[](https://www.acm.org/publications/policies/artifact-review-and-badging-current)

Skelam混合机制：一种新的不同隐私联合学习方法

新加坡国立二古堡大学bao@u.nus.edu

朱一正新加坡国立大学yzhu@nus.edu.sg

肖小葵新加坡国立大学xkxiao@nus.edu.sg

阴阳

哈马德·本·哈利法大学yyang@hbku.edu.qa

BengChinntinOonationalnivestrinpareooibc@comp.nus.edu.sg

谭洪孟

A\*STAR，新加坡benjamin\_tan@i2r.a-明星.edu.sg

Khin Mi Aung A\*STAR，新加坡mmaung@i2r.a-star.edu.sg

# 抽象

深度神经网络具有强大的记忆底层训练数据的能力，这可能是一个严重的隐私问题。这个问题的一个有效解决方案是训练具有差分隐私（DP）的模型，该模型通过向梯度注入随机噪声来提供严格的隐私保证。本文关注的是敏感数据分布在多个参与者之间的场景，这些参与者通过联合学习联合训练模型，使用安全多方计算（MPC）来确保每个梯度更新的机密性，并使用差分隐私来避免生成的模型中的数据泄漏。这一集合中的一个主要挑战是，在深度学习中执行DP的常见机制（注入实值噪声）与MPC（在参与者之间交换有限域整数）根本不兼容。因此，大多数现有的DP机制需要相当高的噪声水平，导致模型实用性差。

**PVLDB工件可用性：**

源代码、数据和/或其他工件已在https://github.com/SkellamMixtureMechanism/SMM.

# 引言

深度神经网络，尤其是GPT-3[10]等大规模神经网络，以其出色的记忆能力而闻名[22，43，52]。然而，很难控制神经网络到底记住了什么，当基础训练数据包含阳性信息时，意外的数据记忆可能是一个严重的问题[12]。例如，考虑一家银行，它在呼叫中心记录上训练类似GPT的语言模型。由于数据记忆，可以通过让模型自动完成前缀来提取敏感信息，例如“我的账号是：

”。显然，如果这种模型（或其API）暴露给对手，

它变成了一台结扎机器，因为攻击者可以尝试使用各种

基于此，我们提出了Skellam混合机制

(

提取敏感数据的前缀，然后起诉银行

SMM），一种在通过feed构建的模型上强制执行DP的新方法-

纯粹的学习。与现有方法相比，SMM消除了输入梯度必须是整数值的假设，从而减少了为保持DP而注入的噪声量。SMM的理论分析是非常不平凡的，特别是考虑到（i）DP深度学习的复杂数学，以及（ii）两个Skellam分布的混合相当复杂。在各种实际设置上的大量实验表明，SMM在所得模型的实用性方面始终显著优于现有解决方案。

**PVLDB参考格式：**

额尔古特宝、朱一正、萧孝奎、阴阳、蚌钦、本雅明、洪孟旦、钦密昂。Skellam混合机制：一种具有差分隐私的联合学习新方法。PVLDB，15（11）：2348-23602022。doi:10.1478/3551793.35551798

本作品获得知识共享BY-NC-ND 4.0国际许可。参观https://creativecommons.org/licenses/by-nc-nd/4.0/以查看此许可证的副本。对于超出本许可范围的任何使用，请发送电子邮件至info@vldb.org.版权归所有者/作者所有。授予VLDB捐赠基金的出版权。

VLDB基金会论文集，第15卷，第11期，ISSN 2150-8097。doi:10.1478/3551793.35551798

侵犯隐私。Shokri等人[42]报告称，简单直观的措施往往无法提供足够的保护，而完全解决这一问题的唯一方法是在严格保证差分隐私（DP）的情况下训练模型[18]。

本文重点讨论了多个个体参与者通过分布式随机梯度下降（SGD）[1，15，17，33]使用联合学习（FL）[34]联合训练机器学习模型的场景。具体地，在每次迭代中，每个个体基于她自己的数据计算相对于当前模型权重的梯度；然后，来自所有参与者的梯度被聚合以更新模型。请注意，来自每个个体的梯度可能会揭示关于她的原始数据集的敏感信息[35，38，41，42，51]。解决这个问题的一种常见方法是采用安全多方计算（MPC）协议[5，7，13，16，23，27，50]，该协议计算聚合梯度，同时保留每个参与者的梯度机密性。MPC的一个优点是，它不需要可信的第三方，这在某些应用中可能很难建立，例如在金融和医疗保健中。

请注意，尽管MPC通过隐藏每个参与者的梯度值来保护个人在梯度更新过程中的隐私，但它并不能提供任何保护，防止意外数据记忆引起的数据提取攻击[20，35，44，45]。

如前所述，抵御此类攻击的有效方法是扰动梯度以满足差分隐私[42]。由于在我们的环境中没有可信的第三方，因此需要以去中心化的方式进行此类动态扰动。在里面

以0.01概率变为1，以0.99概率变为0。虽然舍入梯度的期望值等于原始期望值，但舍入梯度的范数可能明显大于原始期望值-

第一。在我们的示例中，x = {0.01， 0.01， . . .√，0.01} ∈ Rd 可能是

特别地，每个FL参与者首先将噪声添加到她自己的梯度；

四舍五入为 {1， 1， . . . ， 1}，导致 L2 几乎增加

然后，参与者通过加密安全协议（例如，SecAgg[9]）共同聚合他们的噪声梯度，该协议确保（i）稍后基于聚合结果更新模型的服务器除了结果本身之外，什么也不了解受扰动的数据，以及（ii）没有参与者了解其他参与者数据的私人信息，除了聚集性结果。因此，在每次迭代中产生的隐私成本仅取决于梯度之和的灵敏度以及聚集噪声的分布。该框架被称为分布式差分隐私[24，28]，在第2.4节中详细阐述。

尽管在以前的工作中已经研究了DP下的梯度扰动，但将集中式DP解适应于分布式设置绝非易事。例如，考虑经典的DPSGD算法[2]，其中集中式方在模型训练过程的每次迭代中向梯度和注入随机高斯噪声。使DPSGD适应分布式DP有两个主要挑战。首先，高斯分布是在所有实数的域上定义的，而据我们所知，现有的MPC协议要求将输入表示为整数（更准确地说，是有限域元素）[8，9，39]。尽管我们可以从高斯分布中采样一个实数，然后将该值量化为整数，但得到的量化样本将不再遵循高斯分布，严格来说，这是[2]中证明该方法满足DP的有效性。其次，即使我们忽略了在DPSGD[2]中使用量化高斯样本的隐私风险，该算法的隐私分析也依赖于高斯分布的某些数学性质，而这些数学性质在量化样本中不再适用。这些特性包括：

1. 从 i.i.d. 单位方差高斯采样的 n 个值的总和

分布服从方差为 n 的高斯分布，并且

1. 在两个高斯分布之间存在Rényi散度[36]的严格上界。这些问题在很大程度上被早期的分布式DP解决方案所忽视，例如[24，46，47]。

最近，提出了三种方法[3，4，28]来解决上述问题。这些方法的一个常见的高级思想是要求FL的每个参与者在训练过程的每次迭代期间向梯度注入对称整数值噪声（例如[4]中的二项式噪声）。这里，假设梯度的原始值具有有界范数，并且是整数值，稍后将详细说明。接下来，我们指出了现有方法的主要缺点，这也是我们提出的解决方案的动力所在。回想一下上面的两个假设：梯度具有有界范数，并且是整数值的。虽然有界范数假设可以像DPSGD[2]中所做的那样通过梯度剪裁来强制执行，但整数值梯度假设需要更复杂的预处理步骤，并伴随着仔细的隐私分析。在这里，我们简要解释了现有的工作是如何强制执行这一消耗的，我们将详细讨论推迟到第5节。参考文献[4]将梯度随机四舍五入为整数。例如，如果

*𝑥 = {0.01, 0.01, . . . , 0.01} ∈ R𝑑 , then each dimension of 𝑥 is rounded*

标准这种增加的灵敏度导致了满足DP所需的更高的扰动量，这反过来又导致了模型实用性的降低。参考文献[28]提出了一种更复杂的条件舍入过程，以（在有限程度上）缓解灵敏度增加的问题，代价是在所得梯度中引入额外的偏置项，如第5节所述。

**我们的贡献。在这项工作中，我们提出了Skellam混合机制（SMM），这是一种为联邦学习强制执行分布式差分隐私的新解决方案。SMM 的工作原理是注入从两个偏移对称 Skellam 分布的混合物中提取的随机噪声。与现有解决方案不同，SMM不要求其输入（即FL中的梯度）为整数值。这消除了对渐变进行随机舍入过程的需要，从而降低了满足DP所需的噪声水平，从而提高了结果模型的效用。特别是，通过精心选择的混合系数和Skellam分布参数，SMM生成私有且无偏的整数值梯度聚合，用于更新模型。我们证明SMM同时满足第2节中定义的Rennyi-DP和（ε，δ）-DP。同时，SMM与DPSGD [2]框架及其矩计数分析技术兼容，导致隐私丢失分析受到严格限制，类似于我们的竞争对手[3，28]。**

SMM的隐私分析相当具有挑战性

论文的主要贡献。请注意，尽管Skellam分布已在先前的解决方案[3]中使用，但[3]中的隐私分析不适用于我们的设置，因为SMM中的随机噪声是从两个移位对称Skellam分配的混合中采样的，这比[3]中的单个Skellam分布更复杂。此外，尚不清楚如何使用[3]中内置的结果和数学工具推导SMM的严格隐私界限。一个主要原因是参考文献中的隐私界限。[3]对于Skellam噪声，需要输入的1和2灵敏度的上界，并且尚不清楚如何在Skellam混合分布的情况下扩展该界限。为了解决这个问题，我们首先推导出仅涉及2个灵敏度的Skellam噪声的更干净的隐私边界，这是我们对SMM进行隐私分析的基础。我们对Skellam分布的分析技术具有独立的兴趣，并且可以应用于[3]的设置，以通过消除对输入的1灵敏度的依赖来提高其隐私边界。

L

L

双环磁性扫雷

我们将SMM应用于分布式SGD的联合学习，并给出了完整的训练算法。如上所述，SMM通过消除舍入梯度的步骤来提高模型效用，这通常显著增加输入的灵敏度，尤其是对于大型模型。使用基准数据集的大量实验表明，在具有不同隐私和通信限制的各种设置下，SMM比其竞争对手获得了一致且显著的效用收益。

# 初步的

## Skellam分布

随机变量 Y 服从参数 λ 的泊松分布，如果其概率分布为 Pr[Y = k] = exp（−λ）λk， k = 0， 1， 2， . .

