计算机科学
COMPUTER SCIENCE

边缘计算下差分隐私的应用研究综述

孙剑明, 赵梦鑫

引用本文

孙剑明, 赵梦鑫. 边缘计算下差分隐私的应用研究综述[J]. 计算机科学, 2024, 51(6A): 230700089-9.

SUN Jianming, ZHAO Mengxin. Survey of Application of Differential Privacy in Edge Computing[J]. Computer Science, 2024, 51(6A): 230700089-9.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.230700089

边缘计算下差分隐私的应用研究综述

孙剑明　赵梦鑫
哈尔滨商业大学计算机与信息工程学院　哈尔滨 150028
(sjm@hrbcu.edu.cn)

摘　要　为了解决传统云计算模式的延迟和带宽限制,应对物联网和大数据时代的需求,边缘计算开始崭露头角并逐渐受到广泛关注。在边缘计算环境下,用户数据的隐私问题成为了一个重要的研究热点。差分隐私技术有着坚实的数学基础,它作为一种有效的隐私保护算法,已经被广泛应用于边缘计算中,两者的结合有效缓解了隐私保护低和计算成本高的问题。首先介绍了互联网发展带来的问题,其次介绍了边缘计算的基本概念、特点和组成部分,并概括了与传统云计算相比的优势,然后对差分隐私的基本概念和原理进行了概括,进而详细阐述了差分隐私的3种扰动方式和常用的实现机制,最后对边缘计算下差分隐私的应用研究进行了综述,并指出了未来的研究方向。总之,将差分隐私技术应用于边缘计算场景对隐私保护和数据分享都是一种有效保护手段。

关键词:边缘计算;差分隐私;本地化差分隐私;隐私保护;实时数据处理
中图分类号　TP309.2

Survey of Application of Differential Privacy in Edge Computing

SUN Jianming and ZHAO Mengxin
School of Computer and Information Engineering,Harbin University of Commerce,150028,China

Abstract　In order to address the latency and bandwidth limitations of the traditional cloud computing model and to cope with the demands of the Internet of Things and the big data era,edge computing is making its appearance and gaining widespread attention. In the edge computing environment,the privacy of user data has become an important research hotspot. The combination of differential privacy techniques,which have a solid mathematical foundation,has been widely used in edge computing as an effective privacy-preserving algorithm to improve the problem of low privacy protection and high computational cost. The problems brought about by the development of the Internet are firstly introduced,followed by the basic concepts,features and components of edge computing,and the advantages compared with traditional cloud computing are outlined. The basic concepts and principles of differential privacy are again outlined,followed by a detailed description of the three perturbation methods and common implementation mechanisms of differential privacy,and finally the research on the application of differential privacy under edge computing is reviewed. Finally,the research on the application of differential privacy under edge computing is reviewed and future research directions are pointed out. In conclusion,the application of differential privacy techniques to edge computing scenarios is an effective means to protect privacy and data sharing.

Keywords　Edge computing,Differential privacy,Local differential privacy,Privacy preserving,Real-time data processing

1　引言

5G时代网络数据纵横交错,大数据技术带来了更多的信息资源并提高了我们的数据分析能力,以便能够更高效地获取和管理信息,但与此同时,数据的存储、传输和处理速度等也成为了一种挑战和困扰。传统计算方式存在许多限制,如传输延迟、带宽瓶颈、安全问题等,导致云计算模式无法满足实时处理的需求,边缘计算应运而生[1]。边缘计算是一种新兴的分布式计算范式,有着“人工智能的最后一公里”的美称。其主要将计算资源从传统的集中式云端数据中心移至网络边缘,以支持分布式计算和实时处理,且具有部署成本低、效率高、更加节能和智能化等优势。

在当前信息化社会中,数据已成为一种重要的生产要素和基础资源,互联网时代,海量数据迅速传播,人们在快速获取信息的同时其隐私数据也面临着极大风险,而当前人们越来越注重自己的个人隐私,这就要求我们顺应时代的发展和变迁,在信息获取和隐私安全之间找到一种折中的方法以达到两者之间的平衡。而在边缘计算环境下,用户数据的隐私问题也成为了一个重要的研究热点[2],Yahuza等[3]对边缘计算的安全和隐私需求做了一个实质性回顾,研究了面对各种攻击时采用的技术方法并做了一个总结,为研究人员提出新的想法奠定了基础。但传统的隐私保护方案无法满足边缘计算中实时处理的需求,而差分隐私因具有严格定义的数学框架,隐私数据能够得到有效的保护,进而逐渐受到了广泛关注。

基金项目:国家自然科学基金(32201411)
This work was supported by the National Natural Science Foundation of China(32201411).
通信作者:赵梦鑫(1791295445@qq.com)

2 边缘计算

边缘计算[4-5](Edge Computing)的概念最早是由思科(Cisco)提出,目的是解决云计算模式在处理实时应用和大量数据时的延迟和网络拥塞问题。边缘计算指将计算和存储资源放置在接近数据来源的距离更短的设备或者节点上,以便优化网络连接性能、减少数据传输时延、提高可靠性,实现实时数据处理和分布式决策的一种计算模式和架构。通俗地说,边缘计算就是将计算任务从云端转移到离数据源更近的边缘节点。通过图 1 能够直观地了解边缘计算的应用场景和运作模式。

