中国密码学发展报告 2022

中国密码学会 编

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

本书是中国密码学会组织编写的重要文献，从 2008 年起每年出版，集中反映上一年国内外密码学的最新研究进展。本书分为三部分：第一部分对 2021 年国际三大密码年会发表的论文进行了分类整理，分为对称密码、理论基础、公钥密码、（后）量子密码、安全协议和应用密码六大类，每大类又细分为多个研究方向，分别邀请长期从事该方向研究工作的专家进行归纳、总结和点评；第二部分是为本书撰写的关于密码学与机器学习相互作用的 5 篇综述论文，主要涉及机器学习安全与隐私的前沿进展、机器学习的安全外包、基于机器学习的侧信道分析、联邦学习的隐私保护技术及自动化密码分析技术等；第三部分介绍 2021 年中国密码学会 4 篇优秀博士学位论文的主要成果。

图书在版编目（CIP）数据

中国密码学发展报告. 2022 / 中国密码学会编. --北京 ： 电子工业出版社, 2023.6

ISBN 978-7-121-45698-5

Ⅰ. ①中… Ⅱ. ①中… Ⅲ. ①密码术－研究报告－中国－2022 Ⅳ. ①TN918.1

中国国家版本馆 CIP 数据核字(2023)第 098527 号

责任编辑：张　冉
特约编辑：武瑞敏
印　　刷：
装　　订：
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×980　1/16　印张：22.5　字数：576 千字
版　　次：2023 年 6 月第 1 版
印　　次：2023 年 6 月第 1 次印刷
定　　价：118.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

本书咨询联系方式：（010）88254439，zhangran@phei.com.cn，微信号：yingxianglibook。

前　言

《中国密码学发展报告》是中国密码学会组织编写的重要学术文献，从 2008 年起每年都出版，集中反映上一年国内外密码学最新研究进展，是密码学研究者和密码科技工作者重要的学术参考资料，对推动我国密码理论和技术创新具有重要意义。

《中国密码学发展报告 2022》分为三部分。第一部分介绍 2021 年国际三大密码年会（Crypto、Eurocrypt、Asiacrypt）的主要成果；第二部分介绍密码学与机器学习的相互作用；第三部分介绍 2021 年中国密码学会优秀博士学位论文的主要成果。

国际三大密码年会反映了世界先进密码技术和研究的发展方向。整理归纳国际三大密码年会的论文，有助于我国密码工作者掌握最前沿的密码技术，开展具有中国特色的密码学术研究与技术创新。本书对 2021 年度国际三大密码年会发表的论文进行分类整理，分为对称密码、理论基础、公钥密码、（后）量子密码、安全协议和应用密码六大类，每大类又细分为多个研究方向，分别邀请长期从事该方向研究工作的专家进行归纳、总结和点评。希望有助于读者准确掌握每篇论文的研究成果和各研究方向的发展前沿。

以机器学习为代表的人工智能技术在现代信息技术领域有着重要应用。密码学与机器学习相互促进、相互作用，是当前信息安全领域的一个研究热点。本书组织相关领域专家撰写了关于密码学与机器学习相互作用的 5 篇综述论文，主要涉及机器学习安全与隐私的前沿进展、机器学习的安全外包、基于机器学习的侧信道分析、联邦学

习的隐私保护技术及自动化密码分析技术等。希望有助于读者全面、系统地了解密码学与机器学习的相互作用的研究进展。

2021 年中国密码学会评选了 4 篇优秀博士学位论文，论文的作者分别是王森鹏、沈耀斌、温云华、魏春艳，他们均在自己的博士研究领域做出了成绩。本书第三部分介绍了这 4 位优秀博士学位论文的主要思想和观点。

高飞教授、张方国教授和禹勇教授分别负责本书第一、二、三部分的组稿，感谢三位老师的精心组织，感谢执笔的各位老师们；感谢中国密码学会王瑶副秘书长、毕雪妮同志在协调本书的组稿、编辑和出版过程中付出的诸多心力。

中国密码学会副理事长

中国密码学会学术工作委员会主任委员

戚文峰

2022 年 7 月 23 日

目　录

第一部分　2021 年国际三大密码年会主要成果

第二部分　密码学与机器学习的相互作用

第三部分　2021 年中国密码学会优秀博士学位论文主要成果

第一部分
2021 年国际三大密码年会主要成果

本书第一部分首先介绍 2021 年在国际三大密码年会上发表的论文。国际三大密码年会代表了世界先进密码技术和研究方向的发展，通过整理归纳三大国际会议论文的研究方向及其主要贡献，有助于我国密码研究工作者掌握最前沿的密码技术，促进我国密码事业健康快速发展。

2021 年 Crypto 会议共收到 430 篇投稿论文，录用 103 篇，录用率约为 24.0%。Eurocrypt 会议共收到 400 篇投稿论文，录用 78 篇，录用率为 19.5%。Asiacrypt 会议共收到 341 篇投稿论文，录用 95 篇，录用率约为 27.9%。

在 Crypto 录用的文章中，中国高校和科研院所为第一单位的共 14 篇，作者分别来自上海交通大学（沈耀斌、王磊、谷大武、郁昱、刘晗林、赵铄曜、刘炆灵、胡震恺、韩帅、刘胜利）、清华大学（陈一镭、董晓阳、华佳良、王小云、丁津泰）、暨南大学（翁健）、密码科学技术重点实验室（张江）、电子科技大学（王煜宇）、山东大学（陈宇）、华东师范大学（刘富康）、北京航空航天大学（冯翰文）、中国科学院信息工程研究所（孙思维、胡磊、刘美成、鲁小娟、林东岱）、北京工业大学（李铮）、国防科技大学（王毅、陈荣茂、王宝生）、福建师范大学（黄欣沂）、北京应用数学研究院（陶成东）、香港大学（Tsz Hon Yuen、区文浩、T.H. Hubert Chan、Ting Wen）、台湾“中研院”（Kai-Min Chung），研究方向主要集中在对称密码、理论基础、公钥密码和安全协议。此外，国内学者参与完成的论文有 6 篇，海外华人华侨参与完成的有 20 篇。

在 Eurocrypt 录用的文章中，中国高校和科研院所为第一单位的共 7 篇，作者分别来自中国科学院信息工程研究所（于伟、孙思维、许军、史丹萍）、教育部密码技术与信息安全重点实验室（许广午）、清华大学（丁津泰、董晓阳、王小云）、国防科技大学（刘韵雯、李超）、陕西师范大学（来齐齐）、中国科学院数学与系统科学研究院（潘彦斌）、北京工业大学（李铮）、台湾“中研院”（Bo-Yin Yang、Kai-Min Chung）、台湾交通大学（Yu-Hsuan Huang）、台湾大学（Tai-Ning Liao），研究方向主要集中在理论基础、对称密码和公钥密码。此外，国内学者参与完成的论文有 2 篇，海外华人华侨参与完成的有 11 篇。

在 Asiacrypt 录用的文章中，中国高校和科研院所为第一单位的共 13 篇，作者分别来自暨南大学（颜俊、赖俊祚、翁健）、中国科学院软件研究所（江浩东、张振峰）、数学工程与先进计算国家重点实验室（马智）、信息工程大学（叶晨东、田甜）、西安电子科技大学（潘静、陈晓峰）、中山大学（张方国）、中国科学院信息工程研究所（邓燚、马顺利、张心轩、汪海龙、张志宇、胡磊、肖禹亭、张锐、马晖）、中国科学院大学（孙思维）、上海市隐私保护计算重点实验室（宋旭阳、谢翔）、清华大学（董晓阳、魏聪明、王小云、丁津泰）、国防科技大学（王毅、陈荣茂、王宝生）、福建师范大学（黄欣沂、 宁建廷）、山东大学（胡凯、王美琴、王庆菊、陈宇）、上海交通大学（韩帅、刘胜利、谷大武）、中国地质大学（覃悦、程池、张晓涵）、中国科学院数学与系统科学研究院（潘彦斌）、电子科技大学（王煜宇）、香

港大学（杨如鹏）、鹏城实验室（黄正安），研究方向主要集中在对称密码和公钥密码上。此外，国内学者参与完成的论文有 2 篇，海外华人华侨参与完成的有 8 篇。

国际三大密码年会录用的论文基本涵盖了密码技术的主要研究领域。本书第一部分把录用论文的研究领域分为六大类：对称密码、理论基础、公钥密码、（后）量子密码、安全协议和应用密码，每个大类包含多个具体的研究点。根据每个研究点，相应研究领域的专家对会议录用论文进行逐篇归纳、总结和点评，希望有助于我国密码工作者理解和掌握每篇会议论文的知识点。

对称密码

1　对称密码构建

密码分析的作用是评估密码算法的安全界，衡量密码算法抵抗各种攻击的能力。针对对称密码算法，目前有许多密码分析方法，其中最重要的是差分密码分析、线性密码分析、积分攻击及其上述攻击的变形。密码攻击通常可以分为两部分：区分器的构造和密钥恢复攻击两部分。区分器的构造一般通过寻找密码算法中的非随机特性，如内部状态之间的线性相关性，或者给定某些特定的输入差分时，输出差分会出现的异常分布等。密钥恢复攻击通常是针对所构造的区分器之前和之后的轮函数，并利用这种非随机特性来（部分）恢复密钥比特。从本质上来说，攻击者从区分器的外部猜测一些密钥信息，并通过加密或解密若干轮函数后，检查是否存在非随机特性。若统计数据的概率分布符合预期，则认为猜测的密钥可能是正确的。对密钥恢复攻击的优化是密码分析的一个重要环节。

文献[1]提出了一种通用技术来改进分组密码迭代攻击的密钥猜测步骤，将猜测的密钥比特数减少到能保证输出信息仍然确定的严格最小值，避免了对全密钥比特不必要的猜测。具体来说，作者通过定义和研究密码算法中 S 盒的一些新属性，并将它们表示为一种特殊类型的决策树，该决策树对于找到细粒度的猜测策略至关重要。为此，作者首先将一个 S 盒转换为二叉决策树，然后指出所有重要的优化都可作为该树的属性表示出来，其中最重要的属性之一是叶子的数量。因此，找到具有最少叶子的树表示可以直接优化密钥恢复攻击。作者提出并实现了有效找到此类树的一个算法，并将其应用于多个密码算法，改进了对 Noekeon 算法、GIFT 算法和 RECTANGLE 算法目前最好的密钥恢复攻击。

在下一步的研究工作中，如何处理更大规模，且有更多的输入比特的函数是一个值得研究的问题。对树进行启发式的搜索可能会产生更大功能的树，从而使分析不止针对一个 S 盒甚至不止针对一轮密码函数有效，但这有可能需要许多自动化方法以及手工改进方法。此外，对于n比特的任意（平衡）布尔函数，给出最小叶子数的非平凡上界是一个非常值得研究的公开问题。

一般来说，如果 r 轮密码算法不存在某种攻击的区分器，那么一般可以认为任何的n轮（$n > r$）密码算法几乎不存在这类攻击的区分器。然而，目前没有公开正式的论据或证据支

持这一基于经验的结论。积分攻击属于分组密码算法的经典攻击方法，然而，要证明某个给定密码算法是否能够抵抗这些攻击是非常困难的。

在文献[2]中，作者构建了完整的理论来说明任何非平凡子空间上的函数之和与密钥是相关的，并且复合函数的代数次数不小于原函数代数次数，即$\mathrm{minDeg}(S_k \circ F) \geqslant \mathrm{minDeg}(F)$。如果 r 轮分组密码算法可以抵抗积分攻击，那么n轮（$n \geqslant r$）的密码算法自然可抵抗积分攻击。对于分组密码E，为了验证$E_k(x + k_0)$是否具有积分抵抗力，需要探索足够多的三子集划分路径，这可使用 Hebborn 等人在 2020 年亚密会上提出的路径扩展技术来实现。作者还提出了一种使用等效 S 盒的方法来提高搜索路径的效率。文献[2]最后利用给出的方法，对分组密码 CRAFT、PRESENT、SIMON、SIMECK、SKINNY-64、GIFT-64 等算法进行了应用，并给出了针对这些密码算法积分区分器紧的界限。分析结果表明，除 PRESENT 算法和 GIFT-64 算法外，作者给出这些密码算法的安全界与目前已知的最优结果一致。

如何在更自动化以及更少手工优化的情况下，将该方法更广泛地应用于更多的分组密码是一个值得研究的问题，特别是针对 PRESENT 算法和 GIFT-64 算法，作者给出的两个密码算法的安全界分别是 13 轮和 12 轮，但目前已知的最优积分区分器的轮数是 9 轮和 10 轮，存在这种差距的原因是一个值得研究的开放性问题。

