A_l &= [0.3, 0.15, 0.4, 0.15] \\ A_p &= [0.1, 0.7, 0.2] \end{aligned}$$

并根据式(3)计算出一级因素评估结果，构造 d_1 的模糊评估矩阵。

$$R_{d_1} = \begin{pmatrix} B_r \\ B_l \\ B_p \end{pmatrix} = \begin{bmatrix} 0.3187, 0.0585, 0.1346, 0.0887, 0.3995 \\ 0.1500, 0.0975, 0.1500, 0.4025, 0.2000 \\ 0.4811, 0.1233, 0.1374, 0.0748, 0.1834 \end{bmatrix}$$

确定隐私保护效果评估权值向量为 $A_d = [0.35, 0.3, 0.35]$ ，根据式(5)计算 d_1 的评估结果。

$$\begin{aligned} B_{d_1} &= A_d \odot R_{d_1} = [0.35, 0.3, 0.35] \odot \\ &= \begin{bmatrix} 0.3187, 0.0585, 0.1346, 0.0887, 0.3995 \\ 0.1500, 0.0975, 0.1500, 0.4025, 0.2000 \\ 0.4811, 0.1233, 0.1374, 0.0748, 0.1834 \end{bmatrix} = \\ &= [0.3249, 0.0928, 0.1402, 0.1779, 0.2640] \end{aligned}$$

根据最大隶属度原则可以将 d_1 的隐私保护

效果评定为优，由评估过程可以发现影响其隐私保护效果的不利因素主要是各类攻击敌手以及数据中存在较多的偏好信息。

Step5 构建待评估数据 D 的模糊评估矩阵。

$$R_D = \begin{pmatrix} B_{d_1} \\ B_{d_2} \\ \vdots \\ B_{d_{10}} \end{pmatrix}$$

根据文献[9]所提的隐私差异度量方法，将包括 15 种疾病与对应治疗结果作为隐私详情维度下的指标对 D 中元素对应的等级，由此得到 D 的评估权值向量 A_D 。根据式(7)计算得到最终评估结果 B_D ，计算过程省略。

$$\begin{aligned} B_D &= A_D \odot R_D = \\ &= [0.3125, 0.3341, 0.0983, 0.2104, 0.0447] \end{aligned}$$

根据随机抽取的 100 条记录可以认为 MIMIC-III 数据集隐私保护效果为良，这与 MIMIC-III 数据集通过模式识别算法移除个人健康受保护信息的实际情况相符。进一步，可以继续抽取数据集中的记录进行评估，亦可对其他医疗数据集进行评估，通过所提评估结果扩展计算方法得到更全面的临床医疗数据隐私保护效果的评估结果。

5 结束语

本文针对网络环境隐私保护效果评估问题，借助模糊综合评估方法，将网络环境隐私保护效果评估由单一数据层面转向数据全生命周期，从隐私风险来源、隐私类别和隐私保护实施 3 个方面建立评估指标，通过划分最小评估单元，分区块、多层次地进行基于模糊综合评估的隐私保护效果评估，由数据隐私差异确定各因素、各区块权重，整合计算得到网络环境中隐私保护效果评估结果。实例分析表明，所提方法简便可行，在明确网络环境整体隐私保护效果的同时，可得到影响隐私保护效果的主要原因，为优化下一步隐私保护具体实施提供理论依据。

参考文献:

- [1] CHANG V, SUN G, LI J. Guest editorial: security and privacy for

- multimedia in the internet of things (IoT)[J]. Multimedia Tools & Applications, 2018:1-2.
- [2] SWEENEY L. K-anonymity: a model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(5):557-570.
- [3] DWORK C, MCSHERRY F, NISSIM K. Calibrating noise to sensitivity in private data analysis[J]. Proceedings of the VLDB Endowment, 2006, 7(8):637-648.
- [4] LU Q W, WANG C M, XIONG Y. Personalized privacy-preserving trajectory data publishing[J]. Chinese Journal of Electronics, 2017, 26(2): 285-291.
- [5] FUNG B C M, WANG K, YU P S. Top-down specialization for information and privacy preservation[C]//International Conference on Data Engineering. 2005: 205-216.
- [6] JIANG H W, ZENG G S, MA H Y. Greedy clustering-anonymity method for privacy preservation of table data-publishing[J]. Journal of Software, 2017,28(2):341-351.
- [7] XU J, WANG W, PEI J, et al. Utility-based anonymization using local recoding[C]//Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2006.
- [8] 李风华, 李晖, 贾焰, 等. 隐私计算研究范畴及发展趋势[J]. 通信学报, 2016, 37(4): 4-14.
LI F H, LI H, JIA Y, et al. Privacy computing: concept, connotation and its research trend [J]. Journal on Communications, 2016, 37(4): 4-14.
- [9] 俞艺涵, 付钰, 吴晓平. 基于 Shannon 信息熵与 BP 神经网络的隐私数据度量与分级模型[J]. 通信学报, 2018, 39(12): 10-17.
YU Y H, FU Y, WU X P. A metric and classification model for privacy data based on Shannon information entropy and BP neural network[J]. Journal on Communications, 2018, 39(12): 10-17.
- [10] JOHNSON A E W, POLLARD T J, LU S, et al. MIMIC-III, a freely accessible critical care database[J]. Scientific Data, 2016, 3: 160035.

[作者简介]



俞艺涵（1992- ），男，浙江金华人，海军工程大学博士生，主要研究方向为信息安全、模拟仿真、隐私保护。



付钰（1982- ），女，湖北武汉人，博士，海军工程大学副教授，主要研究方向为信息安全风险评估。



吴晓平（1961- ），男，山西新绛人，博士，海军工程大学教授、博士生导师，主要研究方向为系统分析与决策。