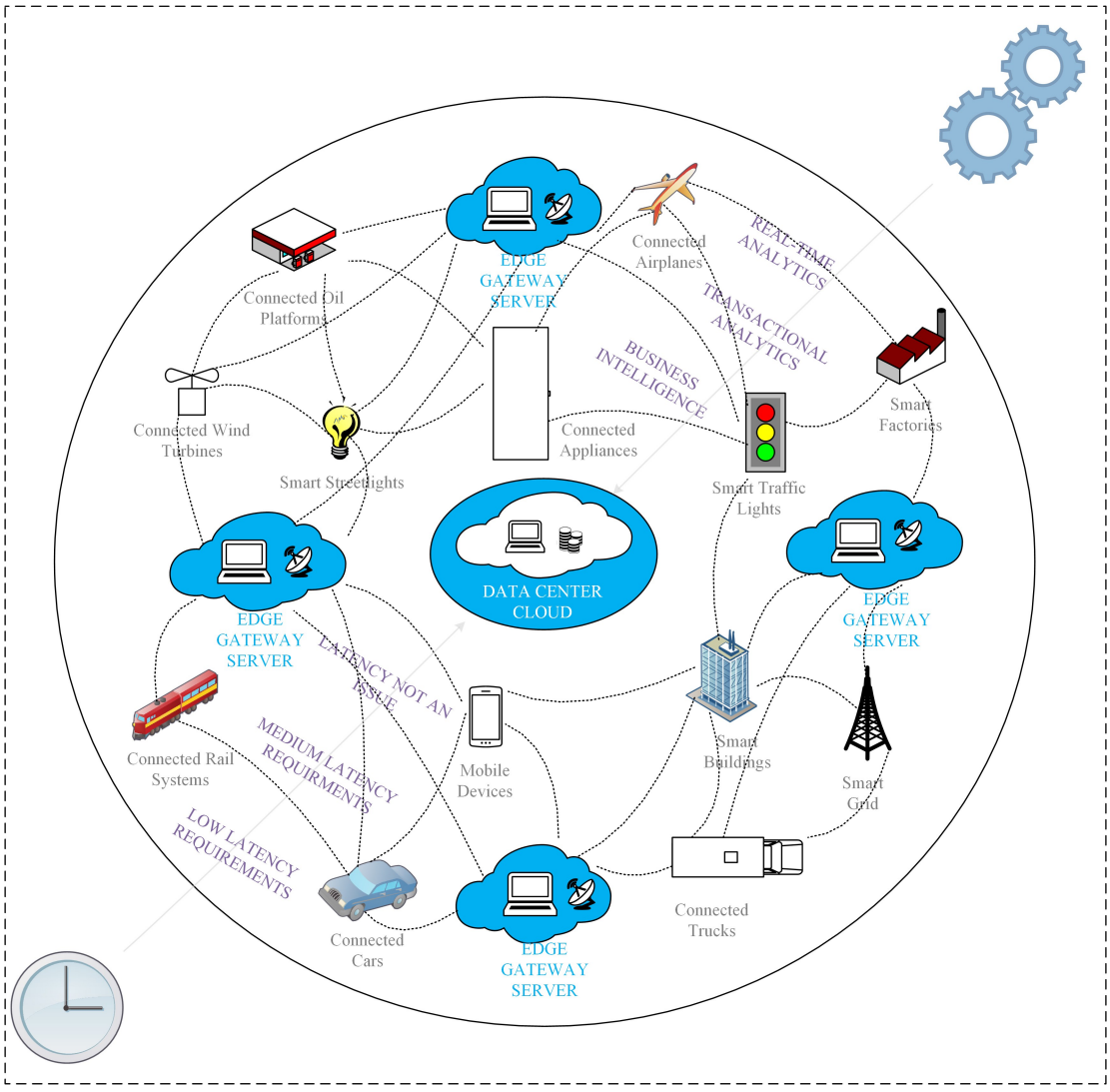

图 1 边缘计算

Fig. 1 Edge computing

边缘计算应用体现在方方面面,如我们生活中随处可见的智能摄像头;在交通系统中实时处理交通数据,提供实时交通流量监测、路况预测等;在医疗设备系统中,实时监测患者生命体征;在零售业系统中,实时进行商品库存管理等。其特点主要包括:

1)具有低延迟。由于数据不必传输到中央服务器进行处理,而是在边缘计算设备上处理,因此能够实现低延迟的数据传输和处理。

2)网络带宽优化。边缘计算可以在本地对数据进行处理和筛选,只将结果和摘要传输到云端,减少了对网络带宽的需求,降低了数据传输时的延迟和网络阻塞率。

3)分布式架构。边缘计算采用分布式架构,将计算任务分散到多个边缘节点上。这种分布式的特点使得边缘计算具有高可扩展性和容错性,可以应对大规模的计算需求和故障情况。

4)数据本地化。边缘计算将计算和存储资源放置在边缘节点,使得数据的处理和存储发生在离数据源或用户更近的位置,避免将所有数据传输到云端,减少了数据传输量和存储成本。

5)弹性计算。边缘计算可以根据实际需求进行动态的资源调配和管理。根据不同的应用场景和负载情况,边缘节点可以灵活地分配计算和存储资源,提供弹性计算能力。

6)离线支持。边缘计算可以在有限或断开的网络连接下工作,支持离线应用和有限带宽环境下的计算需求。

它将计算资源从云端移动到网络边缘,以支持实时处理和分布式计算。常见的边缘计算场景包括智能家居、车联网、位置服务、工业物联网等。通过边缘计算,可以实现对大量终端设备产生的数据进行处理分析,从而使数据量和处理效率得到明显的提升。

与云计算和传统中心化计算相比,边缘计算具有以下优势:

1)低延迟。由于边缘计算设备直接处理数据,不需要将

数据传输到远程服务器上进行处理，因此能够实现更低的延迟，为实时应用提供支持。

2）数据本地性。边缘计算设备通常部署在距离数据源最近的位置，可以快速获取和处理数据，降低数据传输带来的网络瓶颈和安全风险。

3）网络拥塞减轻。边缘计算可以利用薄边界计算机网关设备和云与边缘之间的半连接，使得数据只需流入中心云进行大量冗余计算，从而避免了一些高负载、网络拥塞影响或者成本等问题。

4）稳定性。边缘计算架构下的节点分布在不同地理区域，实现了去中心化管理，可以通过插拔方式访问分级网络，提高了系统的稳定性和故障容错能力。

5）安全性。边缘计算能够保护数据和隐私，以此避免面向公共云、借助云提供商等传统方式的数据泄露风险。

综上所述，边缘计算相比传统的计算架构有明显的优势，因此，在物联网、车联网、视频监控等领域得到了广泛应用，并逐渐成为了推动物联网技术发展的中坚力量。图 2 给出了两种模型的对比图，从图中能够更直观地看出两种模型不同的运作模式。

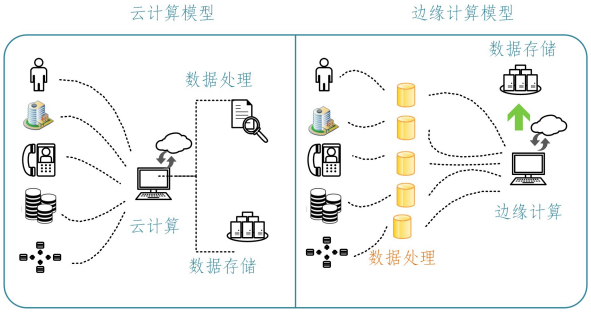

图 2　云计算 vs. 边缘计算

Fig. 2　Cloud computing vs. edge computing

边缘计算的组成主要包括以下几个部分：边缘设备、边缘网络、边缘节点、边缘应用和云端。边缘设备指物理世界即现实世界中的终端设备，如传感器、智能手机等，它通常位于数据源附近，将数据传输到边缘节点以进行进一步处理；边缘节点是位于边缘设备和云端之间的计算和存储节点，通常由路由器、交换机、基站或物联网设备组成；边缘网络是连接边缘设备和边缘节点的网络基础设施，如无线网络（Wi-Fi、蓝牙等）和有线网络（以太网、光纤等），它提供了边缘设备与边缘节点之间的通信和数据传输能力；边缘应用是部署在边缘节点上的应用程序，用于处理和分析边缘设备生成的数据；云端指位于云数据中心的计算和存储资源。随着大数据时代的到来，边缘计算因其能够快速响应用户请求，提供高效能和优质的服务，在未来的发展中有着非常广泛的应用前景。然而，在此背景下，隐私问题也成为了边缘计算领域需要关注和解决的重要问题之一。