.随机变量

isk！的均值和方差都遵循

*𝑌𝜆𝑍*

斯凯拉姆分布，如果它是两个独立泊松变量 Y1 和 Y2 之间的差值。在这项工作中，我们将注意力限制在Y1和Y2具有相同参数λ的情况上。在这种情况下，Z 的概率分布为

Pr[𝑍 = 𝑘] = exp(−2𝜆)𝐼 |𝑘 | (2𝜆), 𝑘 = 0, ±1, ±2, . . . ,

定义2（（ε，δ）-差分隐私[18]）。随机机制 M 满足 （ε， δ） 差分隐私 （DP） 如果

Pr[M（X ] ∈ O] ≤ exp（ε） ·Pr[M（X'） ∈ O] + δ，

*对于任何一组输出范围和任何相邻数据集*

O⊆（M）

*𝑋 and 𝑋 ′.*

请注意，ε，δ -DP 可以被视为机制的最坏情况下隐私保证，因为它对所有可能结果的概率比强制执行上限。另一种定义是仁义差分隐私（RDP）[36]，它建立在仁义分歧的概念之上，而是考虑了平均案例隐私保证。

()

其中 Iv （u） ≜ .∞ 1 u 2h+v 是修改后的贝塞尔函-

*h=0 h！C （h+v+1）*

2

定义3（Rényi Differential Privacy[36]）。随机的-

第一种。我们写 Z ∼ Sk（λ， λ）。线性度为

*ized机制*

*使满意*

*α， τ -仁义差分隐私 （RDP）*

期望值，Z 的均值为 0，方差为 2λ。

*如果 D*

*XM ′*

*τ f（或al）l 相邻数据集 X 和 X ′。*

Skellam分布有一个重要的特性：它们是“ad”-

*α （M（ ） M（X）） ≤*

ditive“，也就是说，对于任何两个独立的 Skellam 随机变量 Z1 Sk λ1， λ1 和 Z2 Sk λ2， λ2，它们的和 Z1 Z2 服从 Skellam 分布 Sk λ1 λ2， λ1 λ2 。此属性对于我们分析解决方案中使用的 Skellam 混合物噪声的隐私保证至关重要。

(++)

~()∼()+

## 雷尼分歧

给定一个感兴趣的函数，使其在本质上私有的规范方法是通过噪声注入来干扰其结果。具体而言，噪声的尺度应校准为感兴趣函数[18]的灵敏度，其形式定义如下。

:D第4（灵敏度）。函数的灵敏度 S （F）

*F D → Rd，记作 S（F），定义为*

定义1（仁义背离[48]）。假设分布 P 和 Q 是在同一域上定义的，并且 P 是绝对的

*𝑆 𝐹 = max*

*𝑋 ∼𝑋 ′*

( )

*F （X ） − F （X ′） ，*

*关于Q的连续性，则P与Q的仁义散度*

*有限阶α ∈ （0， 1） ∪ （1， ∞） 定义为：*

1

"#

*where 𝑋 ∼ 𝑋 ′ denotes that 𝑋 and 𝑋 ′ are neighboring datasets, and*

·是一种常态。

*Dα （P Q） = α − 1 log EX ∼P*

*𝑃 (𝑋 ) 𝛼 −1*

*𝑄 (𝑋 )*

*,*

特别地，注入从

函数 F 的每个维度的 N （0， σ2） 满足 （α， αS 2 （F） ）-

*我们采用的约定是 0 = 0 和 y = ∞ 对于任何 y > 0，*

RDP[36]，其中

*𝑆 (𝐹 )*

代表

2𝜎 2

L2函数灵敏度

00

*对数以 E 为底。*

接下来我们给出Rényi散度的一些有用性质。

定理 1（凸性 [48]）。对于任何订单 α0.和

0∈ [ ∞]

*< 𝜆 < 1, Rényi divergence is convex in its second argument. That is,*

*对于任何概率分布 P、Q0 和 Q1*

*𝐹 .在许多应用中（例如，使用SGD训练神经网络），我们还需要分析由多个组件组成的机械的整体隐私保证。我们有以下RDP机制的组成和子采样引理。*

引理1（RDP机制的组成[36]）。如果机械-

*分别是 M1， . . . ， MT 满足 （α， τ1）， . . . ， （α， τT ）-RDP*

*𝑡 =1*

*Dα （P （1 − λ） · Q0 + λ · Q1） ≤ （1 − λ） ·Dα （P Q0） + λ ·Dα （P Q1）.*

*然后，M1 ◦ . . .◦ MT satisfies （α， .T τi ）-RDP.*

定理2（联合准凸性[48]）。对于 [0.∞] 和 0 < λ < 1 的任何α ∈阶，仁义散度在其参数中是联合准凸的，即对于任意两对概率分布（P0， Q0） 和 （P1， Q1）

*Dα （（1 − λ） · P0 + λ · P1 （1 − λ） · Q0 + λ · Q1）*

引理 2（RDP 的子采样 [37， 53]）。设 M 是一个机甲，满足 （l， τ （l ））-RDP for l = 2， . . . ， α （α ∈ Z， α ≤ 2） 和

*Sq 是一个以概率 q 统一采样输入数据每条记录的过程。则 M ◦ Sq 满足 （α， τ ）-RDP 与*

1

*𝜏 = 𝛼 − 1 ·*

*𝛼*

≤ max{Dα （P0 Q0）， Dα （P1 Q1）}.

)(

## 差异隐私

日志

(1 − 𝑞 𝛼 −1 𝛼𝑞 − 𝑞 + 1) +

∑︁𝑙 =2

*𝛼 (1 − 𝑞)𝛼 −𝑙 𝑞𝑙 𝑒 (𝑙 −1)𝜏 (𝑙 ) ! .*

我们说两个数据集 X 和 X ′ 是相邻的，如果可以通过添加或删除另一个元组来获得一个元组。差分隐私（DP）的主要思想是确保相邻数据集上的随机机制的结果始终相似;直观地说，这为给定的数据记录X是否属于数据集X提供了合理的否认，从而保护了

*𝑙*

记录为 x 的个人的隐私。差分隐私的经典定义是（ε，δ）-DP [18]，如下所示。

最后，任何满足 （α， τ ）-RDP 的机制也满足

（ε，δ）-DP，对于ε和δ的值，如下所示。

引理 3 （将 （α， τ ）-RDP 转换为 （ε， δ）-DP [11]）。对于任何

*𝛼 ∈ (1, ∞), if 𝐷𝛼 (M(𝑋 ) M(𝑋 ′)) ≤ 𝜏 for any neighboring databases 𝑋 and 𝑋 ′, then M satisfies (𝜖, 𝛿)-DP for*

*𝜖 = 𝜏 + log(1/𝛿) + (𝛼 − 1) log(1 − 1/𝛼) − log(𝛼) .*

*𝛼 − 1*

## 分布式差异隐私

原始的集中式差分隐私框架[18]概括为可信的数据管理员，他存储整个私有数据集，并在其对查询的响应中注入随机噪声，例如总和

**算法1：一维Skellam混合机制（1SMM）**

**输入：一组私有值 x1， . . . ， xn xiR .**

{|∈ }

**参数：噪声参数 λ。**

查询，计算给定输入数据集 X = （x1， . . . ， xn） 的 .n xi 。

*𝑖=1*

**1 for i ∈ 1..n do**

当注入的噪声的规模根据查询的灵敏度进行校准（由集中式数据管理器）时，发布的结果满足（集中式）DP。在这项工作中，我们重点关注分布式差分隐私框架[14，21，24，28]，它涉及多个参与者。每个参与者向自己的数据或查询响应注入随机噪声。之后，所有参与者集体

**2pi = xi − ⌊xi ⌋。**

**3来自伯努利试验的样本 yi 具有成功概率 pi**

**4if 𝑦𝑖 = 0 then**

**5xi∗ ← ⌊xi ⌋ + Sk（λ， λ）。**

**6其他**

**7xi∗ ← ⌊xi ⌋ + 1 + Sk（λ，λ）。**

运行MPC协议，通过隐藏参与者的身份来增强隐私保障。我们遵循与先前工作[28]中相同的威胁模型。特别是，所有参与者都是诚实的（即，他们严格遵守协议），但很好奇（即，每个参与者都试图从另一个参与者那里学习私人信息），并且假设没有两个参与者串通。在本文中，我们重点讨论了作为MPC协议的SecAgg[9]，它聚合了输入

根据我们的

**8 𝑥¯∗𝑆𝑒𝑐𝐴𝑔𝑔 𝑥1∗, . . . , 𝑥𝑛∗ .**

**输出：x ̄∗。**

←( ())

噪音 Zi 到确切的总和：

M(𝑥1, . . . , 𝑥𝑛 ) := ∑︁ 𝑥𝑖 + ∑︁ 𝑍𝑖 .

*𝑖=1*

*𝑖=1*

*𝑛𝑛*

威胁模型。具体来说，SecAgg确保除了发布的输出之外，没有人（包括参与者）能够推断出任何关于私有输入的信息。SecAgg的输出应该满足DP，这样它就可以以与集中式DP解决方案的结果相同的方式进行分发。因此，注入到数据或查询结果的总体噪声的规模需要根据每个参与者的输入的灵敏度进行校准。换句话说，分布式DP获得了与集中式DP相同的隐私效用权衡

在隐私方面，我们专注于RDP定义（定义3），可以通过引理3将其转换为经典（ε，δ）-DP（定义2）。特别是，我们希望对于所有相邻数据集 X ， X ′，

*𝐷𝛼 (M(𝑋 ) M(𝑋 ′)) ≤ 𝜏,*

对于某些α > 1.我们测量 M 的误差

*Err= 最大值 1 E ̈M（X） − ∑︁ x ̈ ，*

2

设置，而不依赖可信的第三方。

M𝑋 ⊂R𝑑 𝑑

*𝑥 ∈𝑋 ¨2*

# Skelam混合机制

第3.1节形式化了分布式DP下的分布式和估计问题，第3.2节提出了针对该问题的Skellam混合机制（SMM）。第3.3节介绍了SMM隐私保障的基础。第3.4节规定了SMM的隐私和效用保障。然后，在第4节中，我们将SMM应用于我们的主要问题设置：差异私有联合学习。

## 具有保密性的分布式和估计

假设一个多维数据集 X = x1， . . . ， xn 分配给 n 个人（在下文中称为参与者），其中参与者 i 拥有数据点 xi Rd，因为 i = 1， . . . ， n。

∈

()