积分密码分析是一个有力的对称密码分析工具，利用可分性来寻找积分区分器是目前比较有效的方法之一。可分性可以应用于任意由布尔电路实现的密码算法（例如，由 AND 门和 XOR 门构成）。然而，由于基于比特的可分性理论尚不完善，信息在传播过程中会丢失。考虑到密码算法中的较大规模部件，如 S 盒和扩散层的线性变换，可使得信息的损失减缓。例如，Zhang 和 Rijmen 的研究成果表明，通过基本的 COPY-and-XOR 操作实现的线性映射传播的可分性是不完善的。

文献[3]阐述了对传统基于比特级可分性的新理论和实践的新见解。作者引入了一个新的框架来描述二子集比特可分性理论和实际应用的理解，其中包括公式化已有的二子集比特可分性，刻画二子集比特可分性的凸性，以及列举所有无效、有效和冗余的可分路径。此外，在这个新框架的基础上，还给出了与 Carlet 图形指标界线（代数次数）的关系。区别于 Xiang 等人在 2016 年亚密会上提出的方法，根据这个二子集比特可分性的新框架，从$F_2{}^n$到$F_2{}^m$的任意映射的可分路径可以通过计算一些仅与该映射的真值表相关的集合简单而精确地得到。在基于工具的自动积分攻击中，这种新方法可用于高效准确地建模 S 盒的可分路径，特别是对于大规模的 S 盒，如 AES 算法的 S 盒或 LED 算法的超级 S 盒。基于 SMT/SAT、MILP 对 LED 算法的超级盒，使用传统的方法进行建模几乎是不现实的，但这种新方法使之成为可能。此外，作者提出了一些算法，以提高计算效率。例如，利用给定集的奇偶性集合，可进一步获得可分路径，这对实际应用非常有效。最后，作者对 LED 算法的超级 S 盒进行了建模，并

验证了 8 轮 LED 不存在积分区分器，这解决了 LED 算法是否存在超过 7 轮的积分区分器的公开问题。

AES 算法由 Daemen 和 Rijmen 于 1999 年设计，是目前应用最广泛的分组密码算法之一。AES 算法的安全性从其问世以来，密码学者不断研究其抵抗现有攻击及新型攻击的能力，并评估其安全冗余。AES 有 3 个版本，具有不同的密钥规模和不同的轮数：AES-128 的总轮数是 10 轮，AES-192 的总轮数是 12 轮，AES-256 的总轮数是 14 轮。经过 20 多年对 AES 算法的密码分析，密码学者对该算法的安全性有很强的信心：在单密钥模型下，针对 AES-128 的最优攻击仅达到 7 轮。到目前为止，已知的最优攻击是不可能差分攻击或中间相遇攻击的。然而，在考虑到密钥扩展方案时，在相关密钥模型下，该算法可能会产生新的安全性问题，这是 AES 安全性最弱的部分。例如，在 2010 年 FSE 会议上，Gilbert 和 Peyrin 将两轮 AES 算法的轮函数复合，利用超级 S 盒性质，改进了基于 AES 方案的几个密码分析结果。随后，Gilbert[4]给出了全轮 AES-128 的首个已知密钥攻击。

在文献[5]中，Leurent 等人使用不变子空间攻击技术研究了 AES 密钥扩展方案的等价表示，基于两个主要性质得到了新的密码分析结果。首先，作者发现迭代奇数轮的密钥扩展方案存在一个短周期的置换函数，即使在迭代多轮之后，AES 密钥扩展方案也不能被视为一个随机置换。其次，作者所给出的 AES 密钥扩展方案中的特性，使密码分析者更容易联合从第一轮子密钥到最后一轮子密钥的信息，从而可改进现有的密钥恢复攻击。现有的密码分析方法也尝试过利用密钥扩展方案来降低复杂度。例如，Dunkelman 等人提出的密钥桥技术（key bridging），而文献[5]给出了之前从未利用过的非线性关系，该特性可为进一步利用 AES 算法的密钥扩展方案特点进行密码分析提供新的思路。

在密码设计方面，密码学者也取得了一些新的突破。SPN 结构是一个经典的分组密码算法结构，该结构以非线性层（S 盒）和线性层（P 盒）组成的轮函数进行多次迭代。AES、SKINNY 等许多现代密码算法都采用了 SPN 结构。宽轨迹策略是 AES 设计者评估算法抵抗差分密码分析、线性密码分析时提出的，这种策略使得基于 SPN 结构的密码算法具有易于证明的安全性下界，从而为密码算法抵抗最常见的统计密码分析提供安全保障。在 CHES 2013 年会议上，Gerard 等人提出了部分 SPN 结构（PSPN），即 S 盒在每轮中只应用于部分中间状态。这种结构在各种场景中都具有明显的性能优势，被用于分组密码 Zorro 算法和 LowMC 算法的设计中。然而，宽轨迹策略并不适用于 PSPN 结构，因此现有的分析方法对这些结构的安全性证明是不够的。例如，一些研究结果显示，LowMC 算法初始版本的安全性明显低于设计者所声称的安全界[6-7]。在 2020 年欧密会上，Grassi 等人将经典的 SPN 结构和 PSPN 结构相结合，提出了 HADES 设计策略。在 HADES 设计策略中，一层 PSPN 轮的中间层被两层 SPN 轮包围。HADES 关于统计分析的安全性评估只使用 SPN 结构的轮变换，而忽略了 PSPN 结构的轮变换。

在文献[8]中，Keller 等人的研究显示，线性变换层中 MDS 矩阵的选择将显著影响 HADES 设计提供的安全级别。如果线性变换中的 MDS 矩阵选择恰当，那么 HADES 方案抵抗差分、线性攻击的能力将明显高于设计者所宣称的安全界。另外，MDS 矩阵如果选取得不恰当，那么会导致非常大的不变子空间通过整个中间层，且不产生任何活跃 S 盒。Keller 等人利用基于 HADES 设计的两个实例 Starkad 算法和 Poseidon 算法上印证了上述分析结果。例如，对于 Starkad 算法，在某些参数下所构造的不变子空间可对哈希函数发起原像攻击，从而破坏 HADES 声明的安全性。为了给出 PSPN 结构和 HADES 设计策略更精确的安全性证明，仍有一些问题亟待解决。由于线性变换 MDS 矩阵的选取对 HADES 设计的安全性影响明显，因此选择恰当的 MDS 矩阵，将是推进 PSPN 结构设计与优化的关键。HADES 设计者指出，目前在 PSPN 结构设计中有效的密码分析工具非常稀缺，因此开发新的工具或改进现有的工具都将促进对 PSPN 的深入研究。

安全多方计算（MPC）是现代密码学的重要组成部分和热门研究领域，其主要目标是在一个互不信任的分布式网络中，两个或多个用户能够在不泄露各自隐私数据的前提下合作计算某个约定函数并获得计算结果。作为隐私保护的一个最重要的方法和手段，安全多方计算已成为分布式计算的关键基础性理论与技术，在民用和国防各个领域都有着重要的应用。密码学者针对对称密码体制在安全多方计算协议方面的应用进行了大量研究工作，取得了一些实质性的进展，如 AES 算法或 SHA-256 算法，它们在安全多方计算的实践中得到了广泛应用。然而，这些对称密码算法本身在设计之初并没有考虑针对安全多方计算协议的优化和评估。因此，密码学者开始研究和设计针对 MPC 友好的密码部件。

在文献[9]中，Dinur 等人利用$\mathbb{Z}_2$和$\mathbb{Z}_3$上的线性函数交替使用来设计安全多方计算的快速协议，提出了一套新的、简单的 MPC 友好的设计方案。作者基于交替模技术，设计了新的对称密码，包括候选的单向函数、伪随机数生成器和弱伪随机函数，提供了一个通用的方法来评估不同安全多方计算模型中的模交替体制的安全性，并给出了候选算法对各种应用的有效性。该项工作为下一步的研究工作提供了两个方向：一个是设计具有恶意安全性的安全多方计算协议，同时使其额外成本最小化，文献[10]中给出的最新技术有助于实现这一目标；另一个是基于交替模范式设计和分析其他对称密码体制，如哈希函数、强伪随机函数和分组密码。事实上，Boneh 等人在 2018 年 TCC 会议上已经提出了一个强伪随机函数的候选设计，但分析其具体安全性还有待进一步研究。

密码协议和算法中的乘法数目，通常称为乘法复杂性。低乘法复杂度在安全多方计算（MPC）、完全同态加密（FHE）和零知识证明（ZK）等方案中有重要的应用。低乘法复杂度的密码算法分为两大类：第一类是指最小化$\mathbb{F}_2$中乘法的密码算法，如 Flip、Keyvrium、LowMC 和 Rasta 等算法；第二类是由二元域或素域中构造的密码算法，如 MiMC、GMiMC、Jarvis、Hades、Poseidon 和 Vision and Rescue 算法。这些专门设计的密码算法在这些应用中的性能大

大优于“传统”设计的密码算法。在提供足够的安全级别的同时，寻求最小化乘法复杂度是探索和创新设计策略的一个重要方面。目前，所有新提出的最小化乘法复杂度的方案都只从基于幂映射的 S 盒入手，此外，这些方案大多依赖于复杂和资源密集型的线性层来实现低乘法复杂度。

在文献[11]中，Dobraunig 等人提出了一种加密方案 Ciminion，它在使用非常轻量的线性层的同时，最大限度地减少了二元域或素域的乘法。与其他方案相比，Ciminion 方案利用 Toffoli 门来改进非线性混淆。将 Ciminion 的 MPC 消耗与公开文献中其他设计方案进行对比，结果表明 Ciminion 和 HadesMiMC 是 MPC 应用中最具竞争力的两种低乘数复杂度的设计方案。值得注意的是，用于二元域或素域等密码方案的设计灵感源于 20 世纪 90 年代提出的抵抗差分攻击设计策略的启发。例如，MiMC 类似于 Knudsen-Nyberg 算法，Jarvis 的设计灵感来自 Rijndael 的设计，而 Hades、Vision and Rescue 算法的灵感来自 Shark。此外，在二元域或素域中设计密码算法很容易抵抗差分密码分析。然而，代数攻击似乎是这些算法的主要威胁。因此，有必要探索不同的设计策略来提高这类算法对代数攻击的抵抗能力。

本节作者：刘国强、孙兵（国防科技大学）

参考文献

[1] BROLL M, CANALE F, FLÓREZ-GUTIÉRREZ A, et al. Generic framework for key-guessing improvements[C]. ASIACRYPT 2021(1): 453-483.

[2] HEBBORN P, LAMBIN B, LEANDER G, et al. Strong and tight security guarantees against integral distinguishers[C]. ASIACRYPT 2021(1): 362-391.

[3] UDOVENKO A. Convexity of division property transitions: theory, algorithms and compact models[C]. ASIACRYPT 2021(1): 332-361.

[4] GILBERT H. A simplified representation of AES[C]. ASIACRYPT 2014(1): 200-222.

[5] LEURENT G, Pernot C. New representations of the AES key schedule[C]. EUROCRYPT 2021(1): 54-84.

[6] BAR-ON A, DINUR I, DUNKELMAN O, et al. Cryptanalysis of SP networks with partial non-linear layers[C]. EUROCRYPT 2015(1): 315-342.

[7] DINUR I, LIU Y, MEIER W, et al. Optimized interpolation attacks on LowMC[C]. ASIACRYPT 2015(2): 535-560.

[8] KELLER N, ROSEMARIN A. Mind the middle layer: The HADES design strategy revisited[C]. EUROCRYPT 2021(1): 35-63.

[9] DINUR I, GOLDFEDER S, HALEVI T, et al. MPC-Friendly Symmetric Cryptography

from Alternating Moduli: Candidates, Protocols, and Applications[C]. CRYPTO 2021(4): 517-547.

[10] BONEH D, BOYLE E, CORRIGAN-GIBBS H, et al. Zero-knowledge proofs on secret-shared data via fully linear PCPs[C]. CRYPTO 2019(3): 67-97.

[11] DOBRAUNIG C, GRASSI L, GUINET A, et al. Ciminion: symmetric encryption based on toffoli-gates over large finite fields[C]. EUROCRYPT 2021(2): 3-34.