3　差分隐私

差分隐私（Differential Privacy）是隶属于密码学的一种保护隐私的数据处理技术。其主要是通过在数据处理过程中添加噪声或引入一定程度的随机化扰动，混淆原始数据从而隐藏个人特征和敏感信息。简单来说，就是一种基于数据失真的隐私保护技术。为了更直观地理解差分隐私的运作过程，给出了差分隐私的基本原理图，如图 3 所示。

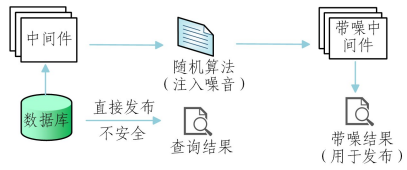

图 3　差分隐私基本原理

Fig. 3　Fundamentals of differential privacy

差分隐私提供了一种数学框架，用于量化在数据发布或查询过程中个体数据对整体结果的影响程度。这种技术被广泛应用于数据共享、隐私保护和数据分析等场景，而在边缘计算环境下，差分隐私同样可以起到重要的保护作用。

3.1　差分隐私的相关定义

在用数学形式对差分隐私[6-9]进行定义前，首先介绍随机算法的定义。

定义 1（随机算法）　给定一个具有域 A 和离散范围 B 的随机算法 M 与映射 $M:A\to\Delta(B)$ 相关联。在输入 $a\in A$ 时，对于每个 $b\in B$，算法 M 以概率 $(M(a))_b$ 输出 $M(a)=b$，概率空间就是算法 M 的抛硬币次数。

其中 $\Delta(B)$ 为离散集合 B 的概率单纯形（Probability Simplex），表示为：

$$\Delta(B)=\{x\in\mathbb{R}^{|B|}:x_i\geqslant 0 \text{ for all } i \text{ and } \sum_{i=1}^{|B|}x_i=1\}。$$

定义 2（差分隐私，DP）　如果对于所有 $S\subseteq Range(M)$ 以及所有 $D,D'\in\mathbb{N}^{\chi}$，且 $\|D-D'\|_1\leqslant 1$（D 和 D' 为临近数据集），有 $\Pr[M(D)\in S]\leqslant e^{\varepsilon\times|D\oplus D'|}\times\Pr[M(D')\in S]+\delta$，则域 $\mathbb{N}^{|\chi|}$ 的随机算法 M 满足 (ε,δ)-差分隐私（松弛差分隐私），若 $\delta=0$，则称算法 M 提供 ε-差分隐私保护（严格差分隐私）。

其中 ε 参数被称为隐私保护预算，通过定义该参数量化隐私损失的允许范围，该参数也决定了在查询结果中引入的噪声的强度。较小的 ε 值表示更严格的隐私保护，但与此同时数据的可用性会降低。在具体的使用场景下，我们需要找到一个平衡点来权衡两者之间的关系，在不损害隐私保护的前提下保障数据的可用性和分析结果的准确性。通常情况下 $|D\oplus D'|=|D\cup D'|-|D\cap D'|=1$；$\Pr[.]$ 为概率函数。差分隐私要求查询结果的概率分布在不同邻近数据集上只有微小的差异，不会因为个体数据的微小变化而产生明显的差别。差分隐私的数据处理框架如图 4 所示。

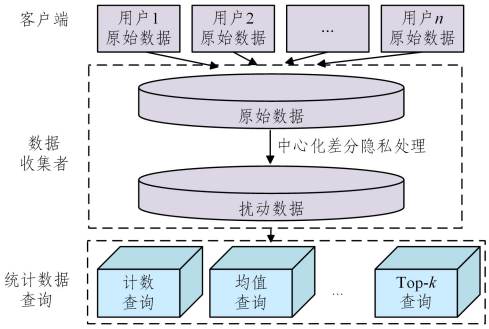

图 4　差分隐私处理框架

Fig. 4　Processing framework of differential privacy

定义 3（本地化差分隐私（Local Differential Privacy，LDP））　本地化差分隐私[10]被认为是 DP 的一种变体，且不需要预设一个可信的第三方[8,11]。它强调在个体设备上进行

数据隐私保护，而不是在集中式的数据收集和处理环境中进行。给定一个随机算法 M，当且仅当在任意两条记录 V 和 V' 上得到的输出结果都为 v^* 的可能性满足 $\Pr[M(V)=v^*]\leqslant e^{\varepsilon}\times\Pr[M(V')=v^*]+\delta$，则 M 满足($\varepsilon$,$\delta$)-本地化差分隐私(松弛本地化差分隐私)，若 $\delta=0$，则满足 ε-本地化差分隐私保护(严格本地化差分隐私)。本地化差分隐私的数据处理框架如图 5 所示。

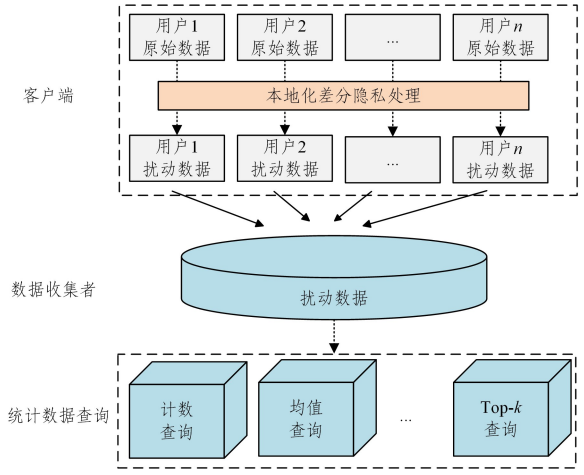

图 5 本地化差分隐私处理框架

Fig. 5 Processing framework of local differential privacy

定义 4(全局敏感度或 ι_1-敏感度) 给定一个查询或计算函数 $f:D\rightarrow\mathbb{R}^d$，对于任意的邻近数据集 D 和 D'，有

$$\Delta f=\max_{D,D'}\| f(D)-f(D')\|_1$$

其中，$\| f(D)-f(D')\|_1$ 为 $f(D)$ 和 $f(D')$ 之间的 1-阶范数距离(曼哈顿距离)。

定义 5(局部敏感度) 设有一个查询或计算函数 $f:D\rightarrow\mathbb{R}^d$，对于给定数据集 D 和它的任意邻近数据集 D'，$\Delta f_{LS}=\| f(D)-f(D')\|_1$ 称为函数 f 在 D 上的局部敏感度。

3.2 差分隐私的扰动方式[12]

差分隐私通过向原始数据添加噪声或扰动混淆原始数据，使侵扰者无法从聚合结果中准确推断出任何关于个体数据的敏感信息。根据对数据集注入噪声的不同，差分隐私的扰动方式可以分为以下 3 种。