不受信任的服务器旨在计算

其中期望接管M中的随机性。

## 斯凯拉姆混合噪音

我们首先考虑每个参与者的数据点 xi 是一维的情况。算法 1 显示了这种情况的一维 Skellam 混合机制 （1SMM） 的伪代码。每个参与者我首先独立掷硬币与正面概率

*pi ：= xixi（第 2 行和第 3 行）。如果是尾巴，则参与者在 Skellam 分布 Sk λ、λ（第 4 行和第 5 行）之后发出噪声;否则，参与者扰动xi1（即*

⌊ ⌋ +

⌊ ⌋()

– ⌊ ⌋

*𝑥𝑖 ) with a noise following the Skellam distribution 𝑆𝑘 𝜆, 𝜆 (Lines 6 and 7). Note that, by the definitions of 𝑥𝑖∗ and the Skellam distribu- tion, 𝑥𝑖∗ is guaranteed to be an integer. Finally, SecAgg aggregates the noisy values from all the participants (Line 9), and the esti-*

⌈ ⌉()

配对总和 x ̄∗ 被释放到服务器。对于每个

参与者的数据点xi是多维的，我们只需为每个维度独立调用算法1即可获得噪声

数据集，即 x ̄ = .n xi ，来自参与者。阿加瓦尔等

*𝑖=1*

该维度的总和，如算法2中所述。

等人[4]提出了一个求解分布和的通用框架

分布式 DP 的估计问题。在这个框架中，每个参与者首先用噪声 Zi 来扰动她的数据 xi：x ∗ xi Zi。接下来，一个安全聚合协议 SecAgg [9]，由参与者作为黑匣子运行，总结来自所有参与者的噪声值 xi∗，

←()

*𝑖 ←+*

并将结果 x ̄∗ SecAgg x1∗， . . . ， xn∗ 输出到服务器。

根据参考文献[9]，SecAgg确保没有参与者（或服务器）了解到任何关于另一参与者私人信息的信息

数据。因此，推导出以下机制 M 的隐私和效用保证就足够了，该机制注入了 n 个独立的随机

算法 1 的结果乍一看似乎很难分析，因为伯努利有 2n 个可能的结果

所有参与者的试验。我们分析中的一个重要见解是，为了推导算法1的效用保证，独立考虑每个参与者就足够了。首先，注意扰动值

*𝑥𝑖∗ follows a mixture of two shifted symmetric Skellam distributions*

其移位平均值分别等于 xi 和 xi，并且

⌊ ⌋⌈ ⌉

每个分布的方差等于 2λ。此外，观察到与混合物分布相关的权重为 1−xi +⌊xi ⌋ 和

*𝑥𝑖 − ⌊𝑥𝑖 ⌋, respectively. Consequently, the expectation of 𝑥𝑖∗ equals*

**算法2：多维Skellam混合机制（dSMM）**

**输入：一组私有值 x1， . . . ， xn xiRd 。**

{|∈}

**参数：噪声参数λ，数据维度d。**

**1 for i ∈ 1..n do**

**2for j ∈ 1..d do**

**3pi， j = xi， j − ⌊xi， j ⌋.**

**4来自伯努利试验的样本 yi，j 具有成功概率**

*皮， J.*

**5if 𝑦𝑖,𝑗 = 0 then**

**6𝑥𝑖∗, 𝑗 ← ⌊𝑥𝑖,𝑗 ⌋ + 𝑆𝑘 (𝜆, 𝜆).**

**7其他**

**8𝑥𝑖∗, 𝑗 ← ⌊𝑥𝑖,𝑗 ⌋ + 1 + 𝑆𝑘 (𝜆, 𝜆).**

**9 𝑥¯∗𝑆𝑒𝑐𝐴𝑔𝑔 𝑥1∗, . . . , 𝑥𝑛∗ .**

←( ())

**输出：x ̄∗。**

FL），在第 4 节中进一步阐述。请注意，尽管我们的分析将 Δ 的值限制为 Δ < 2λ α 1，但这种应变仅影响 Skellam 噪声的效用，而不会影响其隐私保证。这是因为可以通过标准裁剪轻松强制实施约束。此外，在联邦学习设置中，

*λ通常比最优α（RDP的阶数）大得多，因为大量的参与者对整体DP噪声有贡献，并且最佳α通常相对较小（例如在我们的实验中小于10）。因此，上述约束会导致足够大的削波范围，而不会导致太多效用降级。*

L∞

L∞

③〖手稿中删去以下这句话:

我们在定理4中给出的理论结果与[3]中对Skellam噪声的分析之间的一个显著区别是，我们的结果是“更干净”的，因为方程（2）只涉及2范数（类似于连续高斯噪声的情况[36]），而

L

[3] 中的分析还涉及向量 s 的 L1 范数。一般来说，

L1灵敏度的存在可能导致过多的噪声

*𝑥𝑖 . A corner case is that 𝑥𝑖 is an integer. In this case, the perturbed*

对于高维数据，由于L1灵敏度可以√倍

大于L2灵敏度，限制了Skellam的适用性

*𝑥 ∗ can be seen as injecting symmetric Skellam noise 𝑆𝑘 𝜆, 𝜆 to 𝑥𝑖*

*𝑖()*

仅本身。根据期望的线性，x ̄∗的期望也等于 i=1 xi，即 1SMM 产生私有输入总和的无偏估计器。我们将在后面的第3.4节中介绍1SMM和dSMM的详细隐私和效用分析。

.

## Skelam Noise保护隐私

在我们分析SMM的隐私保证之前，我们首先表明其构建块，即单个对称Skellam噪声，保持了隐私，形式化如下。

定理3（Skellam分布的Rényi散度）。

*对于任何整数 s ∈ Z 满足 |s |≤ Δ∞、任何 1 α >和任意 Δ∞*

*满足 α < 2λ/Δ∞ + 1，我们有*

这种应用中的噪声。此外，没有

1范数项还可以显著简化建立在加性Skellam噪声之上的原型和机制的设计，例如，前面提出的算法1SMM和dSMM。为了避免1范数项，我们不使用Rényi散度的已知性质，而是直接使用基本数学工具来解决这个问题，这是一种独立感兴趣的新证明技术（详细介绍在技术报告第[6]页的附录C.1中）。特别是，这种证明技术在一开始就导致了冗长而沉重的公式，但在几个步骤内，大多数项都被抵消了，从而产生了一个干净的边界。

L

L

最后，我们提到存在Skellam分布的精确采样器，如技术报告附录A所述

1.09𝛼 + 0.91 𝑠2

版本[6]）。因此，添加Skellam噪波会严格保留

*𝐷𝛼 (𝑠 + 𝑆𝑘 (𝜆, 𝜆) 𝑆𝑘 (𝜆, 𝜆)) ≤2· 2𝜆 .(1)*

差异隐私。相反，我们不知道连续高斯分布的精确采样器。因此

我们有以下多维扩展。

定理 4（多维斯凯拉姆分布的仁义散度）。设 Skd λ， λ 表示一个 d 维变量，其中每个维度都从 Sk λ， λ 独立采样。然后，对于任何

()

()

*整数值向量 s ∈ Zd 满足 s 2 ≤ c 和 s ∞ ≤ Δ∞、任何 α > 1 和任何 Δ∞满足 α < 2λ/Δ∞ + 1，我们有*

2

*𝐷𝛼 (𝑠 + 𝑆𝑘𝑑 (𝜆, 𝜆) 𝑆𝑘𝑑 (𝜆, 𝜆)) ≤ 1.09𝛼 + 0.91 · 𝑐 .(2)*

2

2𝜆

上述定理的证明可以在技术报告版本的附录C.1中找到[6]。接下来，我们强调我们的理论成果的贡献。首先，根据定理 3，方差为 2λ 的对称 Skellam 噪声提供的隐私保证与添加相同方差的连续高斯噪声的隐私保证相当（即在常数因子内），后者

is 𝛼 ·𝑠2 [36]. Further, as Eq. (1) only involves the quadratic term,

2 2𝜆

·

用不精确采样器采样的随机噪声仅近似遵循高斯分布；严格地说，注入这样的噪声可能违反差分隐私，这是采用我们提出的注入Skellam噪声的方法的另一个动机。

## Skellam混合机理的理论分析

我们对所提出的Skellam混合机制（SMM）进行了理论分析。证明推迟到技术报告版本[6]中的附录C中。在隐私方面，我们对算法1有以下定理。

定理 5.假设每个参与者的数据点 xi 满足

|𝑥𝑖 |2 + (|𝑥𝑖 | − ⌊|𝑥𝑖 |⌋) − (|𝑥𝑖 | − ⌊|𝑥𝑖 |⌋)2 ≤ 𝑐

一维隐私分析可以很容易地扩展到

4𝑛𝜆

*和⌈|xi|⌉ ≤δ∞。然后，只要 1 和 Δ∞ α >满足*

通过用平方2范数代替二次项的多维设置，如定理4所示。

L

同时，由于加性Skellam噪声保留了RDP，由Lem-

2𝑛𝜆

*𝛼 <*

D∞

+ 1 和 （10.9α2

– 1.8a − 9.1） < D2，（3）

∞

mata 1和2，它允许Skellam的严格隐私核算

2

2𝑛𝜆

*噪声参数 λ 的算法 1 满足 （α， τ ）-RDP 且 τ =*

在涉及合成和子采样的应用中的噪声（例如。，

1.2𝛼 +1 · 𝑐 .

我们使用引理1将定理5推广到多维集合。

推论 1.假设每个参与者的数据点 xi 为

*D维并满足*

**算法3：Skellam混合的联合学习**

机械装置

**输入：训练记录的私有数据集 X = x1， . . . ， xn ;初始模型参数 θ ;安全聚合协议 .**

A

()

**参数：采样参数q;迭代次数 T ;噪声**

∑︁j =1

*𝑑*

|𝑥𝑖, 𝑗 |2 +

|𝑥𝑖, 𝑗 | − ⌊|𝑥𝑖, 𝑗 |⌋

– |𝑥𝑖, 𝑗 | − ⌊|𝑥𝑖, 𝑗 |⌋

2

≤ 𝑐,(4)

参数 λ;刻度参数γ ;裁剪阈值 C

和 δ∞;模量 m ∈ N.

**1 表示 h ∈ 1 . . .T do**

*和 ⌈|xi |⌉ ∞ ≤ δ∞。然后，只要α > 1 和 Δ∞满足方程 （3），*

*噪声参数 λ 的算法 2 满足 （α， τ ）-RDP，τ =*

**2服务器将当前模型参数 θ 共享给所有参与者。**

*𝑢.𝑎.𝑟*

使用泊松从所有参与者中均匀随机

1.2𝛼 +1 · 𝑐 .