2 哈希函数/伪随机函数/随机数/消息验证码

哈希函数的一大应用是作为随机数生成器（RNG），在将内部状态置为初始种子后，对于每个输入X，以$R \leftarrow \mathrm{Hash}(R, X)$的方式产生随机数同时更新内部状态。然而由于哈希函数的计算过程较为繁杂，在实际应用中，特别是在频繁中断的操作系统中，每生成一个随机数都调用一次哈希函数是完全不可行的。这就要求设计者在设计时优先考虑效率的随机数生成器，即实用型随机数生成器，其中一种应用广泛的设计是基于移位与异或的实用型随机数生成器：对于每个输入X，生成器先将n比特内部状态R循环右移α位，再将其与X异或后输出，整个过程可以记为$R \leftarrow \mathrm{rot}_{\alpha,n}(R) \oplus X$。Windows 系统就采用了这种实用型随机数生成器，其中的参数设置为$\alpha = 5$与$n = 32$（32 位系统）和$\alpha = 19$与$n = 64$（64 位系统）。

文献[1]从输入熵积累的角度探讨了上述实用型随机数生成器的最佳参数设置问题。由于操作系统的寻址机制，寄存器的高位部分往往为全零，此时实用型随机数生成器的输入X将呈现一种“单峰分布”，即概率密度函数先单调递增后单调递减的分布。在这种分布下，X的熵将集中在低位部分（记其长度为k），而高位部分则倾向于一个固定值。经过i步移位异或的迭代积累，X的熵将扩散至输出R的至多ik位；换言之，基于移位与异或的实用型随机数生成器至少要经过$\lceil n/k \rceil$次迭代，输入熵才能覆盖每位输出，从而产生一个无序的随机数。以此为标准，文献[1]论证了$n = 32$时（32 位系统）$\alpha \in \{5,7,9\}$和$n = 64$时（64 位系统）$\alpha \in \{15,19,23,27\}$均能满足对于每个$1 \leqslant k < n$，输入熵积累的输出覆盖步数均接近$\lceil n/k \rceil$，并通过比较实验给出了最终的最佳参数设置$\alpha = 7$与$n = 32$（32 位系统）和$\alpha = 19$与$n = 64$（64 位系统）。此外，文献[1]还论证了“基于比特反序与异或的实用型随机数生成器$R \leftarrow \mathrm{rot}_n(R) \oplus X$是一种能在不同单峰分布下最小化输出覆盖步数的设计”。

哈希函数在实际应用中可以用来避免口令的明文存储，即只在数据库中存储口令pw的哈希值$H(pw)$，这样即使数据泄露，由于哈希函数的不可逆性，攻击者也无法直接获知口令明文。但如果攻击者收集到大量的口令哈希值数据，那么可以在不针对哈希函数本身进行破解的情况下，借助如彩虹表等一些预处理辅助数据进行口令恢复攻击，因此出现了如在哈希函数的输入中加盐等防御这类攻击的手段。文献[2]对这种带有辅助数据的多实例攻击进行了建模与理论分析。

在文献[2]中首先假设哈希函数是理想的随机函数，针对不同情况设计了相应的交互式攻击模型。其定义了两种不同的敌手：对可猜测性进行攻击的敌手可以进行规定次数的猜测询问，并即时得到是否正确的回答；而对可恢复性进行攻击的敌手只能在最后给出所有的回答。对于辅助数据，文中也分析了两种不同的模型：在 BF-RO（Bit-Fixing Random-Oracle）模型中，随机函数的一部分大小固定的输入输出已经由攻击者确定，其他部分是均匀随机的，预处理时只能根据这部分确定的数据而不能访问整个函数H；而在 AI-RO（Auxiliary-Input Random-Oracle）模型中，预处理时攻击者可以访问函数H，并给出大小固定的预处理辅助数据。对于哈希函数的使用，文献[2]也设计了多种不同的算法模型，如加盐的算法和非自适应的算法。除此之外，由于在实际情况中一部分明文可能泄露，作者在模型中加入了一个特殊的函数可以直接获取一个明文，并将其调用次数也作为成功率的自变量。因此，在文献[2]中通过假设一种模型下的敌手存在来构造另一种敌手，从而通过归约的方式将不同模型相互联系起来，而最简单的直接猜测模型其成功率上界是直接由明文数据的平均信息熵所决定的，由此即可得出各种情况下的攻击成功率上界。

在 2021 年美密会上，针对弱伪随机函数（WPRF）的最小可实现复杂度问题，Boyle 等人在文献[3]中首次提出了一个基于 F2 稀疏多变量多项式（深度为 2 的 XOR、AND 电路）和一个基于奇偶校验门 AC0 电路（AC0◦MOD2）的 WPRF，并证明了所提出的 WPRF 对于代数攻击和线性攻击的抵抗性。该工作还指出基于 AC0[MOD2]的 WPRF 与 LPN（Learning Parity with Noise）假设的关联性。AC0◦MOD2 与 AC0[MOD2]的区别在于，AC0◦MOD2 的奇偶校验门只能位于电路的输入层。对于 F2 稀疏多变量多项式的 WPRF：$F_K(x)=\oplus_{i=1}^{D}\oplus_{j=1}^{w}\wedge_{l=1}^{i}x_{i,j,l}[K_{i,j,l}]$，它有着高有理度（常数项为 0 或 1 的情况下，乘以一个次数最小的非零多项式使得结果全 0 对应的多项式次数），从而可以抗代数攻击。对于基于 AC0◦MOD2 的 WPRF，Boyle 等人将 ABGKR 的 TRIBES 函数 $g(x)=\vee_{i=1}^{\lambda}\wedge_{j=1}^{\log_2\lambda}x_{ij}$ 改为 $g(x)=\vee_{i=1}^{\lambda}\wedge_{j=1}^{\lambda}\vee_{k=1}^{w}x_{ijk}$，其中$w=\lfloor\log_2\lambda-\log_2(\log_2\lambda)\rfloor$，$m=\lambda^2 w$。添加了一层 OR 之后，有理度从$O(\log_2\lambda)$提升到$\lambda$，从而抵抗现有的代数攻击。作为后量子安全的候选困难问题之一，LPN 假设无论在学习理论还是在密码学中都是研究的热点。该工作的重要意义在于它首次给出了基于 Sparse F2-polynomials 和 AC0◦MOD2 的具有亚指数安全性的低复杂度 WPRF，并提出了一种新的 LPN 假设——LPN with Simple Deterministic Noise，即对于随机的 x 和秘密的 s，AC0[MOD2]中的函数$gk(x)$使得$(x,<x,s>\oplus gk(x))$和随机对(x,y)是不可区分的。

格式保留加密（Format-Preserving Encryption，FPE）是一类能保证明/密文格式相同的加密方式。FPE 以密钥、明文数据以及与格式有关的参数作为输入，输出与明文格式（包括数据域和数据长度）完全相同的密文，输入的明文格式可以记为基数为a且长度为l的域Z_a^l。目前大部分 FPE 算法均基于 Feistel 结构，如 FF1 和 FF3 等算法已成为 FPE 标准。也存在一些基于 SPN 结构设计的 FPE 算法，但这些算法在输入明文格式上缺乏灵活性，或者不是纯粹

的 SPN 结构。

文献[4]提出了一种新的真正基于 SPN 结构的 FPE 模式 FAST（Format-preserving Addition Substitution Transformation），支持灵活的数据格式（支持基数$a \geq 4$且数据长度$l \geq 2$）。FAST 接受长度为L的主密钥K，以输入明文Z_a^l的每位为单元进行操作，加/解密过程中使用到m个Z_a上的 S 盒，来自一个事先准备好的 S 盒池。加/解密前，主密钥通过伪随机函数 PRF 分别生成长度为L_1和L_2的子密钥K_{SEQ}和K_{S}，K_{S}通过伪随机发生器 PRNG2 生成包含m个随机生成的 S 盒的 S 盒池，K_{SEQ}通过伪随机发生器 PRNG1 生成加密过程中每层加密使用 S 盒的顺序 SEQ，其中 PRF、PRNG1 和 PRNG2 的选择是开放的。FAST 核心加密过程中每层加密为 SPN 结构，非线性操作涉及Z_a上的模加、模减和 S 盒代换，共迭代n层。Durak 等人基于 AES 算法的加密模式构造 PRF、PRNG1 和 PRNG2 对 FAST 进行了软件实现，并将其处理效率与 FF1 和 FF3 标准算法进行了比较。实验发现，FAST 相比 FF1 和 FF3 标准有更好的加/解密效率和更广的数据格式适用范围。

Durak 等人对 FAST 进行了安全性分析，该分析基于强安全性模型，即在敌手有多个可用目标（未知主密钥）、可任意选择算法参数（包括a、l、m、n等）和任选明/密文进行加密和解密询问的条件下算法与理想 FPE 的不可区分性。Durak 等人证明了在使用的 PRNG1 和 PRNG2 为安全的伪随机发生器以及 PRF 为伪随机函数的条件下，FAST 的强安全性可以约减为其核心加/解密函数已知 S 盒池的弱安全性模型，即敌手在单个主密钥、已知所使用的 S 盒池和未知 S 盒使用顺序 SEQ 的条件下 FAST 与随机置换的不可区分性。Durak 等人给出了 FAST 在不同a、l以及期望达到的比特安全强度s的使用场景下算法的推荐参数，其中$L = s$，$L_1 = L_2 = 2s$，$m = 256$，$n \sim l^{1.5}$。通过设置足够大且合适的n，可以保证 FAST 达到足够的安全强度并拥有较快的处理速度。Durak 等人也给出了 FAST 在量子意义下的安全性分析和推荐参数设置。

在保持数据的某些属性的同时，压缩数据是计算机科学中最基本的任务之一。

局部敏感哈希函数允许将数据点x和y独立地压缩成短摘要$h(x)$和$h(y)$，这样散列值可以用来检查原始点在某种距离度量标准下是否互为近邻。Boyle、LaVigne 和 Vaikuntanathan 给出了鲁棒的属性保留哈希函数（PPH 函数）的研究。PPH 函数结合了哈希函数抗碰撞的安全保证和局部敏感哈希函数的功能。

文献[5]基于 q-strong 双线性离散对数假设，为汉明距离预测这一任务构建了一个鲁棒的 PPH 函数。该哈希函数可以将两个大规模输入x和y压缩成长度为$O(st)$的短摘要（s是安全参数），并且能够使得用户在不访问x、y的情况下，可以准确判断出x、y的汉明距离是否大于等于t。Fleischhacker 等人进一步将 PPH 函数应用范围推广到基于字母表的字符串（如字母数字序列），达成计算字符串中相同位置但是字符不同的数量。基于相同的假设，Fleischhacker 等人构建了一个鲁棒的 PPH 函数，用于集合交集预测。该哈希函数可以将两个包含n个元素

的集合压缩成长度为$O(st)$的段摘要（s是安全参数），用户在不访问集合的情况下，可以准确判断出集合交集包含的元素数目是否大于等于$n-t$。最后文献[5]展示了如何把第二个属性保留哈希函数扩展至多个集合。Fleischhacker 等人提出了一个巧妙的编码方法，将汉明距离预测这一任务转换成集合交集预测。

哈希函数压缩模型一般是采用对单个对称密码函数元件的结构性调用来组成的。对于$m+s-\text{to}-s-\text{bit}$的压缩函数，如果调用了$r$次$n+c-\text{to}-c-\text{bit}$的压缩元件，那么在$2^{\frac{nr+cr-m}{r+1}}$次询问中可以找到一次碰撞。传统的 Merkle Tree 模式并没有达到哈希函数模型的理论抗碰撞下界，因此还存在一定的效率提升空间。

文献[6]构建了一种新的哈希函数压缩模式——ABR 模式。ABR 模式同样为树状压缩模式，但是在节点之间加入更多的消息块。在保持与 Merkle Tree 模式相同的密码原件调用次数的情况下能够压缩更多的消息块。Merkle Tree 模式可以通过调用2^l-1次$2n-\text{to}-n-\text{bit}$的密码元件来压缩$2^l$块消息，ABR 模式在同样的调用次数下可以额外压缩$2^{l-1}-1$块消息，在渐进上达到了最佳的理论抗碰撞下界，但是最基本的 ABR 模式并不能实现抗差分性。文献[6]同时提出了改进的 ABR$^+$模式，ABR$^+$模式是对两个 ABR 树（不需要同等大小）的合并。改进后的 ABR$^+$模式实现了抗差分性，且每次压缩时仅比基本的 ABR 模式少压缩 1 个消息块。

ABR 与 ABR$^+$模式除了哈希函数压缩，同样可用于并行处理器或多核机器对软件更新，图像与视频的认证，大文件系统的完整性检验、长期存档、存储认证、目录分配、种子系统以及匿名密码货币等应用场景。

本节作者：于红波（清华大学）

参考文献

[1] DODIS Y, GUO Siyao, STEPHENS-DAVIDOWITZ N, et al. No Time to Hash: On Super-Efficient Entropy Accumulation[C]. CRYPTO 2021(4): 548-576.

[2] FARSHIM P, TESSARO S. Password Hashing and Preprocessing[C]. EUROCRYPT 2021(2): 64-91.

[3] BOYLE E, COUTEAU G, GILBOA N, et al. Low-Complexity Weak Pseudorandom Functions in AC0[MOD2][C]. CRYPTO 2021(4): 487-516.