3.2.1 输入输出扰动(Input-Output Perturbation)

输入输出扰动技术是一种基于数据独立性的随机化技术，它试图通过扰动原始数据来提高隐私保护水平，即对输入数据或输出结果直接加噪。在输入输出扰动的情况下，每个请求会被扰动生成一个与查询结果相似但不完全准确的响应值。

该方法的优点是易于实现且提供了很好的结果保证；缺点是会导致较大的误差和不确定性。

3.2.2 目标函数扰动(Object Function Perturbation)

目标函数扰动技术是一种基于最优化问题的随机化技术，是针对优化问题中的目标函数进行随机扰动，从而影响最优解的偏移量。通过向目标函数添加噪声来实现模糊化数据，从而保护敏感信息的泄露。目标函数扰动方法过程如下：

1)对于目标函数 $f(x)$ 和输入变量 x，对 x 加噪得到 x'，即 $x'=x+noise$；

2)基于 x' 计算扰动后的目标函数 $f'(x')$，即 $f'(x')=f(x')+noise$；

3)根据实现差分隐私的机制参数 d 计算噪声大小 $\varepsilon=\frac{d}{2\gamma}$，$\gamma$ 为敏感度，从而可以控制噪声大小来保持数据的可用性和隐私性的平衡；

4)将生成的 $f'(x')$ 作为扰动后的目标函数，在求解过程中应用基于差分隐私的优化算法，以便在隐私保护的同时获得最佳的优化结果。

该方法的优点是提高了隐私保护水平，增强了模型可用性，在满足敏感数据隐私保护的同时能够减小误差；缺点是对目标函数进行扰动可能使模型无法收敛到最优，从而影响模型性能。

3.2.3 计算过程扰动(Computation Process Perturbation)

过程扰动技术主要通过对计算过程中的中间结果或执行步骤进行扰动处理，在不泄露原始敏感信息的前提下提供计算结果。在计算过程中应用扰动的方式通常包括以下几个步骤：

1)选择要执行的计算任务；

2)对于每个任务，设计一个相应的计算过程，并将其分解为多个步骤；

3)在计算过程的每个步骤中，根据预定的噪声机制向计算结果添加噪声，以保护数据隐私；

4)将所有经过扰动的中间结果传递到最终计算步骤，生成包含噪声的计算结果；

5)根据需要对结果进行调整，以消除与噪声相关的误差，并尽可能接近原始数据的真实计算结果。

在实际使用中，计算过程扰动可以应用于多种计算任务，在有限的计算资源和数据安全性之间寻找平衡点。然而，由于计算过程扰动会导致计算精度的损失，因此在实际情况中需要权衡数据隐私与计算正确性之间的关系。该方法的优点是数据的安全性更高；缺点是可能会对算法效果产生影响。

以上 3 种扰动方式各有其适用场景和优缺点。应根据实际应用需求，选择不同的差分隐私扰动技术以同时保护数据隐私及查询结果的准确性，提升数据共享和挖掘相关性研究的安全性和效率。

3.3 差分隐私的实现机制

3.3.1 拉普拉斯机制(Laplace Mechanism)

定义 6(拉普拉斯分布) 尺度为 b 的拉普拉斯分布(以 0 为中心)的概率密度函数为：

$$\mathrm{Lap}(x|b)=\frac{1}{2b}\mathrm{e}^{-\frac{|x-\mu|}{b}}$$

其中，μ 为分布的均值(位置参数)。

定义 7(拉普拉斯机制) 假设有一个查询或计算函数 $f:D\rightarrow\mathbb{R}^d$，以数据集 D 为参数，输出查询结果为 $F(D)$，那么随机算法 $M(D)=F(D)+Y$ 提供 ε-隐私保护，其中 $Y\sim\mathrm{Lap}\left(\frac{\Delta f}{\varepsilon}\right)$ 是从拉普拉斯分布中抽样得到的噪声值；拉普拉斯噪声的尺度参数 b 通过计算敏感度 Δf 和隐私参数 ε 得到，即 $b=\frac{\Delta f}{\varepsilon}$，通常用于数值型数据。

3.3.2 高斯机制(Gaussian Mechanism)

定义 8(ι_2-敏感度) 设 $f:D\rightarrow\mathbb{R}^d$ 为任意 d 维函数，对于任意的邻近数据集 D 和 D'，有：

$$\Delta_2 f=\max_{D,D'}\| f(D)-f(D')\|_2$$

其中，$\|\cdot\|_2$ 代表欧氏距离。

定义 9（高斯分布或正态分布） 随机变量 x 服从数学期望为 μ、方差为 σ^2 的高斯分布，具有概率密度函数：

$$P_{\mu,\sigma^2}(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}。$$

定义 10（高斯机制） 设 $\varepsilon\in(0,1)$ 是任意的，对于 $c^2\geqslant 2\ln(1.25/\delta)$，$\sigma\geqslant c(\Delta_2 f)/\varepsilon$ 的高斯机制为 (ε,δ)-差分私有。与另外 3 种机制不同，高斯机制只在松弛差分隐私条件下成立，即 (ε,δ)-差分隐私，多用于连续型查询。

3.3.3 指数机制（Exponential Mechanism）

定义 11（指数机制） 给定数据集 D、查询结果 $\gamma\in R$、查询隐私参数 ε 及可用性函数 $q(D,\gamma)$，当且仅当 $M(D,q)\propto e^{\frac{\varepsilon q(D,\gamma)}{2\Delta q}}$ 时，我们认为随机算法 M 是满足 ε-差分隐私保护的，$\propto$ 表示正比于。但 $e^{\frac{\varepsilon q(D,\gamma)}{2\Delta q}}$ 表示的不是一个概率值，这里需要进行归一化得到相应的概率值，即

$$\Pr[M(D)=\gamma]=\frac{e^{\frac{\varepsilon q(D,\gamma)}{2\Delta q}}}{\sum_{\gamma'\in R}e^{\frac{\varepsilon q(D,\gamma')}{2\Delta q}}}$$

其中，$\Delta q=\max_{\gamma\in R}\max_{D,D'}|q(D,\gamma)-q(D',\gamma)|$ 为敏感度。

其原理是在查询结果中引入适当的噪声，使数据可用性和隐私保护均衡发展。与拉普拉斯机制和高斯机制类似，它基于指数分布，将隐私泄露的概率与查询结果的质量和差异程度相关联，多用于非数值型数据。

3.3.4 随机响应机制（Randomized Response Mechanism）

定义 12（随机响应机制） Dwork 和 Roth 在 Warner 的基础上提出了随机响应的变体，主要用于保护敏感的二元数据，多用于本地化差分隐私。假设有一个活动 A，当被问到本周是否进行过 A 活动，我们执行下述步骤：