2

2𝑛𝜆

**3B ←− {1， 2， . . . ， n }.// 抽样一组参与者**

在实践中，方程 （4） 和 xiΔ 中的约束可以通过裁剪来强制执行，正如我们在第 4 节中所解释的。剪切边界 Δ 的最大值由方程 （3） 计算得出。接下来，我们介绍算法 2 产生的效用保证，该算法来自推论 1。

L∞∞

⌈| |⌉ ∞ ≤ ∞

推论 2.假设每个参与者的数据点 xi 为

*d 维满足方程 （4）、xiΔ、α > 1 和 Δ 满足方程 （3）。然后，当满足α τ -RDP 时，算法 1 产生的误差为*

()

⌈| |⌉ ∞ ≤ ∞∞

使用速率 q 进行采样

**4对于我∈ B 做**

**5𝑔𝑖 ← ∇𝜃 (𝑟𝑖 ).// gradient computation**

**6ziAlgorithm 4 gi .// SMM 在参与者端**

←( )

**7𝑧¯ ← A( {𝑧𝑖 }𝑖∈𝐵).// secure aggregation**

**8𝑔¯∗ ← Algorithm 6(𝑧¯). // gradient sum retrieval by the server**

**9θU pdate θ， g ̄∗ .// 基于近似梯度和的模型更新**

←()

**输出：在X上学习的θ模型参数。**

*𝑛 𝑑*

2· 𝜏 +

*𝑖𝑗*

**算法4：扰动gradi的参与者过程-**

*𝐸𝑟𝑟*

= 1.2𝛼 + 1 𝑑𝑐

∑︁ ∑︁ |𝑥𝑖, 𝑗 | − ⌊|𝑥𝑖, 𝑗 |⌋ − |𝑥𝑖, 𝑗 | − ⌊|𝑥𝑖, 𝑗 |⌋ 2 .

我们简要评论一下推论2。我们定义 pi， j ：= |xi， j |−

=1 =1

M

树人

⌊|𝑥

*𝑖, 𝑗*

|⌋, which is the probability of increasing the absolute value

**输入：私有梯度 giRd**

**参数：噪声参数λ;刻度参数γ ;裁剪**

∈

*𝑥𝑖, 𝑗 by 1 for the 𝑖-th participant. Then, the overall error incurred by dSMM is:*

||

阈值 c 和 Δ ;模量 m N.

**公共随机性：一致随机符号向量**

∞∈

*𝐸𝑟𝑟 M =*

(1.2𝛼 + 1) · 𝑑𝑐 + ∑︁

∑︁（pi， j − pi， j ）.

*ξ ∈ {−1， +1}d .*

**1 𝑔𝑖 ← 𝐻𝑑 𝐷𝜉𝑔𝑖 .// random rotation, where**

*T= 和 /1 0 1 d d 是一个对角矩阵，其中*

*𝑖=1 𝑗 =1*

2𝜏

*𝑛*

*𝑑*

2

*H ∈ {−1/√d， +1 √d }d×d 是一个满足*

*H H ID ∈ {− ， ， + } ×ξ*

Err 的第一个术语可以被视为由于强制实施差分隐私而导致的错误。请注意，前导乘数 1.2α 1 2 仅比该方法略大（即，通过常数因子）

(+ )/

M

*\_；*

对角线

**2 𝑔𝑖 ← 𝛾 · 𝑔𝑖 .// scaling**

注入连续高斯噪声，α 2。第二个误差项是在参与者端进行的所有伯努利试验的总方差。此错误项可视为整数接近错误，即使不强制实施差分隐私也存在。

/

# 基于SKELLAM混合机制的联合学习

在本节中，我们将我们的Skellam混合机制（SMM）应用于分布式SGD的联合学习上的DP。我们假设参与者可以访问黑盒安全聚合协议，遵循[4，28]中的约定。训练过程在算法3中概述。在每次迭代中，服务器向所有参与者发布当前模型参数（算法3中的第2行）。然后，选择服务器不知道其身份的参与者的随机子集（第3行）。然后，所选子集中的每个参与者基于当前模型权重和她自己的数据计算梯度（第5行），并调用算法4进行梯度扰动（第6行）。之后，安全聚合协议计算随机选择的参与者的扰动梯度的总和（第7行）。最后，服务器检索扰动梯度和并更新模型（第8行和第9行）。我们

**3 𝑔𝑖 ← 𝑐𝑙𝑖𝑝 (𝑔𝑖 ) .// clip 𝑔𝑖 as in Algorithm 5**

**4 为 j ∈ 1 . . .d do**

**5pi， j = gi， j − ⌊gi， j ⌋.**

**6来自伯努利试验的样本 yi，j，成功概率 pi，k 。**

**7if 𝑦𝑖,𝑗 = 0 then**

**8𝑔𝑖∗, 𝑗 ← ⌊𝑔𝑖,𝑗 ⌋ + 𝑆𝑘 (𝜆, 𝜆).**

**9其他**

**10𝑔𝑖∗, 𝑗 ← ⌊𝑔𝑖,𝑗 ⌋ + 1 + 𝑆𝑘 (𝜆, 𝜆).**

**11zi，j ← gi∗， j v m.**

**输出：安全聚合协议的 zi ∈ Zd。**

*𝑚*

省略有关更新过程的其他详细信息（例如，学习率计划、权重衰减），因为它们不会影响一般框架或隐私保证。重复上述过程进行 T 迭代后，训练终止，服务端得到最终模型权重 θ 。

在下文中，我们将解释扰动梯度的参与者程序（流程图4）和重建扰动梯度和的服务器程序（流程图6）。每个参与者 i 首先使用 Walsh-Hadamard 矩阵 [26] 和共享的公共随机符号向量 ξ 随机旋转私有向量

**算法5：剪切梯度的参与者过程**

**输入：私人梯度 gi ∈ 路**

**参数：削波阈值 c 和 Δ∞。**

**1 𝑣𝑖 ← 0. // initialize the helper vector for clipping.**

**2 为 j ∈ 1..d do**

参数 M 可以看作是安全聚合协议的每维通信。虽然较大的m有助于保存有关噪声梯度的信息，但这样的m会增加通信成本，减慢聚合过程（特别是使用通信密集型安全聚合协议）

**3𝑣𝑖, 𝑗 = 𝑔𝑖, 𝑗**

|𝑔𝑖, 𝑗 |

· |𝑔𝑖,𝑗 |2 + |𝑔𝑖,𝑗 | − ⌊ |𝑔𝑖,𝑗 | ⌋ + ( |𝑔𝑖,𝑗 | − ⌊ |𝑔𝑖,𝑗 | ⌋)2 .

以及整体模型训练。当

参与者是具有计量互联网连接的移动设备。

将 GI 映射到 VI .

**4 𝑣𝑖 ← min(1, 𝑐 ) · 𝑣𝑖 .// L1 clip and re-scale**

*𝑣𝑖 1*

**5 为 j ∈ 1 . . .d do**

因此，在实践中，通常最好设置一个相对较小的m，例如，在我们的实验中设置28，这相当于每个维度一个字节的通信约束。

**6𝑔𝑖′, 𝑗 = ⌊√︁|𝑣𝑖,𝑘 | ⌋.// compute the integer part**

**7**

*𝑝′*

= 𝑦

//计算分数部分

*𝑖=1*

*𝑖*

2𝑔𝑖′, 𝑗 +1

展开模运算（算法6中的第1行）。在里面

*𝑖, 𝑗*

获得总和（.n z mod m）后，服务器首先

**8𝑔𝑖, 𝑗 ← 𝑣𝑖,**

· (𝑔′

+ p′

).//由两部分组成

特别是，{m/2， m/2 + 1， . . . ， m − 1} 中的值被映射

|𝑣𝑖, 𝑗 |

**9 为 j ∈ 1 . . .d do**

*𝑖, 𝑗*

*𝑖, 𝑗*

分别返回 {−m/2， −m/2 + 1， . . . ， −1};和值

**10𝑔𝑔𝑖, 𝑗**

*𝑖, 𝑗 ←*

|𝑔𝑖, 𝑗 |

·最小值（Δ∞

*, |𝑔𝑖,𝑗 |) .// L∞*

夹子

{0, 1, . . . , 𝑚/2 − 1} remain unchanged. This is because in Line 11

在算法 4 中，{−m/2， −m/2 + 1， . . . ， −1} 中的值映射到

{{C/+-}}-

输出：gi 裁剪的渐变。

**算法6：估计梯度和的服务器过程**

*𝑑*

**输入：私有向量 z ̄ = （.i∈B zi mod m） 通过安全∈ Z**

*𝑚*

聚合

**参数：噪声参数λ;刻度参数γ ;削波阈值 c 和 Δ ;模量 m N.**

∞∈

**公共随机性：一致随机符号向量**

*ξ ∈ {−1， +1}d .*

**1 映射 z ̄ ∈ Zd 到 z ̄′ ∈ [−m/2， m/2]d ∩ Zd .**

*𝑚*

**2 𝑔¯∗ ← 1 · 𝐷𝜉 𝐻𝑇 𝑧¯′.**

*120574；*

*𝑑*

*m 2， m 2 1， . . . ， m 1 ， 分别;和 0， 1， . . . ， m 2 中的值*

1被映射到它们自己。我们请读者参考[28]来详细讨论这个问题。然后，服务器反转在参与者侧执行的缩放（算法6中的第2行），获得梯度和的无偏估计。接下来，我们将解释算法5中概述的裁剪过程。回想一下，削波是DPSGD[2]中引入的一个标准步骤，用于将深度学习的灵敏度与DP结合起来。这项工作中的剪裁程序与DPSGD中的略有不同，后者剪裁梯度的2范数。这种差异是由于SMM的隐私保障不同（见定理5和

}

L

**输出：g ̄∗估计的梯度总和。**

在所有参与者和服务器之间（算法4中的第1行）

推论1）。回想一下，d维的隐私保证

SMM 依赖于输入数据 gi 的以下属性：

⌈|𝑔𝑖 |⌉ ≤ Δ∞, and

*𝑑*

也用于先前的解决方案[3，4，28]。结果梯度的每个维度都遵循具有方差的亚高斯分布

*𝑂 ( 𝑔𝑖 2/𝑑), where 𝑔𝑖 is the participant’s private gradient value.*

2

∑︁j =1

|𝑔𝑖, 𝑗 |2 +

|𝑔𝑖, 𝑗 | − ⌊|𝑔𝑖, 𝑗 |⌋

– |𝑔𝑖, 𝑗 | − ⌊|𝑔𝑖, 𝑗 |⌋

2

≤ 𝑐.