[4] BETÜL DURAK F, HORST H, HORST M, et al. FAST: Secure and High Performance Format-Preserving Encryption and Tokenization[C]. ASIACRYPT 2021(3): 465-489.

[5] FLEISCHHACKER N, SIMKIN M. Robust Property-Preserving Hash Functions for

Hamming Distance and More[C]. EUROCRYPT 2021(3): 311-337.

[6] ANREEVA E, BHATTACHARYYA R, ROY A. Compactness of Hashing Modes and Efficiency Beyond Merkle Tree[C]. EUROCRYPT 2021(2): 92-123.

3 对称密码证明

密码学中有一类可证明安全性的研究，自 20 世纪 80 年代中后期开始出现[1]，并于 20 世纪 90 年代全面铺开[2-4]。这一类研究涉及的密码系统范围很广，涵盖 MAC、哈希函数、对称加密工作模式与认证加密、分组密码结构等。它们的共性在于所使用的假设与证明方法。

（1）这类证明往往不依赖基于数论的难解问题，而仅假设密码系统中存在所谓的随机函数（Random Function）或随机置换（Random Permutation）。这些随机函数/置换或对攻击者保密[1,4]或对攻击者公开[2-3]；前一种场景往往是由伪随机函数/置换建构的密码系统变换所得的，后一种场景则是使用哈希函数、（无密钥的）密码学置换的密码系统在随机谕言/随机置换模型（Random Oracle/ Permutation Model）中的理论模型。

（2）这类证明的关键步骤都是分析随机系统（Random System）中的随机事件与概率，运用的方法来自组合数学、概率论、随机过程等数学分支，得出的结论多是信息论意义下的不可区分性（Information Theoretic Indistinguishability）[1]。

本节收录了 2021 年国际三大密码年会所发表的此类成果。笔者将这些成果进一步划分为哈希函数证明、伪随机函数/伪随机置换证明、MAC 与认证加密证明、量子 Q2 模型中的安全性证明 4 类。这显然没有遵从统一的分类规则，但笔者认为有必要将“量子 Q2 模型中的安全性证明”单独分类，以特别突出此类前沿研究。

3.1 哈希函数证明

2019 年，NIST 启动了轻量级密码标准征集 LWC [5]。为达到更高的效率，LWC 提出将密码杂凑函数和认证加密整合到同一算法家族中，从而使得未来的用户能够用相同或相似的密码函数同时完成两种最重要的对称密码功能，减少实现开销。

实用密码杂凑函数和认证加密往往是基于输入长度固定的密码学置换（Cryptographic Permutations）或（可调）分组密码的工作模式，可以将这一类输入长度固定的密码函数称为块函数（Block Function）。因此，设计基于相同块函数的杂凑与认证加密模式，从而实现上述整合，在技术上是可行的。

另外，用同一个块函数设计认证加密和杂凑函数是有难度的。块函数如果“太小”，搭建的杂凑函数内部状态小，安全性可能不够：NIST 要求杂凑函数至少保证 112 比特安全性，而基于 128 比特（可调）分组密码设计的 Merkle-Damgard 迭代杂凑函数只能达到 128/2=64 比

特（抗碰撞）安全性。块函数如果“太大”，对于轻量级认证加密而言又有浪费之嫌。为克服这个矛盾，NIST 最终轮候选算法 Romulus 采用了所谓双分组长度（Double-Block-Length，DBL）杂凑函数模式，利用分组长度为n比特的可调分组密码构造摘要长度为$2n$比特（固定输入长度）DBL 压缩函数（相对于分组长度翻倍，因此称为“双分组长度”杂凑函数），再用 Merkle-Damgard 结构迭代形成杂凑函数。

具体而言，Romulus 家族[6]的杂凑函数 Romulus-H 采用了 Hirose[7]在 FSE 2006 上提出的 DBL 压缩函数构造，并用带置换增强的 Merkle-Damgard 结构将其转换为迭代杂凑函数。在安全性方面，日本学者 Naito[8]为此方案证明了不可分辨性（Indifferentiability），具体安全性为接近最优的$n-\log_2 n$比特；在性能方面，Hirose 压缩函数优势是每对（可调）分组密码调用使用的密钥都相同，从而节约了一半密钥编排的开销。

但 Romulus-H 需要$3n+k$比特状态（假设所用可调分组密码分组长度为n比特、密钥长度为k比特），在这方面并不是国际范围内最优的：Özen 和 Stam[9]提出将 Hirose DBL 压缩函数中的前向反馈（Feed-Forward）删除，而依靠迭代弥补安全性的损失，从而将杂凑函数内部状态减少到$2n+k$比特；受此启发，Naito[10]提出了另一种不用前向反馈，仅需$2n+k$比特状态的杂凑函数。在状态开销方面，这两种方案才是国际最优的。

Naito 等人针对这一问题，取得了新的突破性进展[11]，提出了新的 DBL 杂凑模式 EXEX-NI 和 EXEX-I，其所需内部状态都降到了$n+k$比特：注意这已经是最优了，因为实现分组长度为n比特、密钥长度为k比特的（可调）分组密码本身就需要$n+k$比特状态（言下之意，这两个杂凑函数在实现的时候没有“额外”的存储需求）。

具体而言，Naito 等人的新方案的核心思路是将传统 DBL 压缩函数并行调用 2 次（可调）分组密码的设计修改为（以某种方式）串行调用 2 次，实现了充分利用每一比特存储的目标。注意，并行改串行是常见的减少状态规模和硬件实现面积的技巧，在轻量级分组密码实现中广泛应用。串行的缺点是更大的计算延时，但是当密码实现被内存占用、硬件面积等指标“卡脖子”时，牺牲一些计算延时显得无关紧要了。

将此新型 DBL 压缩函数代入朴素 Merkle-Damgard 迭代结构，Naito 等人[11]提出了 EXEX-I 迭代杂凑模式，并证明了$n-\log_2 n$比特抗碰撞安全性，这几乎是摘要长度为$2n$比特的杂凑函数所能达到的最优抗碰撞安全性。朴素 Merkle-Damgard 迭代结构不具备不可分辨性安全性，且易受长度扩展攻击（Length-Extension Attack），为此，Naito 等人基于 Coron 等人[12]提出的类 NMAC 方案，在 EXEX-I 基础上添加了 2 次（可调）分组密码调用，提出了 EXEX-NI 迭代杂凑模式，并证明了与 Romulus-H 类似的$n-\log_2 n$比特不可分辨性安全性。同样基于可调分组密码 SKINNY-128-384 实现时，EXEX-NI 和 EXEX-I 的电路面积比 Romulus-H 减少约 2000 GE，总面积不到 Romulus-H 的 70%，表明其在理论与实用层面都实现了瞩目的效果。

3.2 伪随机函数/伪随机置换证明

伪随机函数和伪随机置换的概念分别由文献[13]与文献[1]提出，现已成为密码学核心原语，在绝大部分现代密码系统中都发挥了关键作用。伪随机置换的“结构”性质比伪随机函数更多，因此理论密码学研究通常用单向函数[13]、格困难问题[14]等假设构造伪随机函数，再经 Luby-Rackoff 结构[1]（基于伪随机函数的 Feistel 网络）构造出伪随机置换。但在密码学实践中，经受了密码分析考验的、高度可靠的分组密码，如 AES、SM4 等，就可以看作伪随机置换。因此，分析分组密码理论模型安全性时，通常试图证明模型是伪随机置换。

密钥交替密码或迭代 Even-Mansour 密码正是这样一个分组密码理论模型。它可以看作 AES 等 SP 网络密码中代换层与置换层合并、建模为一个完整置换所得的模型。若假设这些轮函数是随机置换，则可以为建立的密钥交替密码模型证明伪随机性等安全性。这种理论模型与 SP 网络密码的实际设计有非常大的距离，但这一系列理论研究被认为是为 SP 网络密码提供了一定的合理性支持，因此仍然得到了认可和广泛关注。

若使用 t 个不同的 n 比特随机置换、t+1 个相互独立的轮密钥搭建起上述 t 轮密钥交替密码模型，则可以证明，其伪随机安全性是 tn/（t+1）比特。若在 t 轮运算中使用相同的 n 比特随机置换、t+1 个相互独立的轮密钥，则 Chen 等人为 t=2 的情形证明了 2n/3 比特伪随机安全性[15]，与轮函数/随机置换不同的情形安全性一致。

Tessaro 和 Zhang[16]推动改进了这一结论，证明了如下结论：可以从 t–1 个相互独立、均匀分布的“主密钥”导出 t+1 个轮密钥，用于t轮密钥交替密码模型，仍可保证 tn/（t+1）比特伪随机安全性。密钥导出函数所需满足的条件较为复杂，但举例而言，给定 2 个“主密钥”(k_0, k_1)，3 轮密钥交替密码可以依次使用k_0、k_0、k_1、k_1作为 4 轮密钥，仍可保证 3n/4 比特安全性；给定 3 个“主密钥”(k_0, k_1, k_2)，4 轮密钥交替密码可以依次使用k_0、k_1、k_2、$k_0 \oplus k_1$、$k_1 \oplus k_2$作为 5 轮密钥，仍可保证 4n/5 比特安全性。

Tessaro 和 Zhang 进一步证明了如下结论：当轮数$t \geqslant 8$时，可以从 t–2 个相互独立、均匀分布的“主密钥”导出 t+1 轮密钥，用于t轮密钥交替密码模型，仍可保证 tn/（t+1）比特伪随机安全性。可以基于次数为 t–3 的随机多项式构造所需的密钥导出函数。

推进该研究的主要难点在于分析、求解密钥交替密码模型中的随机事件概率，而这些往往又和图论等组合学问题相关。一个著名的图论工具是所谓的“Sum-Capture 引理”，其考虑随机 Cayley 图中边的个数，最早由 Babai 在傅里叶分析的课程课件中证明，后来被发现与基于置换的杂凑函数[17]、基于 4 轮 Feistel 网络的数字签名[18]、2 轮密钥交替密码安全性[15]等众多对称密码安全性证明问题有关。为改进多轮密钥交替密码结论，Tessaro 和 Zhang 对“Sum-Capture 引理”在形式上和结论上都进行了推广和改进。

如前所述，从理论密码角度来看，安全分组密码是伪随机置换。但某些工作模式如 CTR、

Wegman-Carter MAC、GCM 等，当所用的块函数是伪随机函数时，能够保证更高的安全性；而当块函数是n比特伪随机置换/分组密码时，由于伪随机函数/伪随机置换差异的存在（二者的差异是伪随机置换永远不会将不同的输入映射到相同的输出，但伪随机函数可以），因此收集大量不同输入-输出对即可判断它们是伪随机函数还是置换的输出，反而受到生日攻击的影响[19-20]，往往将其在同一密码下所能处理的数据量限制到$2^{n/2}$分组。使用轻量级分组密码时，n往往为 64，系统所能处理的数据量急剧减为$O(2^{32})$，是严重限制。

但在对称密码设计领域，设计安全的置换或分组密码的理论相对成熟，设计伪随机函数则一直是摸黑行路（最近 5 年才出现了一些先行者[21-22]）。Bellare[23]等人最早注意到这个问题，并提出用安全分组密码构造具有“超生日界”可证明安全性的伪随机函数问题。他们提出的一个著名解决方案是所谓 XORP 结构，它基于一个n比特随机置换P，定义一个定义域为$\{0,1\}^{n-1}$、值域为$\{0,1\}^n$的伪随机函数$\mathrm{XORP}(x) = P(0||x) \oplus P(1||x)$。其推广形式为基于$k$个$n$比特随机置换$P_1,\cdots,P_k$的伪随机函数构造$\mathrm{XORP}[k] = P_1(x) \oplus \cdots \oplus P_k(x)$。Lucks[24]是第一位为$\mathrm{XORP}[k]$证明“超生日界”安全性的学者，后续工作更是将其可证明安全界改进到了最优n比特[25]。

传统的安全性定义专注于讨论特定一个密钥的密码系统之特征。在信息时代，密码系统投入实际运用时，会导致网络上的海量用户近乎同时运行同一密码系统，而密钥则是各用户独立选取的。这种场景称为“多用户”场景（Multi-user Setting），而专注于一个特定密钥的传统场景称为“单用户”场景（Single-user Setting）。一个好的伪随机函数F不仅应在传统“单用户”场景中保证足够高的实际安全性，更应该在“多用户”场景中保证高安全性，即使用u个独立选取的密钥$K_1,\cdots,K_u$时，$F_{K_1},\cdots,F_{K_u}$和u个相互独立的随机函数是不可区分的。