1）抛一枚硬币；

2）如果硬币是反面，则如实回答；

3）如果是正面，则抛第二枚硬币，第二枚为正则回答是，为反则回答否。

这种随机化让真实答案产生了不确定性，提供了一种“合理的推诿”，经证明该机制满足 $\varepsilon=\ln 3$ 的 ε-差分隐私。

4 相关工作

目前，差分隐私在边缘计算领域的应用研究主要集中在以下几个方面：基于位置的服务（Location-Based Service，LBS）、隐私保护的数据聚合、个性化推荐系统、车联网、智慧医疗等。在这些研究中，差分隐私主要为边缘计算环境下的隐私保护提供了一种有效的解决方案。

4.1 基于位置的服务

基于位置的服务是一种利用用户的地理位置信息来提供个性化服务的技术。在基于位置的服务[13-43]中，如定位应用或地图应用等，用户的位置信息非常重要，但同时也具有很高的隐私风险。因此，在边缘计算环境下采用差分隐私技术来保护用户的位置隐私是非常必要的。一种常见的做法是，将预先设定的区域划分为多个小区间，并对这些小区间计算出其内部点的平均位置作为所在区间的代表点。然后，通过向这些代表点加入适当的随机扰动量来保证差分隐私。这样，用户在提供位置信息时，可将位置与相邻的区间代表点进行比较，在提供一定程度的位置相关性的同时又可以减少位置信息的细节，从而达到保护用户位置隐私的目的。

Yu 等[15]采用 Trie 数据结构对位置数据进行存储，对选取的频繁位置的支持度进行加噪扰动，然后进后置处理以提高数据可用性，最后通过仿真实验对比了有无差分隐私敏感位置的保护性和可用性，以及该方案采用的一致性约束后置处理与其他两种后置处理方式的数据可用性，证明了该方案极大的改善了共享位置数据可用性低和隐私保护力度弱的问题。Zhang 等[16]对基于 LBS 系统的位置轨迹隐私保护技术做了一个综述，分别介绍了 LBS 现有的基于扭曲、加密、匿名和差分隐私 4 种隐私保护技术，进而分析比较 4 种技术的优缺点，结果表明基于差分隐私的保护技术具有更广阔的应用前景，且该技术有坚实的数学理论作支撑，是当前 LBS 隐私保护的主流技术。

4.1.1 室内定位

室内环境中，由于信号穿透性不佳、易受建筑物遮挡等因素的影响，全球定位系统 GPS 等传统的定位技术提供的定位结果往往精度较低。在这种情况下，基于接收信号强度（Received Signal Strength，RSS）的定位方法因其较高的灵活性和扩展性、可靠性、成本效益、适用于现有无线设备等诸多优势，已经发展成为室内定位技术的主流趋势[16]，其基本思想是根据接收到的信号强度与事先建立的信号强度与位置之间的映射关系，来估计移动设备的位置，以此实现室内用户定位的目的。

Zhang 等[17]针对基于云架构的集中式学习框架需要把大量的 RSS 指纹数据上传到云服务器而耗费大量算力和隐私易泄露的问题，设计了 DP-FlocEC 算法模型，通过将计算任务从中央服务器转移到边缘设备上处理，提高系统的实时性和响应速度；采用差分私有联邦学习模型进行训练，保护用户隐私；对比 4 种基于云架构的集中式模型，在有效的隐私保护前提下定位精度只有很小的损失，且减小了通信成本；DP-FlocEC 基于联邦学习架构，与另外 3 种基于联邦学习的分布式架构模型相比，在时间性能、定位精度和资源通信开销基本相同的前提下，对用户隐私数据的保护更加全面，实验结果证明了该模型的优越性。He[18]首次提出了一种边缘计算下支持动态隐私预算分配的室内定位联邦学习方法 ADP-FLocEC，采用皮尔逊系数对 RSS 数据进行收集和预处理，采用瑞丽差分隐私动态追踪模型训练过程中的隐私损失，有效改变了边缘计算环境下差分隐私预算分配难适用的问题，ADP-FLocEC 模型不仅保证了用户数据和模型参数的隐私保护，而且获得了较好的室内定位精度和较低的处理时延。

4.1.2 兴趣点查询

在基于位置的服务中，兴趣点（Points of Interest，POI）查询涉及用户位置信息的收集和处理，并可能产生与用户行为相关的个人敏感信息，因此需要采取一些差分隐私技术来保护用户隐私[19]。

Zhang 等[20]针对不同上下文添加统一随机噪声导致的处理效率低、LBS 可用性下降的问题，提出了 DP3-SLOC 算法模型。该模型引入了地理不可区分性，使得构建的语义位置信息不会受用户所在地区的干扰，结合语义位置信息和 d_x-隐私来计算其敏感度，进而根据敏感度为不同兴趣点添加

相应的噪声，使开销、服务质量和隐私保护水平三者得到了均衡发展。收集用户的签到信息可以丰富推荐系统的数据来源，以此为用户推荐新的兴趣点，但收集用户签到记录容易造成隐私泄露、信息滥用和偏见与歧视（局限了用户的兴趣范围）的问题。Jong 等[21]针对隐私保护推荐系统只考虑用户 POI 间的关系而不考虑签到历史中人类的行为动向的问题，设计了一个具有本地差分隐私的连续兴趣点推荐框架 SPIREL。SPIREL 联合学习用户-POI 和 POI-POI 的关系，通过在 4 个数据集上的实验证明，SPIREL 与之前的隐私保护推荐系统相比具有更高的准确性。

4.2 车联网

车联网（Internet of Vehicles，IoV）是互联网络的新形态，是互联网与物联网的融合，也被认为是车载自组织网络和物联网的一个特殊类别[22]。它包括车载通信、车辆感知、网络通信、数据处理和应用服务等技术的综合应用，极大地促进了智能交通领域的发展。然而，车辆数据隐私问题被认为是车联网应用和发展的主要障碍，引起了人们的极大关注[23]，Zhao 等对差分隐私在车联网中的应用做了一个调查，并对现有的应用技术做了一个全面的分析与总结[24]。

车联网的发展需要更大的存储空间和更低信息处理时延，传统的云计算很难满足这一要求，而边缘计算被认为是一种更适合于提高车辆环境计算能力的解决方案[25]。边缘计算与差分隐私技术的结合可以在车联网中更好地保护用户的隐私数据[26]，在保护数据隐私和可用性的同时提高计算效率、减少数据传输并赋予用户更多的数据所有权和控制权。差分隐私技术的应用可以保护用户的个人信息和数据安全[27]，提高车联网的可靠性和服务质量，同时促进其健康发展。