具体来说，当 d 很大时，每个维度都集中在 0 左右，例如，神经网络的数万个维度。从本质上讲，此操作使梯度变平，并限制了在计算梯度总和时溢出的概率。我们请读者参考[4，28]进行详细讨论。

−−

之后，参与者缩放旋转后的矢量并剪辑

缩放矢量（第2行和第3行）。我们将解释剪辑过程

因此，这需要不同的剪裁程序。第一个属性很容易强制实施。例如，对于 Δ = 1 和 xi = 1.9，我们只需将 xi 增加到 1。第二个属性的执行更加复杂，如下所述。对于每个参与者，我们首先构造一个辅助向量 vi。具体而言，vi 的每个维度计算如下：

很快（在算法 5 中概述）。然后，对于裁剪向量中的每个第 k 个坐标，参与者从

*𝑣𝑖, 𝑗*

= 𝑔𝑖, 𝑗

|𝑔𝑖, 𝑗 |

· |𝑔𝑖, 𝑗 |2

+ |𝑔𝑖,

| − ⌊|𝑔𝑖, 𝑗

|⌋ + (|𝑔𝑖,

| − ⌊|𝑔𝑖, 𝑗

|⌋)2 ,

成功概率的伯努利分布 gi，k − ⌊gi，k ⌋，其中

0

对于 j = 1 . . . ， d（算法 5 中的第 3 行）。为了完整性，我们定义

*𝑔𝑖,𝑘*

is the 𝑘-th element of the rotated vector 𝑔𝑖

（第5行和第6行

0 = 1。接下来，我们根据标准中的 L1 范数裁剪向量 vi

算法 4）。如果伯努利试验失败，参与者按照 Skellam 分布 Sk λ、λ 对噪声进行采样，并将结果转换为 gi，k（第 7 行和第 8 行）;否则，参与者将相同的结果转移到gi，k（第9行和第10行）。最后，参与者对噪声向量应用逐元素模运算（第 11 行）。从本质上讲，此步骤限制了参与者的输出

⌈⌉

⌊⌋

()

到 Zd ，并强制执行每个 log m 位通信约束

方式（算法5中的第4行）。最后，我们将修剪后的向量重新映射到其原始形式（算法5中的第5行至第8行），并通过Δ∞修剪向量的每个维度（算法5的第10行）。

## 隐私分析

在本节中，我们分析了算法3的隐私保障。观察算法3的每次迭代都可以看作是在运行

*米2*

维度，这两者都是安全聚合协议所要求的。

接下来，参与者共同计算他们的输出-

通过安全聚合协议放置有噪声的向量，并揭示

梯度的随机子集上的Skellam混合机制。这是因为模型共享（算法3的第2行）、梯度和重建（算法3中的第8行）或模型更新（算法3中排第9行）过程都不会产生任何额外的

服务器的总和 （.n zi mod m）。正如我们已经提到的，

∞

*𝑖=1*

隐私损失，因为更新后的模型可以通过

构造的扰动梯度和，该扰动梯度和又可以从安全聚合协议释放的扰动梯度总和中计算出来。此外，由于参与者的随机子集的身份对服务器来说是未知的，因此隐私保证得益于子采样的放大。（关于这个问题的详细讨论，我们请读者参考[28]。）因此，算法3的隐私保证通过将com定位（引理1）和放大（引理2）结果应用于SMM的隐私分析（推论1）来实现。算法3的隐私保证的正式声明如下。

定理6（算法3的隐私保证）。对于参数 q;抽样子集 B;迭代次数 T ;噪声参数 λ;和削波阈值 c 和 Δ∞，对于任何α > 1 和 Δ∞ 满足

*𝛼 < 2|𝐵|𝜆 + 1, and (10.9𝛼2 − 1.8𝛼 − 9.1) < 4|𝐵|𝜆 ,(5)*

四舍五入后，x 的范数从 0.01 √d 增加到 √d。换句话说，即使每个参与者的梯度向量都有一个

范数最多为01√，所有参与者的四舍五入梯度之和

2 灵敏度可能为 √d。这显着增加了

L

二级

·

*.·𝑑*

cpSGD实现差分隐私所需的噪声量，导致隐私和效用之间的不利权衡。

**分布式离散高斯（DDG）机制[28]。为了限制cpSGD的局限性，Kairouz等人[28]提出了DDG，这是一种利用离散高斯分布[11]而不是二项式分布来产生噪声的方法。特别地，离散高斯分布具有与连续高斯分布类似的PDF，但是是在整数域上定义的。使用离散高斯噪声的主要优点是它可以实现RDP，这使得为具有子采样的迭代算法推导DDG的严格隐私边界变得更加容易。**

D∞

*算法 3 满足 （α， τ ）-RDP*

1

*𝜏 = 𝑇 · 𝛼 − 1 ·*

）+∑花旗

2类似于cpSGD，DDG还假设输入为整数-

Δ∞值，因此，需要对非整数进行随机舍入。

为了缓解舍入引起的灵敏度增加，DDG

应用条件舍入方法如下。首先，给定一个

输入 x ∈ Rd，带有界 L2 范数 Δ2（否则为 DDG 削波

日志 （1 − q）

*α −1 （αq − q −*

*𝛼*

1

*l =2*

𝛼 (1

– 𝑞)

*𝛼 −𝑙 𝑞𝑙 𝑒 (𝑙 −1)𝜏 (𝑙 ) ! ,*

输入）和刻度参数γ，DDG 缩放输入 x 和

获得 γx 。之后，DDG 对 γx 执行随机四舍五入。如果缩放输入的舍入版本的 L2 范数更大

*𝑙*

*其中 τ （l） 定义为 τ （l） ：= 1.2l +1 · c ， 对于 l = 2， . . . . ， α.*

22 |𝐵 |𝜆

# 相关工作

如第1节所述，关于具有差分隐私的联合学习的现有工作主要考虑了可以使用实值噪声的非MPC设置。据我们所知，关于在联合学习中使用整数噪声来实现DP，先前只有四项研究[3，4，28，30]。在下文中，我们重新审视[3，4，28，30]中的解决方案，并将它们与我们的SMM进行比较。

**cpSGD[4]cpSGD让每个参与者将二项式噪声（即从独立的伯努利试验中得出的多个二进制值的总和）注入到她的离散梯度中，以满足DP。类似于连续域中的高斯噪声，二项式噪声也可以很容易地聚合，因为多个i.d.二项式值的总和也遵循二项式分布。这一特性简化了cpSGD的隐私推理，因为它使我们能够专注于所有参与者的梯度总和中的聚合二项式噪声，而无需分析每个参与者的二项式噪音分离率。然而，[4]中对cpSGD的隐私分析是基于**

*ε，δ -DP 而不是 RDP，这导致联邦学习的隐私边界相对宽松，因为很难得出确切的*

()

*ε，δ具有子采样（例如 SGD）的迭代算法的 -DP 保证。*

()

cpSGD 的另一个限制是它假定输入是整数值。对于任何非整数输入x，该方法需要x的随机舍入[4]，这通常会导致灵敏度的显著增加。例如，如果 x = 0.01， 0.01， . . . ， 0.01 Rd ，则 x 的每个维度四舍五入 1，概率为 0.01，四舍五入为 0，概率为 0.99。此方法可确保每个舍入值的期望值等于原始值，但舍入值

{} ∈

可能比原来的要大得多。特别是在最坏的情况下，当 x 的每个维度四舍五入为 1 时，L2

比

√︃𝛾 2Δ2 + 𝑑/4 + √︁2 log(1/𝛽)(𝛾 Δ2 + √𝑑/2),(6)

2

对于一些固定β（即将解释），然后 DDG 丢弃它并重新生成另一个随机舍入版本。重复该过程，直到满足上述要求。超参数

*从0到1的β控制了条件舍入过程中偏差和敏感性增加之间的权衡。要看到这一点，请注意，舍入值的期望值通常不等于原始值（因为不符合方程 （6） 中上述条件的舍入值被拒绝），这会对 DDG 输出的准确性产生不利影响。较小的β导致较低的偏置，但更高的灵敏度增加，这反过来又导致满足DP所需的更高噪声量，反之亦然。这种条件舍入方法可确保舍入运算不会产生*

灵敏度显著提高，但仍在增加

二级

*𝑂 √𝑑 . In addition, the conditional rounding operation introduces*

()

超参数β，在 DP 下很难调谐。[28] 的身份验证建议将β固定为 e−0.5，这在我们的

实验。

**斯凯拉姆机制[3]。在参考文献[3]中，Agarwal等人。建议从Skelam分布而不是离散Gauss-Ian分布中采集噪声。由于独立的skellam噪声之和仍在skellam分布之后（见第2.1节），因此分布式skellam噪声的隐私推理是直接的，与DDG不同。特别是，该文件表明，向整数添加Skelam分布式噪声也实现了RDP。然而，对于非整数输入，[3]中的skelam机制仍然需要[28]中引入的条件循环方法。因此，它的精度也足以满足灵敏度增加以及条件循环带来的偏差。**

**与SMM的比较。与上述方法相比，SMM的一个主要优点是它不依赖于额外的随机舍入[4]或条件舍入[3，28]步骤来处理非整数输入。相反，SMM直接接受任何**

*𝑥 R𝑑 as input, and outputs a noisy version 𝑥 ∗ of 𝑥 whose expecta-*

∈

等于x，而不会显著增加灵敏度。因此，SMM注入的噪音量更小，同时实现了与其竞争对手相同级别的隐私。因此，SMM能够获得比cpSGD [4]，DDG [28]和Skellam机制[3]更准确的结果，特别是在通信受限于低位宽的环境中。特别是，在这种情况下，量化粒度被设置为粗略级别（即，小尺度参数γ）以避免溢出。与量化梯度相比，这种粗略的量化粒度导致灵敏度折痕相对较大。因此，在这种低位宽设置下，由于舍入而导致的cpSGD、DDG和Skellam中的扰动噪声相当高，导致模型效用比SMM低得多。我们将在下一节中通过实验来验证这一说法。

此外，与DPSGD[2]相比，SMM仅涉及一个

附加 HYPER 参数：刻度参数 γ ，用于控制通信成本和效用之间的权衡。请注意，DPSGD 中不存在这种权衡，因为它是集中设置的解决方案。确定γ后，我们可以计算SMM的削波阈值c，因为对于某个常数，c = γ 2 Δ2