印度学者 Bhattacharya、Nandi 考虑了以下场景[26]：使用u个相互独立的随机置换$P_1,\cdots,P_u$，构成u个伪随机函数$F_1(x) = P_1(x||00) \oplus P_1(x||01) \oplus P_1(x||10),\cdots,F_u(x) = P_u(x||00) \oplus P_u(x||01) \oplus P_u(x||10)$，作为$u$个用户使用的伪随机函数。攻击者对每个用户进行（最多）$q_{\max} \leqslant 2^n/12$次询问，合计$u \times q_{\max}$次询问。在这样的条件下，他们证明了XORP[3]的多用户 PRF 安全界为$\frac{20\sqrt{uq_{\max}}}{2^n}$。这个结论表明，XORP[3]可以供$O(2^n)$数目的用户同时使用，而攻击者对每个用户都进行$O(2^n)$次询问后仍不能被攻破。而在传统“单用户”场景中，这个结论表明询问次数达到$q = O(2^n)$的攻击者成功的概率仍是可忽略函数$O(1/2^{n/2})$。

Bhattacharya、Nandi 进一步提出了一个效率稍高的方案$\mathrm{XORP}'[3]: \{0,1\}^{n-2} \to \{0,1\}^{2n}$：$\mathrm{XORP}'[3](x) = P(x||000) \oplus P(x||001) \oplus P(x||010) \,||\, P(x||000) \oplus P(x||101) \oplus P(x||110)$。该方案用 5 次随机置换调用产生$2n$比特伪随机输出，显然比并行执行 2 次XORP[3]更高效。但是它的安全性和XORP[3]很接近：在传统“单用户”场景中，进行q次询问的攻击者之攻击

优势上界为$\frac{5\sqrt{q}}{2^n}+\frac{256q}{2^{2n}}+\frac{8192q}{2^{3n/2}}$。本文不仅提出了新的伪随机函数构造方案，更为XORP这一经典方案提供了新的认识。

安全多方计算是近 10 年里最重要的密码研究方向之一。在理论和实践中，往往用所谓相关健壮（Correlation Robust）杂凑函数[27]搭建电路乱码（Circuit Garbling）、不经意传输及其扩展等密码学方案，再用这些方案进一步搭建安全多方计算协议与系统。进一步，Bellare 等人[28]在 S&P 2013 会议上提出用固定密钥 AES 搭建混淆电路，充分利用 AES-NI 指令与流水线实现加速，解决了混淆方案消耗 CPU 时间高的问题。受此启发，2013 年后设计的新开源安全多方计算组件、系统广泛使用固定密钥 AES。对此，郭淳等人[29-30]对基于（固定密钥）AES 的相关健壮杂凑函数进行了系统的研究，提出了系列方案，包括以下 3 个。

（1）调用一次固定密钥 AES 的循环相关健壮杂凑函数$\mathrm{MMO}(x)=\mathrm{AES}_0\big(\sigma(x)\big)\oplus\sigma(x)$，$\sigma$为正型变换（$\sigma$和$\sigma'(x)=\sigma(x)\oplus x$均为$\{0,1\}^n$上的置换变换）。该方案用于设计安全多方计算系统时，可保证半诚实（Semi-Honest）条件下的安全性。

（2）调用两次固定密钥 AES 的可调相关健壮杂凑函数$\mathrm{TMMO}(t,x)=\mathrm{AES}_0(\mathrm{AES}_0(x)\oplus t)\oplus\mathrm{AES}_0(x)$。该方案用于设计安全多方计算系统时，可保证针对恶意（Malicious）攻击者的安全性。

（3）调用一次完整 AES（带密钥编排）的可调相关健壮杂凑函数$\mathrm{MMO}(t,x)=\mathrm{AES}_t\big(\sigma(x)\big)\oplus\sigma(x)$，$\sigma$为正型变换。该方案用于设计乱码电路时，可保证“超生日界”安全性。

从理论上说，带密钥编排的完整 AES 性能比固定密钥 AES 低，因此一个自然的问题是能够使用 1~2 次固定密钥 AES 调用达到与完整 AES 方案相当的“超生日界”安全性。Chen、Tessaro[31]研究了这一问题，并提出以下两个新方案。

（1）调用一次固定密钥 AES、并执行一次$\mathrm{GF}(2^{128})$上有限域乘法$\otimes$的可调相关健壮杂凑函数$\mathrm{MMO}_{\mathrm{mul}}(t,x)=\mathrm{AES}_0(m\otimes t)\oplus(m\otimes t)$。

（2）调用两次固定密钥 AES 的可调相关健壮杂凑函数$\mathrm{FPTP}(t,x)=\mathrm{AES}_0\Big(t\oplus\mathrm{AES}_0\big(\sigma(x)\big)\Big)\oplus\sigma(x)$。

方案$\mathrm{MMO}_{\mathrm{mul}}$与可调 Even-Mansour 密码[32]有相似之处，仅用一次固定密钥 AES 调用，即可达到与两次固定密钥 AES 方案TMMO相当的安全性。但不幸的是，这并不意味着它在实际运用时的性能更好：在文献[29]的工作过程中，郭淳已与合作者讨论过$\mathrm{MMO}_{\mathrm{mul}}$方案，并因$\mathrm{GF}(2^{128})$乘法$\otimes$的低效率而将其淘汰。

方案FPTP性能与TMMO相当，而安全界在 OT 扩展的具体应用场景中得到了大幅改善。

简而言之，当它的调柄（Tweak）输入为随机值而“明文”输入x不重复时，它的可证明安全界为$\frac{B\sqrt{D}T}{2^n}+\frac{D^2}{2^n}$，其中$B$是每个特定调柄值下不同“明文”输入$x$的最大数目，$D$表示数据复杂度，$T$表示计算复杂度。要证明这个结论，需要考虑攻击者针对D个随机调柄值选取$2p$个值，所选值与调柄值所能构成的特定形式等式的数目最大值，此问题与随机 Cayley 图的性质有关，Chen 等人最早将相关结论（称为“Sum-Capture 引理”）引入对称密码证明中[15]，而 Chen、Tessaro 做了一个更精巧的应用。如前所述，这一结论事实上提出了新的（具备特定性质的）伪随机函数构造，因此也推动了“伪随机密码函数建构”研究的前进。

尽管上述“超生日界”结论对输入的限制很多，但这些限制在 OT 扩展的具体应用场景中是满足的。为此，Chen、Tessaro 进一步详细讨论了 OT 扩展的实际安全性（Concrete Security），并基于FPTP杂凑方案，提出了一个实际安全性更高的 OT 扩展方案。

3.3　MAC 与认证加密证明

基于分组密码和泛杂凑函数（Universal Hash Function），建构消息认证码 MAC 方案的思想起源于 Wegman 与 Carter[33]，这类方案的优势是以黑盒方式运用泛杂凑函数，在实际部署时可以根据平台的具体性能选择最适合的泛杂凑实例（或基于分组密码，或基于可调分组密码，或基于有限域乘法等）。具体而言，Wegman-Carter 方案基于块函数$F_K:\{0,1\}^n\to\{0,1\}^n$和泛杂凑$H_{K_h}:\{0,1\}^*\to\{0,1\}^n$建构，给定新鲜值（Nonce）和消息$M$，其计算认证码的方式为$\mathrm{WC}_{K,K_h}(N,M)=F_K(N)\oplus H_{K_h}(M)$。当泛杂凑$H_{K_h}$有充分好的组合学性质、$F_K$是伪随机函数且新鲜值不重复使用时，Wegman-Carter 方案可以保证最优的n比特安全性：其含义是可以安全地生成约2^n个不同认证码，同时可以抵抗攻击者对伪造的认证码进行约2^n次验证。

然而，如“伪随机函数/伪随机置换证明”所述，当块函数F_K是n比特伪随机置换/分组密码时，所得的 Wegman-Carter 方案变体（称为 Wegman-Carter-Shoup 方案）反而受到生日攻击的影响[19-20]，能够安全地生成的不同认证码个数减少到$2^{n/2}$个。解决此安全性问题的思路包括增加分组密码调用次数、使用随机初始向量等。在 2016 年美密会上，Cogliati 和 Seurin 提出了基于 2 次分组密码调用的 EWCDM（Encrypted Wegman-Carter with Davies-Meyer）MAC 方案 [34]，实现了“超生日界”安全性，引起了学界对此类方案研究的兴趣。后续研究对此进行了深度扩展，提出了 DWCDM（Decryted Wegman-Carter with Davies-Meyer）[35]、nEHtM [36]（nonce-based Enhanced Hash-then-Mask）等“超生日界”安全 MAC 方案。

基于上述进展，Chen 等人[37]对基于 2 次伪随机置换调用的 MAC 方案进行了全面的归类，提出了统合这类方案的框架；对其“新鲜值伪随机安全性”（Nonce-based PRF Security）进行了讨论，对每个类别方案分别给出了安全性的上界与下界。

具体而言，仅调用分组密码/置换一次的伪随机函数看起来是都存在生日攻击的。Chen 等

人没有对此进行讨论，但基于他们提出的框架可以找出这样的生日攻击。简而言之，仅调用分组密码/置换一次的伪随机函数通式为$\mathrm{F1}_K^E(x) = E_K(a_{11} \cdot x) \oplus (a_{21} \cdot x)$。选出$\lambda$个使$a_{11} \cdot x_1, \cdots, a_{11} \cdot x_\lambda$互异的输入$x_1, \cdots, x_\lambda$，并询问得$y_1 = \mathrm{F1}_K^E(x_1), \cdots, y_\lambda = \mathrm{F1}_K^E(x_\lambda)$。此时可以解得$E_K(a_{11} \cdot x_1) = y_1 \oplus (a_{21} \cdot x_1), \cdots, E_K(a_{11} \cdot x_\lambda) = y_\lambda \oplus (a_{21} \cdot x_\lambda)$。于是，当$\lambda = O\left(2^{\frac{n}{2}}\right)$时，即可通过解得的$y_1 \oplus (a_{21} \cdot x_1), \cdots, y_\lambda \oplus (a_{21} \cdot x_\lambda)$中有无碰撞的性质判断所询问的谕言机是伪随机函数$\mathrm{F1}_K^E(x)$还是真随机函数。

作为 Chen 等人本身的结论，他们先考虑了“两置换伪随机函数方案”，即由两次分组密码调用和任意次异或运算组成的伪随机函数方案，进行了统一化分析。为了进行这种分析，Chen 等人基于一个3×3矩阵定义了这类伪随机函数方案的通式：这一思路起源于文献[37]之第二、第三作者在 2012 年美密会发表的哈希函数安全性分析[17]。针对伪随机函数的问题，Chen 等人的结论是能够保证“超生日界”安全性的“两置换伪随机函数方案”只有以下 3 类。

（1）$\mathrm{SoP}_{K_1,K_2}(X) = E_{K_1}(X) \oplus E_{K_2}(X)$（于 1998 年欧密会提出[23]）。

（2）$\mathrm{EDM}_{K_1,K_2}(X) = E_{K_2}\left(E_{K_1}(X) \oplus X\right)$（于 2016 年美密会提出[34]）。

（3）$\mathrm{EDMD}_{K_1,K_2}(X) = E_{K_2}\left(E_{K_1}(X)\right) \oplus E_{K_1}(X)$（于 2017 年美密会提出[38]）。

其中，SoP_{K_1,K_2}的n比特安全性由 Dai 等人证明，所用证明方法称为χ^2方法；EDMD_{K_1,K_2}的n比特安全性由 Mennink、Neves 证明，方法是将其安全性归约至SoP_{K_1,K_2}；EDM_{K_1,K_2}的$n - \log_2 n$比特安全性由 Mennink、Neves 证明，该证明运用了 Patarin 关于方程组解的个数下界的“镜像理论”（Mirror Theory）。

Chen 等人进而考虑了“两置换一哈希 MAC 方案”，即由两次分组密码调用和一次泛哈希调用组成的 MAC 方案，对其“新鲜值伪随机安全性”进行了统一化分析。为了进行这种分析，Chen 等人基于一个3×4矩阵定义了这类 MAC 方案的通式，结论如下。

（1）有 9 种 MAC 方案能够保证“超生日界”安全性，具体如下。

①其中 3 种是将 Wegman-Carter MAC 中的伪随机函数替换成前述三种“超生日界”安全伪随机函数SoP_{K_1,K_2}、EDM_{K_1,K_2}、EDMD_{K_1,K_2}所得的。新鲜值不重用时，这些 MAC 能保证最优的n比特 MAC 安全性；但新鲜值一旦重用，其安全性即完全丧失。

②其余 6 种，包括 4 种基于EDM_{K_1,K_2}的 MAC 方案与 2 种基于SoP_{K_1,K_2}的 MAC 方案，在新鲜值不重用时，可以保证$3n/4$比特“超生日界”安全性。而其中 4 种“最优方案”在新鲜值（有限）重用时仍可保证“超生日界”安全性。