Xie 等[28]针对车联网中虚假轨迹信息可用性低的问题，结合位置语义提出了一种基于位置语义的差分隐私轨迹保护方法 DPSL，文中采用 K-均值对具有相同语义的虚假位置进行聚类，采用 Hausdorff 距离衡量轨迹相似度，经实验结果证明，在保证用户轨迹隐私的前提下，生成的虚假轨迹信息可用性也得到了明显提升。Xu 等[29]着重考虑用户的个性化隐私偏好需求，根据用户的偏好和隐私需求进行定制化的隐私预算分配；引入多属性决策理论为用户选择效用最高的路线，利用 ε-地理不可区分性生成用户可接受的虚假位置范围信息传送给服务器，在不暴露隐私信息的情况下尽力贴合真实的服务请求位置。通过真实数据集上的仿真实验以及和其他方案的对比，该方案在满足用户个性化隐私保护的同时服务质量也有明显的提升。Wu 等[26]针对攻击者通过训练反卷积网络的方式发动图像还原攻击的问题，提出了 3 种面向车路协同推断的差分隐私防御算法，且推断精确度都达到了 90% 以上。

4.3 智慧医疗

智慧医疗（Smart Healthcare）[30]是利用信息技术、物联网技术、大数据和人工智能等新兴技术进行医疗保健的一种创新模式。它旨在通过建立数字化医疗服务平台、优化医疗流程、提高医疗质量、促进医疗创新，实现医疗数据和资源的生态共享。

在医疗健康领域，患者健康数据的收集和共享是一个非常敏感和涉及隐私的问题。差分隐私可以帮助医疗机构在保护患者隐私的前提下对数据进行分析和利用[31]。在未来的智能医疗保健应用程序中，期望设备能够做出决策并对任务做出相应的反应，以减少实时应用程序的延迟。通过实现远程监控、个性化治疗和更快、更准确的诊断，边缘计算有可能彻底改变医疗保健的提供方式[32]。同时，将数据处理和分析工作移到边缘端设备上，不仅可以降低网络负载、提高数据隐私和安全性，还可以提高实时性、节约计算成本。

Li[33]提出了一个基于边缘计算的安全医疗诊断系统，采用基于边缘计算的联邦学习框架对本地数据进行保护，采用差分隐私梯度下降算法对训练数据进行保护，采用环签名技术对医疗诊断机器学习模型进行保护，最后对 Kaggle 库中标注处理后的原始淋巴切片细胞图片进行实验验证，准确率达到了 91.90%。Liu 等[34]提出了一种基于区块链的本地差分隐私上下文在线学习模型，用于移动边缘计算中对冠心病的诊断。该模型基于情境感知的冠心病在线学习诊断算法，能够对患者进行实时个性化诊断；基于自适应扩展树结构，保证了诊断推荐结果的准确性；采用本地差分隐私方法防止患者隐私受到攻击，利用区块链模型保证诊断记录共享和医疗交易的安全性；采用边缘计算，减少了计算开销和空间成本。基于理论分析和实验证明，该模型为亚线性遗憾患者提供了实时、宝贵的冠心病诊断，实现了高效的隐私保护。

随着智慧医疗技术和应用的不断完善，未来智慧医疗将有更广泛的应用，为医疗资源的合理配置以及医疗服务的升级带来新的机遇和挑战。

4.4 推荐系统

推荐系统指根据用户兴趣或历史行为等信息，预测并推荐符合用户需求的商品、服务、信息等内容的系统。在基于边缘计算的推荐系统中[35]，很多数据处理和推荐计算等操作都可以在本地进行，从而避免了将所有数据传输到云端进行处理的问题，同时还能有效减少延迟和降低网络压力，因此边缘计算在推荐系统中得到了广泛应用。但是，由于涉及对用户敏感信息的处理，在推荐系统中保护用户隐私成为了一个重要问题[36]。

Zheng 等[37]提出一种分布式用户隐私保护推荐框架，同时兼顾对用户位置隐私（基于保序加密函数）和用户偏好隐私（在奇异值分解算法的基础上结合差分隐私技术）的保护，与传统的推荐算法相比推荐性能更好；Du 等[38]提出一个基于联邦矩阵分解的隐私保护推荐系统，在用户级分布式矩阵分解框架的基础上采用同态加密和两阶段随机响应机制进行框架增强，解决了原有框架侧重于保护模型隐私和价值隐私不受不可信推荐者的侵害，但对存在隐私的考虑有限的问题，且大量实验结果表明，该方法可以更有效地保护用户的隐私，模型性能下降更小，计算量也更小；Du 等[39]认为现有的考虑隐式反馈数据隐私安全的推荐算法在构建矩阵分解模型时实际上是基于线性回归模型，并且认为建立分类模型能更好地利用隐式反馈数据，基于上述原因提出了 DPLRMF 算法，引入了 Sigmoid 非线性函数的对称性解决了差分隐私安全性证明的问题，且实验结果表明了该算法的优越性。

4.5 数据聚合

数据聚合指将来自多个来源的数据收集、组合和定义成更有意义和有用的数据集合的过程。它通常涉及收集和整合

在不同时间和位置生成的数据，进而为企业提供更全面、详细和准确的信息，以便做出更明智和有效的商业决策。

在许多场景中，如物流、工业监控、摄像头监测等，需要对数据进行实时采集、处理、分析和应用，以提供时间敏感性的结果和决策，因为这些场景中的数据聚合需要高度的实时性。使用差分隐私技术在边缘设备上实现数据聚合，可以降低数据在传输过程中被窃取和私自改动的风险，还能够降低数据成本，提高数据的质量和完整性，使数据分析和应用更加可信和健壮。

Yang 等[11]提出一种具有本地差分隐私的流数据聚合框架 SPoFC，该框架不利用原始数据流，而是拟合一个类似数据流，极大地节省了隐私预算；基于数据流的凹凸趋势来选择数据的骨架点，而不仅仅只是依靠极值点，有效降低了计算偏差，提高了统计精度。经实验表明，该框架不仅能极大地保护可穿戴设备产生的用户数据的隐私，且相比目前最先进的方法性能更优。Gai 等[40]提出了一种基于本地化差分隐私的智能电网数据聚合方案，用户动态地加入和退出系统时，不会影响其他用户的使用，且不需要重新协商调整参数；该方案设计了基于条件概率的数据离散化算法，减小了聚合结果与实际数据聚合结果之间的差异性，提高了统计精度；最后对该方案进行隐私、效用和性能分析，能较好地估计智能电网的供需状况，且能有效进行隐私保护和减少计算及通信开销。Khan 等[41]提出了一个关于智能电网的容错安全数据聚合方案 FTSDA，假如某些智能电表发生一些故障，没有提交数据，则应用该方案不会影响数据聚合活动，该方案能够保证智能电表高效准确地运行。

上述各场景下的差分隐私保护方案所使用的主要技术如表 1 所列。

表 1　边缘计算下各场景的差分隐私保护方案

Table 1　Differential privacy protection scheme for each scenario in edge computing