2

·

Δ2，对应于在 DPSGD [2] 中将 2 削波范数设置为 Δ2。此外，SMM 的削波边界 Δ 由公式 （5） 计算得出。相比之下，DDG [28] 和 Skellam 机制 [3] 都包含一个额外的超参数β。在这些算法中，参数β控制其条件舍入过程中偏差和灵敏度之间的权衡，如前所述

L∞∞

L

早些时候。选择不当β可能会对性能产生不利影响

## 分布式和估计

作为一个简单的应用，我们首先在给定私有d维输入数据集的情况下，评估so- lution SMM在第3.1节中描述的分布式和估计问题上的性能。根据[28]中的实验设置，我们生成了一个合成数据集，其中包含从d维2球体均匀采样的n = 100个数据点。我们将维度设置为 d = 65536，半径设置为 r = 1（即输入的 2 灵敏度为 1）。参与者在分布式DP下释放他们的噪音金额。我们报告所有维度的均方误差 （mse）。我们的评估使用ε，δ-DP（定义2）定义而不是RDP（定义3，因为ε，δ-DP是差分隐私的经典定义，竞争对手cpSGD支持前者，但不支持后者。我们将δ固定为 10−5，并将隐私参数ε从 1、2、3、4、5 改变。对于 DDG、Skellam 和 SMM，我们首先使用 RDP 计算隐私保证，然后将保证转换为

L

L

()

()

{}

*ε，使用引理 3 δ -DP（从 2 到 100 的整数中选择最佳 RDP 顺序）。*

()

对于DDG、Skellam、cpSGD和SMM，我们改变了通信-

每个维度的位宽为10、12、14、16、18。Corre公司-

{}

从 210、212、214、216、218 开始变化（见

{}

算法 4）。对于 m 等于 210、212、214、216 和 218，我们变化

刻度γ参数在 4、8、16、32、64、128、256、512 和

{} {} {} {}

1024, 2048 , respectively (see line 2 in Algorithm 4). For SMM, the clipping threshold 𝑐 is set to 𝛾 2𝑟 2 with 𝑟 = 1. Additionally, we compute the L∞ clipping bound for SMM using Eq. (3), based

{}

按照最佳RDP顺序。对于Skellam和DDG，L2剪裁

√︃√︁√

绑定设置为 Δ2 = γ 2r 2 + d/4 + 2 log（1/β）（γr + d/2），其中

*r = 1，d = 65536，β = exp（−0.5），如 [28] 所示。就 Skellam 和 DDG 绑定的 L 1 剪辑而言，我们有*

2

L

Δ1≤最小值

√ ·Δ2， Δ2 ， 在 [28] 之后。请注意，我们不

这些算法；同时，在差分隐私要求下，超参数调整是相当具有挑战性的。

我们还注意到，我们对SMM的理论分析与[3]中的理论分析有很大不同，这是由于SMM中使用的Skellam混合物分布与[3]中使用的Skelam分布之间存在固有差异。此外，即使对于输入中的特殊情况，SMM的隐私保证（见定理3和4）也与[3]中的Skellam机制不同，因为我们使用了与[3]中不同的证明技术。虽然[3]中使用的技术也是非平凡的，但正如我们在第3.3节中提到的，我们对整数输入的结果更干净，并且具有独立的兴趣。

# 实验

我们评估了SMM在分布式求和问题和两个基本机器学习任务上的性能。为了简便起见，所有实验都是使用TensorFlow库中离散高斯和Skellam的近似采样器进行的，这些采样器基于浮点近似。与精确采样器相比，近似采样器更快。我们在技术报告版本[6]的附录a中对此问题进行了详细讨论。

形成用于舍入梯度的实际1个削波步骤，因为2和1范数之间的上述关系对于所有整数值向量自动成立。类似地，对于cpSGD，范数为

以 2 个规范的 √ 倍为界，遵循其原始的实现-

L1级

L

双环磁性扫雷

站。我们还包括连续高斯，这是集中式DP设置的解决方案，作为强基线。

结果如图1所示。当通信位宽受到限制时（即，当m = 210、212、214时），SMM明显优于其所有竞争对手，如图1（a）、（b）、（c）、（f）、（g）和（h）所示。当m = 216且γ = 256时，SMM的性能与DDG和Skellam相当，如图1所示

当m和γ都很大时，SMM的性能略差于DDG和Skellam，后者获得的精度几乎与强基线连续高斯相同，如图1所示。

（i） ，（e）和（j）。最后，Skellam和DDG在所有设置下都具有相似的性能，并且cpSGD会产生相当高的误差（>104），并且超出了图中所示的误差范围。下面，我们简单介绍一下

解释产生上述结果的原因。

如第 5 节所述，由于随机舍入（以 cpSGD 为单位）或条件舍入（以 DDG 和 Skellam 为单位），分布式 DP 的现有解决方案会产生高灵敏度开销。更具体地说，灵敏度开销大约为每个维度 1，与缩放数据相比，这是不可忽略的，特别是当数据维度很大（例如，d = 65536）和量化时

粒度在小位宽（即小位宽（即小）下是粗略的（即小γ

*𝑚).这种灵敏度的提高导致cpSGD、DDG和Skellam的扰动更强，并解释了为什么SMM在小位宽的设置中表现最佳。随着位宽的增加*

*γ，与缩放数据相比，上述灵敏度开销可以忽略不计。因此，Skellam 和 DDG 的精度几乎与连续高斯相同。同时，SMM的性能略差于连续高斯，DDG和Skellam。这是因为根据推论2，SMM总是比连续高斯误差略大，其中有一个额外的因子1.2导致SMM的误差项。*

## 联盟学习

接下来，我们在两个经典基准数据集上评估了所提出的具有DP（算法3）的FL解决方案SMM的性能：MNIST [31]和时尚MNIST [49]，它们分别包含手写数字和服装的灰度图像。这两个数据集都代表 10 个类分类任务，其中包含 60， 000 条训练数据记录。我们将训练数据中的每个数据记录视为参与者。如前所述，我们的评估使用 ε δ -DP。我们将δ固定为 10−5，并将隐私参数ε从 1、2、3、4、5 改变。特别是，对于cpSGD，我们应用线性组合和高级组合[19]进行隐私会计，并在它们之间选择更强的保证。我们还将强大的中心模型DPSGD [2]作为基线。

()

{}

对于MNIST [31]和Fashion MNIST [49]，我们按照之前的工作[4]训练了一个具有完全连接层和ReLu激活的三层神经网络。我们将每层的神经元数量设置为 80，得到一个权重为 d = 63， 610 的模型。为

DDG，Skellam，cpSGD和SMM，我们改变通信约束m从26，28，210，其中m = 28对应于每个参数一个字节。对于每个 m，我们改变缩放参数γ m 32、m 16、m 8、m 4、m 2、m （参见算法 4 中的第 2 行）。为

{}

{ /////}

cpSGD、DDG、Skellam和集中式算法DPSGD，我们对原始实值使用相同的L2削波范数1

梯度。对于DDG和Skellam中的缩放梯度，我们设置L2

并且性能间隙随着比特宽度的增加而逐渐闭合。

具体来说，当m = 26时，SMM是唯一实现

在隐私参数ε、批量大小 B 和比例比γ的所有设置下都具有有意义的准确性（参见图 2（a）、（b） 和 （c） 以及图 3（a）、（b） 和 （c））。这是因为在DDG、Skellam和cpSGD中注入的噪声尺度太大，导致浮点数溢出，破坏了所得梯度和的效用。

| |

当m = 28（即每个参数一个字节）时，SMM还可以实现

与竞争对手相比，精度明显更高。具体而言，在图2（d）和3（d）中，我们将比例参数固定为γ = 64，批量大小为B = 240。当ε = 1时，由于浮点数溢出，DDG和Skellam的效用非常低，而SMM的效用要高得多，接近DPSGD（即差距小于10%）。随着ε的增加（表明隐私保护较弱），SMM与其竞争对手之间的性能差距变得不那么明显。这是因为随着更高的

| |

*𝜖, the required noise scale for the competitors becomes smaller, to the point that it no longer causes floating point number overflows. Nevertheless, there remains a noticable performance gap, since the noise scale of SMM is still significantly lower than that of its competitors. In particular, when 𝜖 = 3, the accuracy improvement of SMM over DDG and Skellam is around 6% and 10% for MNIST and Fashion MNIST, respectively, while the accuracy gap between SMM and the centralized baseline DPSGD is only around 3%.*

SMM 和集中式 DPSGD 算法之间存在性能差距，即使ε高达 5。并不是说在这一点上，满足DP所需的噪声不再主导扰动总量;相反，相对粗糙的量化粒度（即由小γ引起的）成为一个重要因素。正如我们稍后演示的那样，这种精度差距随着更大的位宽和/或大比例比γ逐渐缩小。

在图 2（e） 和 3（e） 中，我们固定了隐私参数 ε = 3 和刻度参数 γ = 64，并将批量大小 B 从 120、240、480、960 变化。SMM 是唯一在所有设置下始终如一地实现与 DPSGD 相当的精度的算法

{}

| |

对于MNIST和

剪裁绑定到√︃ 2 24√︁2对数1√2，带

*c D2 + d/*

+

（ /b）（c D2 +

*𝑑/ )*

|B|.特别是当 |B|=960，精度提高

Δ2 = 1，d = 65536，β = exp 0.5 。对于SMM，我们将削波阈值c设置为γ 2Δ2，Δ2 = 1，类似于其竞争对手。在

（-）

分别是时尚MNIST。最后，我们修复了隐私参数和批量大小，并改变比例参数γ如图 2（f） 所示

L条款

2

∞剪裁界

SMM，我们计算Δ∞

从…起

和3（f）。结果显示，与

等式（3）使用最优α序。我们还将批量大小 B 从 120、240、480、960 变化。该模型训练了 4 个 epoch，即当 B 等于 120、240、480 和 960 时，我们分别训练模型 2000、1000、500 和 250 轮。对于所有实验，我们使用学习率 η = 0.005 的 Adam 优化器 [29]。我们不会调整超参数以支持任何特定的解决方案，并且省略了超参数调整的其他实验，例如模型结构、学习率、裁剪范数、优化器、训练周期等。我们指出，我们的方法与现有的差分私有参数调谐技术[25，32，40]兼容，这是本文的一个正交主题。我们报告了 5 次运行的平均测试准确性。结果如图 2 和图 3 所示。

| |

{}

| |

总的来说，结果和分布和估计的结果一致，并得出了和以前类似的结论，即SMM比其竞争对手具有明显的性能优势

变化γ .我们还注意到，随着γ从8增加到256，SMM的精度先增加后减少。一方面，作为

*γ增加，梯度权重变得更加细粒度，包含更多信息，从而提高精度;另一方面，随着γ的增加，需要更大的噪声来满足DP。当γ = 256时，噪声梯度权重超过单字节通信约束，导致效用降级。Skellam和DDG可以观察到相同的性能模式。*