（2）有一种 MAC 方案在“新鲜值伪随机安全性”定义下，是存在生日攻击的[39]；但其仍有可能保证“超生日界”MAC 安全性，目前对此既无攻击又无证明。

更具体地说，在 4 种“最优方案”中，有两种是基于SoP_{K_1,K_2}建构的并行结构，是 Dutta

等人提出的 nEHtM 方案的变体；另两种则是基于EDM_{K_1,K_2}建构的新型串行结构，是 Chen 等人提出的新 MAC 方案。Chen 等人的方案的安全性证明运用了 Kim 等人[40]证明的$3n/4$比特安全“镜像理论”变体。

Chen 等人的方案有待后续解决的问题包括以下几点。

（1）Chen 等人仅对选出的结构证明了“新鲜值伪随机安全性”（Nonce-based PRF Security），这样只能确定可能具备“超生日界”MAC 安全性的方案，而不能最终证明 MAC 安全性。文中所提出的候选结构的“超生日界”MAC 安全性有待后续工作予以最终证明。

（2）Chen 等人对 EDM 伪随机函数的安全性证明用到了 Patarin“镜像理论”，但 Patarin 的证明尚未通过同行评审，被认为无法验证。目前已经过同行评审的“镜像理论”部分结论仅仅证明到$3n/4$比特安全性，距离“镜像理论”断言的n比特安全性还很遥远。“镜像理论”的证明是对称密码安全性证明中的一个关键组合学问题。

花开两朵，各表一枝。先前介绍的研究专注于基于“新鲜值”（Honce）的 MAC，下面介绍确定性 MAC（Deterministic MAC）的进展。确定性 MAC 的认证码生成函数直接将密钥-消息组合(K, M)映射为对应的认证码，不依赖额外的随机值/随机初始向量或“新鲜值”。这样设计的坏处是无法利用随机初始向量/“新鲜值”不重复出现的特性获得更强的安全界，但好处是不用消耗额外存储空间维护随机初始向量/“新鲜值”，因而有利于轻量级应用场景。因此，ISO/IEC 29192-6:2019 标准推荐的 3 种轻量级 MAC 均为确定性 MAC，分别是基于分组密码的工作模式 LightMAC，基于密码杂凑函数的工作模式 Tsudik keymode（以$\mathrm{Hash}(M \,||\, K)$为认证码的朴素思路）和基于密码学置换的方案 Chaskey-12（这或许体现了 ISO 标准对于“设计理念多样性”的看重）。

其中，LightMAC 是 Luykx 等人在会议 FSE 2016 上提出的方案[41]。假设其使用n比特分组密码，且认证的消息M长度为$(n-s)$比特的倍数时，将M写作ℓ个$n-s$比特的分组$M[1],\cdots,M[\ell]$，并利用两个分组密码密钥K_1、K_2计算认证码如下。

$\mathrm{LightMAC}_{K_1,K_2}(M)$:

$$= E_{K_2}\left(E_{K_1}(M[1] \,||\, \langle 1\rangle_s) \oplus \cdots \oplus E_{K_1}(M[\ell-1] \,||\, \langle \ell-1\rangle_s) \oplus M[\ell] \,||\, 10^{s-1}\right)$$

其中，$\langle i\rangle_s$为整数i的s比特表示，10^{s-1}为比特 1 和$s-1$个比特 0 连接而成的s比特串。其具备以下突出特点。

（1）计算过程只包括分组密码调用和异或运算，不需要任何有限域乘法。

（2）支持分组密码调用并行化。

（3）可证明安全界在典型场景中与消息长度无关。具体而言，当其使用n比特分组密码、认证的消息长度不超过$(n-s)2^s$比特时，对其进行q次询问的攻击者之伪造成功概率不超过$O(q^2/2^n)$，成功概率与攻击者询问的消息长度无关。

上述第 3 点是 LightMAC 的突出特点，是 PMAC [42]、OMAC、Tsudik keymode 等确定性

MAC 标准所不具备的。事实上，LightMAC 可以看作将消息分组$M[1],\cdots,M[\ell-1]$依次与一个计数器的数值“连接”来产生$\ell-1$次分组密码调用的输入，利用不同的计数值实现了某种程度上的定义域分离（Domain Separation）；这是其能够不借助有限域乘法而实现“消息长度无关”安全界的关键点之一。当然，代价是理论效率有所下降。例如，当计数器数值长度$s=n/4$时，LightMAC 消耗 4 次分组密码调用才能为$3n$比特消息（消息中的n比特分组数目为 3）生成认证码，消息中（n比特）分组数目与分组密码调用次数的比值为$3/4$（该比值通常称为“rate”：一般认为“rate”接近 1 的分组密码模式是最优的）。此外，一旦消息长度超过$(n-s)2^s$比特，伪造就易如反掌。但是，这一缺点在轻量级应用场景中并非特别严重：由于摆脱了有限域乘法等运算，节约了软硬件实现开销，稍低的“rate”作为一个代价并非不可接受；而很多轻量级应用场景中的消息长度本身就不太可能超过$(n-s)2^s$比特的限制。

LightMAC 另一个关键设计理念是使用两个不同分组密码密钥K_1、K_2（如前所述）：在吸收明文分组的“前半部分”中的分组密码调用使用密钥K_1，最后用K_2将明文分组形成的摘要加密生成认证码。由于使用密钥K_1与K_2的分组密码调用可以看作是相互独立的，因此 LightMAC 可以看作经典的“先哈希再加密”（Hash-then-PRF）范式的一个实例，其安全性证明可以专注于论证“前半部分”（有密钥的）杂凑函数的“弱抗碰撞性”（Weak Collision Resistance），这显著简化了其安全性证明。但存储、运用两个密钥开销不小，对于 LightMAC 的设计而言属于不可忽略的代价了。

针对以上内容，印度学者 Chattopadhyay 等人[43]提出以下两个问题。

（1）只用一个密钥的 LightMAC 变体（记作 1k-LightMAC）是否还能实现 “消息长度无关”安全界？

（2）若只用一个密钥会导致 LightMAC 安全性显著下降，则能否通过引入极小的改动来恢复与原版 LightMAC 相当的安全界？

“前半部分”与最终的分组密码调用使用相同密钥时，它们的输入值间的碰撞也会导致最终输出的认证码失去（伪）随机性。因此，解决上述问题需要在安全性证明里小心地处理更多类型的碰撞事件，而朴素的处理方法只能得到不尽如人意的$O(q^2\ell/2^n)$安全界。

Chattopadhyay 等人首先证明，当消息长度不超过$(n-s)\min\{2^{n/4},2^s\}$比特时（比 LightMAC 稍强的限制），1k-LightMAC 的安全界也是$O(q^2/2^n)$，与原版 LightMAC 是相当的。在通信报文长度存在上限的物联网等场景中，1k-LightMAC 更强的消息长度限制不会是弱点，因此会严格优于原版 LightMAC。为克服前述安全性分析障碍，Chattopadhyay 等人提出了一种在基于 H-系数（H-coefficient）方法的安全性证明中更改部分随机中间值的证明方法，并称为 reset-sampling。

为了克服 1k-LightMAC 对消息长度更严格的限制，Chattopadhyay 等人进一步提出了 1k-LightMAC 的一个改进方案（称为 LightMAC-ds），并证明，当消息长度不超过$(n-s)2^{s-1}$比

特时，LightMAC-ds 的安全界也是$O(q^2/2^n)$。LightMAC-ds 的消息长度限制和安全界就与原版 LightMAC 几乎完全一致了。Chattopadhyay 等人的结论与新方案预期会对 ISO/IEC 29192-6:2019 标准产生积极的影响。

认证加密（Authenticated Encryption）的概念由 Bellare、Namprempre [44]在 2000 年亚密会上正式提出，是能够同时完成数据完整性校验与加密的对称密码学方案。自其提出后，学术界与产业界都认识到了 AE 的重要性，并迅速展开了研究与部署，特别是 2014 年启动的 CAESAR 认证加密设计竞赛[45]与 2018 年启动的 NIST 轻量级 AE 标准竞选工程[5]，将此研究两度推向高潮，最终催生了 AES-OCB、Ascon、ASE-GCM-SIV 等成熟设计。

抗新鲜值误用[46]（Nonce-Misuse Resistance）和“超生日界”安全性是新型认证加密设计所追求的两个重要安全性目标。在 2016 年美密会上，Thomas Peyrin、Yannick Seurin 提出了 NSIV、SCT 等基于可调分组密码的认证加密工作模式[47]，首次用同一个方案同时实现了抗新鲜值误用与“超生日界”安全性两大目标。

韩国学者 Choi 等人[48]提出了一个新的基于（传统）分组密码的认证加密工作模式 SCM，同时首次实现了理论最高效率、最优安全性（新鲜值不重用时）、安全性随新鲜值重用的下降率平缓三大目标。SCM 设计的关键点包括以下内容。

（1）“加密新鲜值”（Encrypting Nonces）：在加密伊始，利用 CENC 模式（一个“超生日界”安全的密钥流生成模式）从新鲜值导出 3 个伪随机密钥Δ、Δ'、Δ''。

（2）基于Δ''，用一个 nEHtM MAC 模式的变体生成认证码T。这一步骤的“超生日界”安全性由 Dutta 等人证明[36]。

（3）用分组密码E、伪随机密钥Δ与Δ'组成一个类似 XE 模式[42]的可调分组密码，用该可调分组密码加密认证码T形成密钥流，并用密钥流加密明文。

在性能方面，SCM 支持并行实现，且不需要实现分组密码的解密运算。

与 AES-GCM-SIV（RFC 8452）相比，SCM 的优势包括以下两个方面。

（1）AES-GCM-SIV 的可证明安全性假设为“AES 在多密钥情形下有着接近理想密码的安全性”，而 SCM 仅需依赖经典假设“AES 是伪随机置换”。

（2）当新鲜值不重用且明文不太长时，SCM 能够实现远高于 AES-GCM-SIV 的实际安全性。

另外，作者称设计基于分组密码的认证加密模式一大动机是“AES-NI 指令的存在使得分组密码（AES）能够具备远优于可调分组密码的性能”，此观点却是值得商榷的：基于 AES 轮函数设计的可调分组密码 Deoxys-TBC 已入选 ISO/IEC 标准 18033-7:2022，该设计同样可以依托 AES-NI 指令实现极高的性能。在此方面，笔者认为分组密码 AES 等相对可调分组密码的优势是更可靠的安全性，原因是 AES 等经历了强度远高于新型可调分组密码算法的密码分析与安全性检验。

3.4 量子 Q2 模型中的安全性证明

量子算法与量子计算机的高速发展是近年最受瞩目的新兴技术之一。量子计算机运行量子算法，它基于量子物理原理，可以在一定程度上发挥并行计算的能力，加速特定问题的求解，严重威胁经典密码系统的安全性。例如，Shor 算法对 RSA 等的攻击推动了 NIST 抗量子公钥密码标准征集[49]。

很多对称加密标准在量子计算机上运行并处理叠加态（superposition）输入时，可以基于它们定义有周期的数学函数，并用 Simon 算法[50]在多项式时间内求解周期，从而用多项式规模的复杂度破解它们。例如，Kaplan 等人[51]在 2016 年美密会上提出了针对 CBC-MAC（ISO/IEC 9797）、PMAC（IACR Fellow P. Rogaway 设计）、GCM（NIST SP 800-38D 标准）、OCB（RFC 7253）等重要加密或认证标准的攻击。这种模型往往称为量子 Q2 模型。这与学界“增加密钥就能搭建抗量子攻击对称密码”的旧有印象大相径庭。由于这些发现，伪随机函数、MAC、对称加密方案等的量子 Q2 安全性已得到了广泛关注：自 2010 年起，一系列研究确立了量子伪随机模型[52]、量子随机谕言模型[53]、量子语义安全[54]等重要结论，提供了部分解决方案及分析技术，成果持续发表于 FOCS、美密会、欧密会等密码理论顶级会议。

本节关注对称密码方案的量子 Q2 安全性，介绍的两篇论文分别推动了认证加密和 HMAC/NMAC 的量子 Q2 安全性研究。

学界已对（不带认证功能的）加密模式 CBC、CTR 等的量子 Q2 安全性进行过讨论[55]。关于认证加密，OCB 是三大认证加密模式（GCM、CCM、OCB）中效率最高的[56]，处理m个明文和关联数据（Associated Data）分组仅需调用分组密码$m+2$次（这种性质一般称为“rate-1”）。如前所述，在量子 Q2 模型中，OCB 可以被基于 Simon 算法的攻击破解，复杂度为多项式。NIST 轻量级密码标准征集第二轮候选方案 Saturnin 具备量子 Q2 安全性，但为了达到此安全目标，其使用“先加密后认证”（Encrypt-then-MAC）方法，将 CTR 加密模式与类似于 HMAC/NMAC 的量子安全 MAC 相结合，其效果不理想。因此，基于 OCB 的思路，设计具备量子 Q2 安全性的“rate-1”认证加密方案，是对称密码与抗量子密码领域最吸引人的问题之一。