应用场景		参考文献	方案	主要技术
基于位置的服务	室内定位	[15]	基于差分隐私的 LBS 用户位置隐私保护方案	Trie 树结构、差分隐私、一致性约束后置处理
		[17]	DP-FLocEC	DP、联邦学习、卷积神经网络
	兴趣点查询	[18]	ADP-FLocEC	Rényi 差分隐私、联邦学习、皮尔逊相关系数
		[20]	DP3-SLOC	地理不可区分性、差分隐私、d_x-隐私机制
		[21]	SPIREL	LDP、连续 POI 推荐框架、迁移学习
车联网		[28]	DPSL	DP、K-均值、Hausdorff 距离
		[29]	车联网中基于差分隐私的个性化位置隐私保护方案	差分隐私、地理不可区分性、多属性决策理论、个性化隐私预算分配算法 PPBA
		[26]	模型、输入、输出扰动 3 种差分隐私防御机制	深度卷积网络、反卷积网络、差分隐私
智慧医疗		[33]	基于边缘计算的安全医疗诊断系统	卷积神经网络、联邦学习 Flower 框架、可追踪环签名技术、差分隐私随机梯度下降算法
		[34]	LCOL	移动边缘计算、本地化差分隐私、自适应扩展树、区块链模型、上下文信息感知
推荐系统		[37]	DDP-SVD	差分隐私、基于奇异值分解的推荐算法、分布式框架、无约束和等比约束随机分片算法
		[38]	基于联邦矩阵分解的隐私保护推荐系统	联邦矩阵分解、两阶段随机响应、同态加密、差分隐私
		[39]	DPLRMF	逻辑回归、矩阵分解、差分隐私、目标扰动技术、sigmoid 非线性函数的对称性
数据聚合		[11]	SPoFC	本地化差分隐私、流数据聚合框架、骨架点法
		[40]	基于随机响应的满足本地化差分隐私的数据聚合方案	本地化差分隐私、k-随机响应、基于条件概率的数据离散化算法
		[41]	FTSDA	非对称和对称加密技术、DP

结束语　本文对边缘计算下差分隐私的应用研究进行了综述。研究表明，差分隐私是一种保护隐私的有效方法，在现有的边缘计算环境下具有广泛的应用前景。然而，在实践中如何平衡数据隐私和数据利用的问题仍需要进一步探索和解决。未来的研究可以着眼于以下几个方面：

1）针对边缘设备的资源限制和计算能力不足的问题，探索如何设计更加轻量级的差分隐私算法。现有的研究中也有针对轻量级的差分隐私算法的，例如基于哈希函数的差分隐私机制、基于压缩算法的差分隐私机制等轻量级的差分隐私机制。在未来的研究中，我们也可以通过压缩数据、降低噪音量化的复杂度等方法探索设计更加轻量级的差分隐私算法。

2）在边缘设备端集成差分隐私算法时，尽量将数据的处理和噪声添加等过程在本地完成，可以考虑在应用差分隐私算法前，在边缘设备上进行数据预处理，通过本地生成噪声并添加到数据中，利用本地计算资源进行部分计算，最大程度地减少数据传输开销。

3）对于异构边缘设备数据间的协作或共享场景，研究如何保证合作过程中每个参与者的隐私安全，并基于差分隐私技术设计合适的协议。设计和实施合适的协议来保证每个参与者的隐私安全是一项复杂的任务，需要综合考虑多方面的因素，包括数据类型、共享目的、隐私需求和技术限制等。Alaggan 等[42]很早就对异构差分隐私有了一定的研究，且实验表明该机制在考虑不同的隐私态度的同时也有良好的效用；Yang[43]对异构信息网络间的差分隐私保护方法也进行了一定的研究，设计了面向相关性的差分隐私保护机制，实验结果显示相关性差分隐私保护机制相比传统方法和交互式方法具有更强的保护效果。

4）在边缘设备上实现更加灵活的数据响应方式，提供定制化的数据服务，使得不同用户可以获得他们需要的数据而不会影响整个系统的隐私保护。要实现上述目标，可以根据

数据分类和敏感性评估的结果设计合适的差分隐私算法，并根据实际需要和隐私保护要求对隐私参数进行调优，在数据访问的过程中实施严格的访问控制和身份验证，并定期评估和更新该方案以适应变化的需求。推荐系统、基于兴趣点的查询等都属于此方面的研究，未来我们需要研究出更加健壮的算法，使隐私安全更加有保障。

5)研究如何在边缘计算应用中建立可信的隐私保护机制，验证由差分隐私所提供的隐私保护能力并保证其正确性。基于边缘计算应用中设计的差分隐私机制已有不少的研究，像应用于物联网、智能交通等的差分隐私机制现如今已经有了很大的发展。要达到上述目标，我们应该根据应用的具体需求和数据类型选择合适的差分隐私机制并进行详细的设计，对该机制进行相应的评估和优化，定期更新。合适的工具和技术的选择是保证验证的准确性和可靠性的重要因素。

总的来说，将差分隐私技术应用于边缘计算场景具有广阔的发展前景，未来还有很多需要探索和发展的研究方向。在这一领域中，如何在隐私保护和数据处理之间寻找平衡，提高技术的可行性和效率是一个关键点。

参 考 文 献

[1] GE B, WU C, ZHANG T H, et al. Privacy Protection Method of Edge Computing Based on Federated Learning [J]. Journal of Anhui University of Science and Technology (Natural Science Edition), 2022, 42(6): 79-86.

[2] LI X W, CHEN B H, YANG D Q, et al. Review of Security Protocols in Edge Computing Environments [J]. Journal of Computer Research and Development, 2022, 59(4): 765-780.

[3] MYKTAR Y, MOHD Y I I, AINUDDIN W B A W, et al. Systematic Review on Security and Privacy Requirements in Edge Computing: State of the Art and Future Research Opportunities [J]. IEEE Access, 2020, 8: 76541-76567.

[4] LIU X F, QU Y X, WANG Y, et al. Privacy intelligent inference prediction in Edge Computing environment [J]. Artificial Intelligence, 2019(5): 45-54.

[5] SHEN C N. Research progress of edge computing security and privacy Protection [J]. Network Security and Data Governance, 202, 41(8): 41-48.

[6] DWORK C. Differential privacy: A survey of results [J]. Foundations and Trends © in Theoretical Computer Science, 2011, 9(3/4): 211-407.

[7] DWORK C, ROTHA. The Algorithmic foundations of differential privacy [J]. Found. Trends Theor. Comput. Sci., 2014, 9(3/4): 211-407.

[8] ZHAO Y Q, YANG M. A review of Differential privacy Research [J]. Journal of Computer Science, 2023, 50(4): 265-276.

[9] XIONG P, ZHU T Q, WANG X F. Differential Privacy Protection and its Application [J]. Chinese Journal of Computers, 2014, 37(1): 101-122.

[10] YE Q Q, MENG X F, ZHU M J, et al. A review of localized differential privacy [J]. Journal of Software, 2018, 29(7): 1981-2005.