当通信位宽足够大时（例如，当m = 210时），我们观察到，虽然DDG和Skellam达到与DPSGD几乎相同的精度，但精度差距很小。

在SMM和DPSGD之间。例如，如图2（g）和图3（g）所示，MNIST的精度差距为0.5%和1%

以及SMM和DPSGD之间的时尚MNIST ε ≥2。此外，当|B|= 960

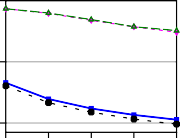
连续高斯SMMSkellamDDGcpSGD

104

102

英语

100



12345

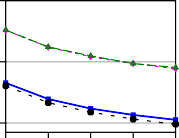
*120598；*

104

102

英语

100



12345

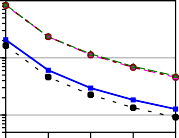
*120598；*

102

101

英语

100



12345

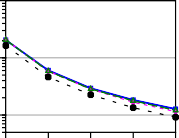
*120598；*

102

101

英语

100



12345

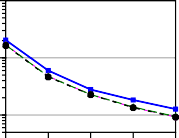
*120598；*

102

101

英语

100



12345

*120598；*

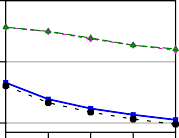
104

102

英语

100

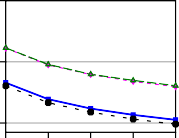
1. *m = 210，γ = 4。*



12345

*120598；*

1. *m = 210，γ = 8。*
2. *m = 212，γ = 16。*

104

102

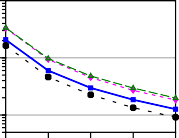
英语

100

12345

*120598；*

1. *m = 212，γ = 32。*
2. *m = 214，γ = 64。*

102

101

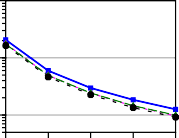
英语

100

12345

*120598；*

1. *m = 214，γ = 128。*
2. 216, 𝛾 = 256.

102

101

英语

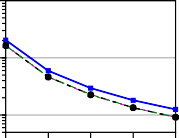
100

12345

*120598；*

**（一） m = 216，γ = 512。**

1. 218, 𝛾 = 1024.

102

101

英语

100

12345

*120598；*

**(j) 𝑚 = 218, 𝛾 = 2048.**

### 图 1：对具有不同隐私参数ε、尺度参数γ和通信约束 m 的合成数据的评估。

DPSGDSMMSKELAMDDGCPSGD

100

90

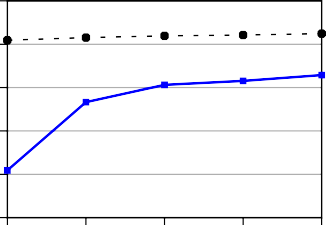
测试准确度%

80

70

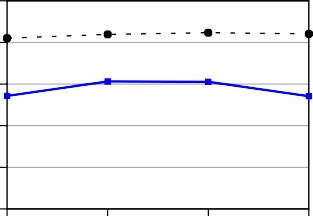
60

50



12345

100

90

测试准确度%

80

70

60

50

120240480960

100

90

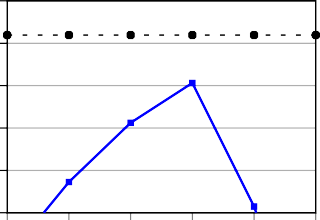
测试准确度%

80

70

60

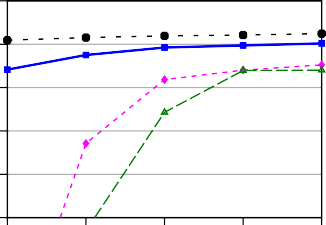
50



248163264

6𝜖6|𝐵 |6𝛾

**（a） m = 2 ， γ = 16， |B |= 240。不同的ε。**

100

90

测试准确度%

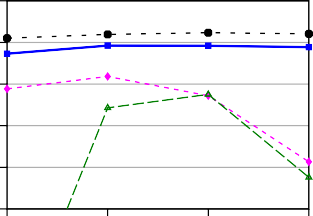
80

70

60

50 12345

**（b） m = 2 ， ε = 3， γ = 16。变化 |b |.**

100

90

测试准确度%

80

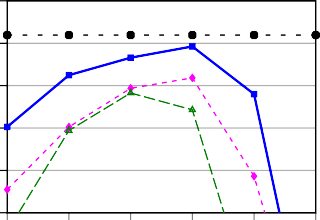
70

60

50

120240480960

**(c) 𝑚 = 2 , 𝜖 = 3, |𝐵 | = 240. Varying 𝛾 .**

100

90

测试准确度%

80

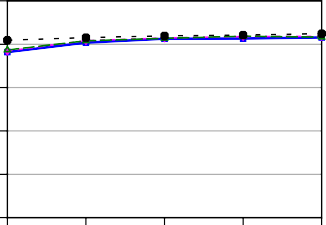
70

60

50 8163264128 256

8𝜖8|𝐵 |8𝛾

**（d） m = 2 ， γ = 64， |B |= 240。不同的ε。**

100

90

测试准确度%

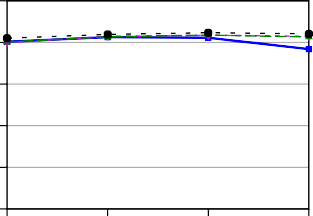
80

70

60

50 12345

**（e） m = 2 ， ε = 3， γ = 64。变化 |b |.**

100

90

测试准确度%

80

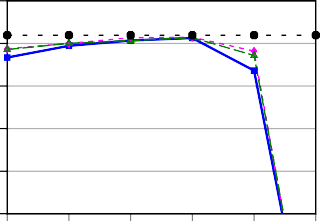
70

60

50

120240480960

**（f） m = 2 ， ε = 3， |B |= 240。变化γ .**

100

90

测试准确度%

80

70

60

503264128 256 512 1024

**（g） m = 210**

*120598；*

*， γ = 256， |B |= 240。不同的ε。*

**（高） 米 = 210**

|𝐵 |

*, 𝜖 = 3, 𝛾 = 256.变化 |b |.*

**（一） m = 210**

*120574；*

*， ε = 3， |B |= 240。变化γ .*

### 图 2：具有不同通信约束 m、隐私参数 ε、规模参数γ和批量大小 |B| 的 MNIST 上的评估。

DPSGDSMMSKELAMDDGCPSGD

90

80

测试准确度%

70

60

50

40 12345

90

80

测试准确度%

70

60

50

40

120240480960

90

80

测试准确度%

70

60

50

40 248163264

6𝜖6|𝐵 |6𝛾

**（a） m = 2 ， γ = 16， |B |= 240。不同的ε。**

90

80

测试准确度%

70

60

50

40 12345

**（b） m = 2 ， ε = 3， γ = 16。变化 |b |.**

90

80

测试准确度%

70

60

50

40

120240480960

**(c) 𝑚 = 2 , 𝜖 = 3, |𝐵 | = 240. Varying 𝛾 .**

90

80

测试准确度%

70

60

50

40 8163264128 256

8𝜖8|𝐵 |8𝛾

**（d） m = 2 ， γ = 64， |B |= 240。不同的ε。**

90

80

测试准确度%

70

60

50

40 12345

**（e） m = 2 ， ε = 3， γ = 64。变化 |b |.**

90

80

测试准确度%

70

60

50

40

120240480960

**（f） m = 2 ， ε = 3， |B |= 240。变化γ .**

90

80

测试准确度%

70

60

50

403264128 256 512 1024

**（g） m = 210**

*120598；*

*， γ = 256， |B |= 240。不同的ε。*

**（高） 米 = 210**

|𝐵 |

*, 𝜖 = 3, 𝛾 = 256.变化 |b |.*

**（一） m = 210**

*120574；*

*， ε = 3， |B |= 240。变化γ .*

### 图 3：具有不同通信约束 m、隐私参数ε、规模参数γ和批量大小 |B| 的时尚 MNIST 评估。

当图2（h）和（i）以及3（h）和（i）中的γ= 512时。总体而言，随着位宽的增加，DDG和Skellam的性能提升远比SMM显着，SMM在不同位宽下性能相对稳定。最后，在所有设置下，cpSGD的准确度都相当低（<20%），并且超出了图中所示的准确度范围。

# 结论

本文提出了Skellam混合机制（SMM），这是一种通过使用分布式随机梯度下降的基于MPC的联合学习过程建立的在机器学习模型上增强差分隐私的新解决方案。与现有的解决方案相比，SMM在不增加输入灵敏度的情况下实现了可组合和可扩展的隐私保证。在合成数据集和两个经典基准数据集以及各种实际设置上进行的大量实验表明，在限制性通信约束下，SMM比现有解决方案具有一致性和显著的准确性。

对于未来的工作，我们计划在相同的隐私保护水平下，进一步减少SMM隐私分析中的常数因子，以提高模型的实用性。另一个有希望的方向是打开MPC协议的黑匣子，并在考虑MPC协议细节的情况下进行仔细的隐私分析，这可能有助于进一步降低噪声水平，从而为联合学习带来更有利的隐私效用权衡。

# 鸣谢

这项工作得到了新加坡教育部（编号MOE2018-T2-2-091）、新加坡A\*STAR（编号A19E3b0099）和卡塔尔国家研究基金卡塔尔基金会（编号NPRP1C-1229-170007）的支持。本材料中表达的任何意见、发现、结论或建议均为作者的意见、发现和结论或建议，并不反映资助机构的意见。