Bhaumik 等人[57]研究了上述问题，首先推广了针对 OCB 的量子 Q2 攻击，证明在（基于分组密码的）OCB 基础上，修改每个明文分组对应的“偏移值”（Offset）是不能实现量子 Q2 安全性的。同时，提出了针对ΘCB（相当于基于可调分组密码的 OCB）的量子 Q2 攻击，证明纵然用（比分组密码更强的）可调分组密码设计认证加密方案，也无法沿用即有结论ΘCB。

基于这些发现，Bhaumik 等人提出基于可调分组密码的“rate-1”认证加密模式 QCB。该方案整体结构仍与ΘCB相似，其中克服量子攻击的关键在于，QCB 处理第i个明文分组/关联数据分组时，会将下标号i与初始向量IV连起来作为相应可调分组密码调用的调柄。Bhaumik

等人证明了 QCB 的量子 Q2 安全性，所用的量子 Q2 安全性定义为 IND-qCPA[54]和“BZ-不可伪造性”[58]。其中，IND-qCPA 的含义是攻击者通过进行叠加态加密询问，仍不能区分不同的经典态挑战明文（Classical Challenge Messages，实际含义为“不能破获所加密的经典态明文的信息”）；“BZ-不可伪造性”的含义则是攻击者进行q次叠加态加密询问后，无法利用所得信息制造$q+1$个合法的明文/密文/认证码三元组。

为了将 QCB 转化为基于经典分组密码的认证加密模式，Bhaumik 等人考虑了量子安全可调分组密码。由于 QCB 内部可调分组密码调用的限制，它使用的可调分组密码仅需在攻击者使用经典调柄和叠加态明文进行询问时保证安全。针对这一有限制的可调分组密码量子安全定义，Bhaumik 等人提出了密钥调整插入（Key-tweak Insertion）TBC $\tilde{E}_K(T,x)=E_{K\oplus T}(x)$，并在分组密码$E$具备相关密钥安全性的假设下，为其证明了量子安全性。

综上所述，QCB 及其基于分组密码的实例构成了国际范围内首个“rate-1”，且支持明文分组并行处理的量子 Q2 安全认证加密方案。

消息认证码 MAC 有两种著名构建方式：其一为 CBC-MAC 等基于分组密码的工作模式，其二为 HMAC、NMAC 等基于抗碰撞密码杂凑函数的结构。为介绍后一种技术路线，先介绍 Merkle-Damgard 迭代杂凑函数结构。具体而言，给定压缩函数$h:\{0,1\}^{m+n}\to\{0,1\}^n$，Merkle-Damgard 构造$\mathrm{MD}^h$定义如下：设$\mathrm{IV}\in\{0,1\}^n$为固定且公开的初始化向量。首先对消息$M$进行填充，将填充后的消息拆分为$m$比特消息分组$M[1],\cdots,M[\ell]$。将状态初始化为$S_0:=\mathrm{IV}$，然后迭代计算$S_{i+1}:=h(M[i+1]\,||\,S_i)$，最终输出$S_\ell$。

由此，可引出 HMAC、NMAC 的定义方式。当密钥长度为$k\leqslant m$时，$\mathrm{HMAC}^h:\{0,1\}^k\times\{0,1\}^n\times\{0,1\}^*\to\{0,1\}^n$定义为$\mathrm{HMAC}^h(K,\mathrm{IV},M):=\mathrm{MD}^h(\mathrm{IV},K_{\mathrm{out}}\,||\,\mathrm{MD}^h(\mathrm{IV},K_{\mathrm{in}}\,||\,M))$，其中$K_{\mathrm{in}}:=(K\,||\,0^{m-k})\oplus\mathrm{ipad}$，$K_{\mathrm{out}}:=(K\,||\,0^{m-k})\oplus\mathrm{opad}$，$\mathrm{ipad}、\mathrm{opad}\in\{0,1\}^n$是固定、公开但不相等（$\mathrm{ipad}\neq\mathrm{opad}$）的常量。

$\mathrm{NMAC}^h:\{0,1\}^n\times\{0,1\}^n\times\{0,1\}^*\to\{0,1\}^n$则是 HMAC 的双密钥变体，定义为$\mathrm{NMAC}^h(K_1,K_2,M):=\mathrm{MD}^h\big(K_2,\mathrm{MD}^h(K_1,M)\big)$，其中$K_1$、$K_2\in\{0,1\}^n$是两个独立随机选择的密钥。HMAC、NMAC 实际上也是最重要的、基于 Merkle-Damgard 迭代杂凑函数的伪随机函数，已被 ISO、IETF 等重要标准收录。

在量子 Q2 安全模型中，HMAC、NMAC 各自存在两种攻击：其一为基于 Grover 搜索的密钥穷举攻击，对 HMAC、NMAC 的攻击复杂度分别为$O(2^{k/2})$、$O(2^{2n/2})$（符合通常所认为的“密钥长度折半”效应），其二为利用 HMAC、NMAC 输出碰撞的区分攻击，对 HMAC、NMAC 的攻击复杂度均为$O(2^{n/2})$。由此，HMAC 的量子 Q2 安全性上界为$\min\{k/2,n/3\}$比特，NMAC 则为$\min\{n,n/3\}=n/3$比特。证明与这两个数值一致的安全界是有意义且有挑战性的问题。

在非量子场景中，HMAC、NMAC 安全性证明往往假设$h(\cdot\,||\,K):\{0,1\}^m\to\{0,1\}^n$是伪随

机函数。在量子 Q2 场景中，Song 和 Yun[59]考虑了类似的假设，即$h(\cdot \,||\, K):\{0,1\}^m \to \{0,1\}^n$是量子伪随机函数（qPRF，即在量子 Q2 模型中安全的伪随机函数）；并在此假设下，证明了 HMAC、NMAC 的量子 Q2 安全性下界为$n/8$比特或$n/5$比特（后一结论在一个特定的猜想下成立）。但即使是更好的、基于猜想的$n/5$比特下界，也和上述$n/3$比特上界有显著的距离。对实际运用而言，这意味着若要根据 Song-Yun 结论选取参数n，则须保证$n/5$不少于所需的安全性，如$n/5 \geqslant 80$以保证 80 比特量子 Q2 安全性，这意味着$n \geqslant 400$，这样大的参数对于减少实现开销显然是不利的。

日本学者 Hosoyamada、Iwata [60]进一步研究了 HMAC、NMAC 的量子 Q2 安全性，将底层压缩函数h看作量子随机谕言（Quantum Random Oracle），并在量子随机谕言模型[53]（Quantum Random Oracle Model，即考虑协议实体和攻击者均可自由询问一个公开的量子随机谕言h的场景，并在此假设下论证协议的安全性）中，为证明了量子 Q2 紧界（Tight Bounds）。具体而言，假设：

①$m \geqslant n$；

②对 HMAC、NMAC 询问的消息长度不超过ℓm比特；

③$h:\{0,1\}^{m+n} \to \{0,1\}^n$是（对所有实体）公开的量子随机谕言；

④攻击者对 HMAC 进行Q次询问，对h进行q_h次询问。

则：

①以常数概率区分 HMAC 和随机函数的必要条件是$q_h \cdot \ell^{5/3} + Q \cdot \ell^{5/3} \geqslant \Omega(2^{n/3})$或$q_h + Q \cdot \ell \geqslant \Omega(2^{k/2})$；

②以常数概率区分 NMAC 和随机函数的必要条件是$q_h \cdot \ell^{5/3} + Q \cdot \ell^{5/3} \geqslant \Omega(2^{n/3})$。

简言之，以上结论表明，在所处理消息不是特别长的情形下，HMAC、NMAC 的量子攻击复杂度不能低于$O(2^{n/3})$，这和此前从攻击得出的上界一致，因而是紧界。以具体场景的参数举例而言，以太网协议使用 HMAC-SHA-256（$n = 256$）验证 TCP/IP 数据包时，有$\ell < 32$，此时 Hosoyamada-Iwata 结论保证了$256/3 \approx 85$比特 Q2 安全性，而 Song-Yun 结论只保证了约 52 比特或 32 比特 Q2 安全性。

如前所述，所处理消息较长时，Hosoyamada-Iwata 结论不再为紧界。但在此情形下，也并不存在复杂度低于$O(2^{n/3})$的攻击。因此，针对长消息改进 Hosoyamada-Iwata 结论是公开问题。

本节作者：郭淳（山东大学）

参考文献

[1] LUBY M, RACKOFF C. Pseudo-random Permutation Generators and Cryptographic

Composition[C]. ACM STOC 1986: 356-363.

[2] EVEN S, MANSOUR Y. A Construction of a Cipher from a Single Pseudorandom Permutation [C]. ASIACRYPT 1991: 210-224.

[3] BELLARE M, ROGAWAY P. Random Oracles are Practical: A Paradigm for Designing Efficient Protocols[C]. ACM CCS 1993: 62-73.

[4] BELLARE M, KILIAN J, ROGAWAY P. The Security of Cipher Block Chaining[C]. CRYPTO 1994: 341-358.

[5] NIST: Lightweight Cryptography Standardization: Overview.https://csrc.nist.gov/projects/lightweight-cryptography.

[6] GUO C, IWATA T, KHAIRALLAH M,et al. Romulus v1.3.https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/finalist-round/updated-spec-doc/romulus-spec-final.pdf.

[7] HIROSE S. Some plausible constructions of double-block-length hash functions[C]. FSE 2006: 210-225.

[8] NAITO Y. Optimally indifferentiable double-block-length hashing without postprocess-ing and with support for longer key than single block[C]. LATINCRYPT 2019: 65-85.

[9] ÖZEN O, STAM M. Another glance at double-length hashing[C]. IMACC 2009: 176-201.

[10] NAITO Y.Indifferentiability of double-block-length hash function without feedforw-ard operations[C]. ACISP 2017: 38-57.

[11] NAITO Y, SASAKI Y, SUGAWARA T. Double-Block-Length Hash Function for Minimum Memory Size[C]. ASIACRYPT 2021(3): 376-406.

[12] CORON J-S, DODIS Y, MALINAUD C, et al. Merkle-Damgård Revisited: How to Construct a Hash Function[C]. CRYPTO 2005: 430-448.

[13] GOLDREICH O, GOLDWASSER S, MICALI S. How to Construct Random Functions (Extended Abstract)[C]. FOCS 1984: 464-479.

[14] YU Y, STEINBERGER. J Pseudorandom Functions in Almost Constant Depth from Low-Noise LPN[C]. EUROCRYPT 2016(2): 154-183.

[15] CHEN S, LAMPE R, LEE J, et al. Minimizing the Two-Round Even-Mansour Cipher[C]. CRYPTO2014(1): 39-56.

[16] TESSARO S, ZHANG X. Tight Security for Key-Alternating Ciphers with Correlated Sub-keys[C]. ASIACRYPT 2021(3): 435-464.

[17] MENNINK B, PRENEEL B. Hash Functions Based on Three Permutations: A Generic Security Analysis [C]. CRYPTO 2012: 330-347.

[18] KILTZ E, PIETRZAK K, SZEGEDY M. Digital Signatures with Minimal Overhead from Indifferentiable Random Invertible Functions[C]. CRYPTO 2013(1): 571-588.

[19] LEURENT G, SIBLEYRAS F. The Missing Difference Problem, and Its Applications to Counter Mode Encryption[C]. EUROCRYPT2018(2): 745-770.

[20] LUYKX A, PRENEEL B. Optimal Forgeries Against Polynomial-Based MACs and GCM[C]. EUROCRYPT2018(1): 445-467.

[21] MENNINK B, NEVES S. Optimal PRFs from Blockcipher Designs. IACR Trans. Symmetric Cryptol. 2017(3): 228-252.

[22] BANIK S, ISOBE T, LIU F, et al. Orthros: A Low-Latency PRF. IACR Trans. Symmetric Cryptol. 2021(1): 37-77.

[23] BELLARE M, KROVETZ T, ROGAWAY P. Luby-Rackoff backwards: Increasing security by making block ciphers non-invertible[C]. EUROCRYPT 1998: 266-280.

[24] LUCKS S. The Sum of PRPs Is a Secure PRF[C]. EUROCRYPT 2000: 470-484.

[25] DAI W, HOANG V T, Stefano Tessaro. Information-Theoretic Indistinguishability via the Chi-Squared Method[C]. CRYPTO 2017(3): 497-523.