[11] YANG M M, LAM K Y, ZHU T Q, et al. SPoFC: A framework for stream data aggregation with local differential privacy [J]. Concurrency and Computation: Practice and Experience, 2022, 35(5). https://doi.org/10.1002/cpe.7572.

[12] TAN Z W, ZHANG L F. Review of Machine learning privacy Protection [J]. Journal of Software, 2020, 31(7): 2127-2156.

[13] CHEN Q. Research on Privacy Protection of Private Location by Edge Computing Difference [D]. Lanzhou: Lanzhou Jiaotong University, 2020.

[14] ZHANG L, LIU Y, WANG R Z. Data publishing technology based on differential privacy in Location Big Data Service [J]. Journal of Communications, 2016, 37(9): 46-54.

[15] YU N W, YANG S J, CHEN Z G, et al.. LBS user location Privacy Protection scheme based on Differential privacy [J]. Journal of Hebei University of Science and Technology, 2021, 42(3): 222-230.

[16] ZHANG Q Y, ZHANG X, LI W J, et al. Overview of Location Trajectory Privacy Protection Technology Based on LBS System [J]. Computer Application Research, 2020, 37(12): 3534-3544.

[17] ZHANG X J, HE F C, GAI J Y, et al. Differential private federal learning model for fingerprint indoor location under edge computing [J]. Computer Research and Development, 2022, 59(12): 2667-2688.

[18] HE F C. Research on key technologies of indoor location federated learning model supporting privacy protection in edge computing environment [D]. Lanzhou: Lanzhou Jiaotong University, 2022.

[19] ZHANG Q Y, ZHANG X, LI W J, et al. Design of Interest Point Recommendation Algorithm Based on Differential Privacy Protection [J]. Computer Applications and Software, 2019, 36(9): 243-248, 269.

[20] ZHANG X J, YANG H Y, LI Z, et al. Differential Private Location Privacy Protection Method Integrating Semantic Positions [J]. Computer Science, 2021, 48(8): 300-308.

[21] JONG S K, JONG W K, YON D C. Successive Point-of-Interest Recommendation with Local Differential Privacy [J]. arXiv: 1908.09485, 2019.

[22] AFZAL, KIRAN, TARIQ, et al. An Optimized and Efficient Routing Protocol Application for IoV [J]. Mathematical Problems in Engineering, 2021, 2021(Pt. 23): 9977252.1-9977252.32.

[23] DENG Y K, ZHANG L, LI J. Summary of privacy protection research on the Internet of Vehicles [J]. Computer Application Research, 2022, 39(10): 2891-2906.

[24] ZHAO P, ZHANG G L, WAN S H, et al. A survey of local differential privacy for securing internet of vehicles [J]. The Journal of Supercomputing, 2019, 76(11): 8391-8412.

[26] WU M Q, HUANG X M, KANG J W, et al. Research on Differential Privacy Protection for Vehicle Road Collaborative Inference [J]. Computer Engineering, 2022, 48(7): 29-35.

[25] RAZA S, WANG S G, AHMED A, et al. A Survey on Vehicular Edge Computing: Architecture, Applications, Technical Issues, and Future Directions [J]. Wireless Communications and Mobile Computing, 2019, 2019: 3159762: 1-3159762: 19.

[27] ZHONG X Y, LI M H, LI L H. Summary of Research on Privacy Protection of Vehicle Internet Location [J]. Internet of Things Technology, 2023, 13(3): 77-79.

[28] XIE S S, LIU H L, ZHAO G S. Differential Privacy Protection Method Based on Location Semantics in the Internet of Vehicles

[J]. Small Micro Computer Systems,2023,45(4):984-990.

[29] XU C,DING Y Y,LUO L,et al. Personalized Location Privacy Protection Based on Location Services in the Internet of Vehicles [J]. Journal of Software,2022,33(2):699-716.

[30] QU P R. The Application of Internet of Things Technology in Smart Healthcare [J]. Internet Weekly,2022(22):45-47.

[31] AN J,NING H J. Research on Security and Privacy Protection for Smart Medical Networks [J]. China New Communications,2022,24(17):122-124,134.

[32] AHAD A,ALI Z,MATEEN A,et al. A Comprehensive review on 5G-based Smart Healthcare Network Security: Taxonomy, Issues,Solutions and Future research directions[J/OL]. Array,2023,18. https://doi.org/10.1016/j.array.2023.100290.

[33] LI B S. Privacy Protection System in Edge Intelligent Collaborative Computing Mode [J]. Electronic Testing, 2022,36(20): 59-62.

[34] LIU X,ZHOU P,QIU T,et al. Blockchain-Enabled Contextual Online Learning under Local Differential Privacy for Coronary Heart Disease Diagnosis in Mobile Edge Computing[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 20(8): 2177-2188.

[35] SUN C,LI H,LI X,et al. Convergence of Recommender Systems and Edge Computing: A Comprehensive Survey[J]. IEEE Access,2020,8:47118-47132.

[36] FENG H,YIN H W,LI X H,et al. Summary of privacy protection research in recommendation systems [J]. Computer Science and Exploration,2023,17(8):1814-1832.

[37] ZHENG X Y,LUO Y L,WANG X S,et al. Research on Distributed Differential Privacy Recommendation Method Based on Location Services [J]. Journal of Electronic Science, 2021, 49(1):99-110.

[38] DU Y J,ZHOU D Y,XIE Y,et al. Federated matrix factorization for privacy-preserving recommender systems[J/OL]. Applied Soft Computing Journal, 2021, 111. https://doi.org/10.1016/j.asoc.2021.107700.

[39] DU M K,PENG J J,HU Y J,et al. A Logistic Regression Matrix Decomposition Recommendation Algorithm Satisfying Differential Privacy [J]. Journal of Beijing University of Posts and Telecommunications,2023,46(3):115-120.

[40] GAI N,XUE K P,ZHU B,et al. An efficient data aggregation scheme with local differential privacy in smart grid[J]. Digital Communications and Networks,2022,8(3):333-342.

[41] KHAN H M,KHAN A,KHAN B,et al. Fault-Tolerant Secure Data Aggregation Schemes in Smart Grids: Techniques, Design Challenges,and Future Trends [J]. Energies, 2022, 15(24): 9350.

[42] ALAGGAN M,GAMBS S,KERMARREC M A. Heterogeneous Differential Privacy[J]. Journal of Privacy and Confidentiality, 2017,7(2).

[43] YANG L. Research on Differential Privacy Protection Methods for Heterogeneous Information Networks [D]. Harbin: Harbin Institute of Technology,2019.

SUN Jianming, born in 1980, postdoctor, professor, master supervisor. His main research interests include pattern recognition, intelligent agriculture, machine vision, image information processing and automatic control.

ZHAO Mengxin, born in 1999, postgraduate. Her main research interests include differential privacy and data mining.