# 参考信息

1. 马丁·阿巴迪、保罗·巴勒姆、陈建民、陈志峰、安迪·戴维斯、杰弗里·迪恩、马修·德文、桑杰·格马沃特、杰弗里·欧文、迈克尔·伊萨德、曼·朱纳思·库德勒、乔什·莱文伯格、拉贾特·蒙加、雪莉·摩尔、德里克·G。默里、贝诺伊特·施泰纳、保罗·塔克、维杰·瓦苏德万、皮特·沃登、马丁·威克、袁宇和郑晓强。2016年。张量流：大型机器学习系统。在奥斯迪。265-283。
2. 马丁·阿巴迪、安迪·朱、伊恩·J。古德费罗、H.布伦丹·麦克马汉、伊利亚·米罗诺夫、库纳尔·塔瓦尔和李章。2016年。以不同的隐私进行深入学习。在CCS中。308-318。
3. 说出阿加瓦尔、彼得·凯鲁兹和刘子瑜的名字。2021.不同私立联合学习的Skellam机制。在NeurIPS中。5052–5064。
4. Agarwal、Ananda Theertha Suresh、Felix Yu、Sanjiv Kumar以及H.Bren和McMahan的名字。2018.CpSGD：通信高效和差异私有分布式SGD。在NeurIPS中。7575-7586页。
5. Prabhanjan Ananth、Arka Rai Choudhuri、Aarushi Goel和Abhishek Jain。2018.具有诚实多数的圆形最优安全多方计算。在CRYPTO。395–424页。
6. 额尔古特宝、朱一正、萧孝奎、阴阳、蚌钦、本雅明、洪孟旦、钦密昂。2022.Skellam混合机制：一种具有差分隐私的联合学习新方法（技术报告）。检索日期：2022年5月15日https://drive.google.com/file/d/1k6HILAQC5\_是否查看？usp=共享
7. 阿莫斯·贝梅尔、阿里尔·加比松、尤瓦尔·伊沙伊、埃亚尔·库什莱维茨、西格德·梅尔德加德和阿纳特·帕斯金·切尔尼亚夫斯基。2014.非交互式安全多方计算。在CRYPTO。387–404之间。
8. James Henry Bell、Kallista A.Bonawitz、AdriàGascón、Tancrède Lepoint和Mar-iana Raykova。2020.具有（多）对数开销的安全单服务器聚合。在CCS中。1253年至1269年。
9. Keith Bonawitz、Vladimir Ivanov、Ben Kreuter、Antonio Marcedone、H.Bren-dan McMahan、Sarvar Patel、Daniel Ramage、Aaron Segal和Karn Seth。2017.保护隐私的机器学习的实用安全聚合。在CCS中。1175年至1191年。
10. Tom Brown、Benjamin Mann、Nick Ryder、Melanie Subbiah、Jared D Kaplan、Prafulla Dhariwal、Arvind Neelakantan、Pranav Shyam、Girish Sastry、Amanda Askell、Sandhini Agarwal、Ariel Herbert Voss、Gretchen Krueger、Tom Henighan、Rewon Child、Aditya Ramesh、Daniel Ziegler、Jeffrey Wu、Clemens Winter、Chris Hesse、Mark Chen、Eric Sigler、Mateusz Litwin、Scott Gray、Benjamine Chess、Jack Clark，Christopher Berner、Sam McCandlish、Alec Radford、Ilya Sutskever和Dario Amodei。2020年。语言模型是少数射击学习者。在NeurIPS中。1877年至1901年。
11. Clément L.Canonne、Gautam Kamath和Thomas Steinke。2020.差异隐私的离散高斯人。在Neurips中。
12. Nicholas Carlini、Chang Liu、Úlfar Erlingsson、Jernej Kos和Dawn Song。2019.秘密分享者：评估和测试神经网络中的意外记忆。在第267–284节中。
13. David Chaum、Ivan Damgård和Jeroen van de Graaf。1987.多方协商保证各方输入的隐私和结果的正确性。和CRYPTO。施普林格，87–119岁。
14. Albert Cheu、Adam D.Smith、Jonathan R.Ullman、David Zeber和Maxim Zhilyaev。2019.通过Shuffing实现分布式差异隐私。在EUROCRYPT。375–403之间。
15. Adam Coates、Brody Huval、Tao Wang、David J.Wu、Bryan Catanzaro和Andrew Y.Ng.2013。使用COTS HPC系统进行深度学习。在ICML中。1337年至1345年。
16. Ronald Cramer、Ivan Damgård和Jesper Buus Nielsen。2015。安全多方计算和秘密共享。剑桥大学出版社。
17. Jeffrey Dean、Greg Corrado、Rajat Monga、Kai Chen、Matthieu Devin、Quoc V.Le、Mark Z.Mao、Marc’Aurelio Ranzato、Andrew W.Senior、Paul A.Tucker、Ke Yang和Andrew Y.Ng.2012。大规模分布式深度网络。在NeurIPS中。1232–1240页。
18. 辛西娅·德沃克、弗兰克·麦克雪莉、科比·尼西姆和亚当·斯密。2006年，私人数据分析中校准噪声对灵敏度的影响。在TCC中。265–284页。
19. 辛西娅·德沃克和亚伦·罗斯。2014.差异隐私的算法基础。建立趋势理论。计算。科学。9，3-4（2014），211–407。
20. 辛西娅·德沃克、亚当·D·史密斯、托马斯·斯坦克、乔纳森·R·厄尔曼和萨利·P·瓦丹。2015年。跟踪量的强大可追溯性。在FOCS中。650–669之间。
21. Ulfar Erlingsson、Vitaly Feldman、Ilya Mironov、Ananth Raghunathan、Kunal Talwar和Abhradeep Thakurta。2019.Shuffing的放大：通过匿名从地方到中央的差异隐私。在SODA中。2468–2479页。
22. 维塔利·费尔德曼。2020.学习需要记忆吗？关于长尾的短篇小说。在STOC。954年至959年。
23. 罗萨里奥·根纳罗、尤瓦尔·伊沙伊、埃亚尔·库什莱维茨和塔尔·拉宾。2002.关于2轮安全多方计算。在CRYPTO。178–193之间。
24. Goryczka、李雄和Vaidy Sunderam。2013.具有差异隐私的安全多方聚合：比较研究。在EDBT/ICDT 2013联合研讨会上。155–163之间。
25. 阿努帕姆·古普塔、卡特里娜·利格特、弗兰克·麦克雪莉、亚伦·罗斯和库纳尔·塔瓦尔。

2010，差分私有组合优化。在SODA中。1106年至1125年。

1. A.赫达亚特和W.D.瓦利斯。1978.Hadamard矩阵及其应用。

*《统计年鉴》6，6（1978），1184-1238。*

1. Yuval Ishai、Eyal Kushilevitz和Anat Paskin。2010.具有最小交互的安全多方计算。在CRYPTO。577–594页。
2. 彼得·凯鲁兹、刘子瑜和托马斯·斯坦克。具有安全聚合的分布式离散高斯联合学习机制。在ICML中。5201–5212。
3. Diederik P.Kingma和Jimmy Ba.2015。亚当：一种随机光学方法。在ICLR中。
4. Antti Koskela、Joonas Jälkö、Lukas Prediger和Antti Honkela。2021.离散值机制和使用FFT的次采样高斯机制的紧微分隐私。和AISTATS。3358-3366中。
5. Y、 Lecun，L.Bottou，Y.本吉奥和P.哈夫纳。1998.应用于文件识别的基于梯度的学习。程序。IEEE 86，11（1998），2278-2324。
6. 刘敬成和库纳尔·塔尔瓦尔。2019.从私人候选人中进行私人选拔。在STOC。298–309页。
7. Ryan T.McDonald、Keith B.Hall和Gideon Mann。2010.结构化感知器的分布式训练策略。在HLT-NAACL中。456–464之间。
8. Brendan McMahan、Eider Moore、Daniel Ramage、Seth Hampson和Blaise Agüera y Arcas。2017.深度网络的通信高效学习-基于去中心化数据的工作。在AISTATS中。1273–1282页。
9. 卢卡·梅利斯、宋从政、埃米利亚诺·德·克里斯托法罗和维塔利·什马蒂科夫。2019.利用协作学习中的意外特征泄漏。在标准普尔691–706页。
10. 伊利亚·米罗诺夫。2017.Rényi差异隐私。在CSF中。263–275页。
11. 伊利亚·米罗诺夫、库纳尔·塔尔瓦尔和李章。2019.采样高斯机制的Rényi差分隐私。CoRR abs/1908.10530（2019）。
12. 米拉德·纳斯尔、雷萨·肖克里和阿米尔·胡曼萨德尔。2019.深度学习的全面隐私分析：被动和主动白盒推理攻击下的独立和联合学习。在标准普尔。739–753页。
13. 帕斯卡·派利尔。1999。基于复合度残差类的公钥密码系统。在EUROCRYPT。223–238页。
14. 尼古拉斯·帕普诺特和托马斯·斯坦克。2021.使用仁义差分隐私进行超参数调整。CoRR abs/2110.03620（2021）。
15. Apostolos Pyrgelis、Carmela Troncoso和Emiliano De Cristofaro。2018.敲门，谁在那里？聚合位置数据的成员关系推断。在NDSS中。
16. 雷萨·肖克里、马尔科·斯特罗纳蒂、宋从政和维塔利·什马蒂科夫。2017.针对机器学习模型的成员推断攻击。在标准普尔。3-18。
17. 宋从政，托马斯·里斯滕巴特，维塔利·什马蒂科夫。2017.记忆太多的机器学习模型。在CCS中。587–601页。
18. 宋从政和维塔利·什马蒂科夫。2019.文本生成模型中的审计数据来源。在KDD中。196–206之间。
19. 宋从政和维塔利·什马蒂科夫。2020.过度收入暴露敏感属性。在ICLR中。
20. 斯泰西·特鲁克斯、纳塔莉·巴拉卡尔多、阿里·安瓦尔、托马斯·斯坦克、海科·路德维希、张锐和周。2019.一种保护隐私的联合学习的混合方法-（扩展摘要）。知会光谱42，5（2019），356-357。
21. Filipp Valovich和Francesco Aldà。2017.基于格的密码学的计算差分隐私。在NuTMiC中。121-141中。
22. Tim van Erven和Peter Harremoës。2014年，Rényi分歧和Kullback-Leibler分歧。IEEE传输。Inf.理论60，7（2014），3797–3820。
23. 韩晓、卡希夫·拉苏尔和罗兰·沃尔格拉夫。2017.时尚MNIST：一个用于基准机器学习算法的新型图像数据集。CoRR abs/1708.07747（2017）。
24. 姚。1986.如何生成和交换秘密（扩展摘要）。在FOCS中。162–167页。
25. Samuel Yeom、Irene Giacomelli、Matt Fredrikson和Somesh Jha。2018.机器学习中的隐私风险：分析与过度拟合的联系。在CSF中。268–282之间。
26. Chiyuan Zhang、Samy Bengio、Moritz Hardt、Benjamin Recht和Oriol Vinyals。2021.理解深度学习（仍然）需要重新思考泛化。常见的ACM 64，3（2021），107-115。
27. 朱玉清和王玉祥。2019.Poission子采样Rényi差分隐私。在ICML中。7634–7642。