[26] BHATTACHARYA S, NANDI M. Luby-Rackoff Backwards with More Users and More Security[C]. ASIACRYPT 2021(3): 345-375.

[27] ISHAI Y, KILIAN J, NISSIM K, et al. Extending Oblivious Transfers Efficiently[C]. CRYPTO 2003: 145-161.

[28] BELLARE M, HOANG V T, KEELVEEDHI S, et al. Efficient Garbling from a Fixed-Key Blockcipher[C]. IEEE Symposium on Security and Privacy 2013: 478-492.

[29] GUO Chun, KATZ J, WANG Xiao, et al. Efficient, Secure Multiparty Computation from Fixed-Key Block Ciphers[C]. IEEE Symposium on Security and Privacy 2020: 825-841.

[30] GUO C, KATZ J, WANG X, et al. Better Concrete Security for Half-Gates Garbling (in the Multi-instance Setting)[C]. CRYPTO 2020(2): 793-822.

[31] CHEN Y, TESSARO S. Better Security-Efficiency Trade-Offs in Permutation-Based Two-Party Computation[C]. ASIACRYPT 2021(2): 275-304.

[32] COGLIATI B, LAMPE R, SEURIN Y. Tweaking Even-Mansour Ciphers[C]. CRYPTO 2015(1): 189-208.

[33] WEGMAN M N, CARTER L. New Hash Functions, Their Use in Authentication and Set Equality. J. Journal of Computer & System Sciences, 1981, 22(3): 265-279.

[34] COGLIATI B, SEURIN Y. EWCDM: An Efficient, Beyond-Birthday Secure, Nonce-Misuse Resistant MAC[C]. CRYPTO 2016(1): 121-1493.

[35] DATTA N, DUTTA A, NANDI M, et al. Encrypt or Decrypt? To Make a Single-Key Beyond Birthday Secure Nonce-Based MAC[C]. CRYPTO 2018 (1): 631-661.

[36] DUTTA A, NANDI M, TALNIKAR S. Beyond Birthday Bound Secure MAC in Faulty Nonce Model[C]. EUROCRYPT2019(1): 437-466.

[37] CHEN Y, MENNINK B, PRENEEL B. Categorization of Faulty Nonce Misuse Resistant Message Authentication[C]. ASIACRYPT2021(3): 520-550.

[38] MENNINK B, NEVES S. Encrypted Davies-Meyer and Its Dual: Towards Optimal Security Using Mirror Theory[C]. CRYPTO 2017(3): 556-583.

[39] NANDI M. Mind the Composition: Birthday Bound Attacks on EWCDMD and SoKAC21[C]. EUROCRYPT2020(1): 203-220.

[40] KIM S, LEE B, LEE J. Tight Security Bounds for Double-Block Hash-then-Sum MACs [C]. EUROCRYPT2020(1): 435-465.

[41] LUYKX A, PRENEEL B, TISCHHAUSER E , et al. A MAC Mode for Lightweight Block Ciphers[C]. FSE 2016: 43-59.

[42] ROGAWAY P. Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC[C]. ASIACRYPT 2004: 16-31.

[43] CHATTOPADHYAY S, JHA A, NANDI M. Fine-Tuning the ISO/IEC Standard LightMAC[C]. ASIACRYPT 2021(3): 490-519.

[44] BELLARE M, NAMPREMPRE C. Authenticated Encryption: Relations among Notions and Analysis of the Generic Composition Paradigm[C]. ASIACRYPT 2000: 531-545.

[45] Cryptographic competitions: Introduction. https://competitions.cr.yp.to/index.html.

[46] ROGAWAY P, SHRIMPTON T. A Provable-Security Treatment of the Key-Wrap Problem[C]. EUROCRYPT 2006: 373-390.

[47] PEYRIN T, SEURIN Y. Counter-in-Tweak: Authenticated Encryption Modes for Tweakable Block Ciphers[C]. CRYPTO 2016(1): 33-63.

[48] CHOI W, LEE B, LEE J, et al. Toward a Fully Secure Authenticated Encryption Scheme from a Pseudorandom Permutation[C]. ASIACRYPT2021(3): 407-434.

[49] NIST: Post-Quantum Cryptography Standardization: Overview. https://csrc.nist.gov/projects/post-quantum-cryptography.

[50] SIMON D R. On the Power of Quantum Computation[C]. FOCS 1994: 116-123.

[51] KAPLAN M, LEURENT G, LEVERRIER A, et al. Breaking Symmetric Cryptosystems Using Quantum Period Finding[C]. CRYPTO2016(2): 207-237.

[52] ZHANDRY M. How to Construct Quantum Random Functions[C]. FOCS 2012: 679-

687.

[53] BONEH D, DAGDELEN Ö, FISCHLIN M, et al. Random Oracles in a Quantum World[C]. ASIACRYPT 2011: 41-69.

[54] BONEH D, ZHANDRY M. Secure Signatures and Chosen Ciphertext Security in a Quantum Computing World[C]. CRYPTO 2013(2): 361-379.

[55] ANAND M V, TARGHI E E, TABIA G N, et al. Post-Quantum Security of the CBC, CFB, OFB, CTR, and XTS Modes of Operation[C]. PQCrypto 2016: 44-63.

[56] KROVETZ T, ROGAWAY P. The Software Performance of Authenticated-Encryption Modes[C]. FSE 2011: 306-327.

[57] BHAUMIK R, BONNETAIN X, CHAILLOUX A, et al. QCB: Efficient Quantum-Secure Authenticated Encryption[C]. ASIACRYPT2021(1): 668-698.

[58] BONEH D, ZHANDRY M. Quantum-Secure Message Authentication Codes[C]. EUROCRYPT 2013: 592-608.

[59] SONG F, YUN A. Quantum Security of NMAC and Related Constructions-PRF Domain Extension Against Quantum attacks[C]. CRYPTO 2017(2): 283-309.

[60] HOSOYAMADA A, IWATA T. On Tight Quantum Security of HMAC and NMAC in the Quantum Random Oracle Model[C]. CRYPTO 2021(1): 585-615.

4 对称密码攻击

安全多方计算（MPC）、零知识证明（ZK）和全同态加密（FHE）是近几年密码学界的研究热点，显现了良好的应用前景。LowMC（欧密会 2015）、Rasta（美密会 2018）等都是面向 MPC/FHE/ZK 的应用场景而专门设计的对称密码算法。其中，分组密码 LowMC 作为底层加密方案被应用于后量子签名方案 PICNIC（NIST 后量子密码竞赛第三轮备选算法）中。在 PICNIC 数字签名方案中，LowMC 生成的明-密文对用作签名公钥，相应的加密密钥用作签名私钥。因此，在对 LowMC 的密码分析时，研究攻击者只能访问单个已知明-密文对的攻击情境，对讨论 PICNIC 数字签名方案的安全性尤为重要。但这是一个具有挑战性的数学问题，因为在这种极端情况下，攻击者无法使用绝大多数对称密码标准分析技术，如线性分析和差分分析。Rasta 是一族流密码，具有极低的 ANDdepth，且等于每个加密比特的 AND 门数。由于在 Rasta 中每次加密生成仿射层非常耗时，Hebborn 和 Leander 设计了 Dasta，其线性层被替换为一个不断变化的比特置换和一个确定性线性变换。Rasta 和 Dasta 的一个特点是，分组长度/密钥长度n比要达到的安全级别κ大得多。为了鼓励更多的密码分析，Rasta 的设计者还提出了一个激进的版本，称为 Agrasta，其分组大小仅略大于安全级别，$n = \kappa$+1。

在亚密会 2021 上，Liu 等人[1]指出，Rasta 和 Dasta 的设计者忽略了 x 操作的一个重要性质。利用这一性质，并结合 Rasta 和 Dasta 的特殊结构，作者极大地改进了代数攻击。特别是作者从理论上攻破了 2 个（共 3 个）全轮 Agrasta 实例。作者还进一步揭示了，由于 Dasta 使用一个由不断变化的比特置换和确定性线性变换组成的线性层，因此 Dasta 比 Rasta 更容易受到该攻击。根据作者的分析，(327,80,4)、(1877,128,4)、(3545,256,5)这几个版本的 Dasta 和 Rasta 的安全边界减少到只有 1 轮，其中 3 个参数分别表示分组大小、声称的安全级别和轮数。和其他可以在合理时间内实现并针对相同安全级别的版本相比，这几个版本的算法的 ANDdepth 是最低的。因此，研究他们的安全性具有特殊的重要性。

在 Banik 等人的论文中（IACR ToSC 2020:4），作者使用 S 盒线的线性化技术对 LowMC 的某些实例进行了攻击。在亚密会 2021 上，Banik 等人[2]首先对该线性化攻击给出了更精确的复杂度分析。然后，他们展示了如何对 LowMC 分两阶段进行中间相遇攻击，即 2-阶段 MITM 攻击。第一阶段，筛除一些对应于主密钥部分比特的候选密钥；第二阶段，对这个削减的候选集合和剩余的部分密钥比特，进行中间相遇攻击，可以成功地恢复主密钥。这两个阶段的综合计算复杂度显著低于 Banik 等人在 ToSC 2020 论文中报告的复杂度。

在美密会 2021 上，Liu 等人[3]重新考虑了对 LowMC 算法的差分枚举技术，并提出了新的代数技术来实现有效的密钥恢复攻击。在最初的差异枚举攻击框架中，一个不可避免的步骤是预先计算并存储一组中间状态的差分，以便通过二叉树搜索进行有效检查。作者的第一个发现是，Bar-On 等人为具有部分非线性层的 SPN 开发的通用代数技术可以用来完成相同的任务，但这项技术不需要存储大量的状态差分，因此可以使内存复杂度忽略不计。借助 Bar-On 等人的技术，对于分组大小远大于密钥大小的情况，作者可以显著改进对 LowMC 的攻击，甚至可以完全攻破 LowMC。另外，作者提出了新的密钥恢复技术，在只有一个输入-输出消息对以及它们的差分轨迹的情况下，可以极大地缩短检索全部密钥的时间。结合这两种技术，只选择 2 个明文，他们可以攻击使用完整非线性层的 4 轮 LowMC，分组长度分别为 129 比特、192 比特、255 比特。注意，这是 PICNIC 3 推荐的 3 个参数。需要指出的是，他们的攻击并不表明 PICNIC 3 被攻破了，因为 PICNIC 用例非常特殊，所以攻击者无法自由选择 2 个明文来加密一个具体的 LowMC 实例。但是，这些参数在最新的 LowMC 版本中被认为是安全的。此外，对于 Peyrin 和 Wang[4]在美密会 2020 上提出的后门密码 LowMC-M 的 7 个实例，在未找到后门的情况下，通过充分利用允许的 2^{64} 个数据，作者可以攻击更多的轮数。上述攻击所占用的内存很低，都是可忽略的。

面向 MPC/FHE/ZK 的对称密码设计是一个较新的课题。从设计的角度来看，这些攻击表明，这个课题还未被很好地认识和理解，设计理论还不成熟，现有的密码算法往往存在安全性问题。另外，从分析的角度来看，对于新型密码算法，也需要开发新型的攻击技术。

Double-block Hash-then-Sum MACs（DbHtS MACs）是一类旨在实现超越生日界安全性的

消息认证码，包括 SUM-ECBC、PMAC_Plus、3kf9 和 LightMAC_Plus。最近，Datta 等人（FSE'19）及 Kim 等人（Eurocrypt'20）证明 DbHtS 结构在单用户情境下是超越生日界安全的，然而，在多用户情境下，他们的安全性证明结果降级到生日界，甚至更差。

在美密会 2021 上，Shen 等人[5]重新考虑了多用户情境下 DbHtS MAC 的安全性。他们提出了一个通用框架来证明 DbHtS 结构的超越生日界安全性。为了展示该框架的可用性，他们提出了几个 DbHtS MAC 的密钥缩减的变体，包括 2k-SUM-ECBC、2k-PMAC_Plus 和 2k-LightMAC_Plus。他们的研究结果表明，这些结构的安全性不会随着用户数量的增加而降低。另外，在先前的工作中，研究者使用域分离来简化证明，而 Shen 等人的研究结果还表明，无须额外的域分离，也可以证明这些结构在单用户和多用户情境下都是超越生日界安全的。此外，作者发现了 2kf9 一个严重缺陷，而在此之前，Datta 等人（FSE'19）证明在生日限制之外是安全的。Shen 等人可以在不进行任何查询的情况下成功伪造概率为 1 的标签，并进一步展示了对 2kf9 的几个变体具有生日限制复杂性的攻击。

ChaCha 是由密码学家 Bernstein 在 2008 年设计的流密码算法。ChaCha 和 Poly1305 消息认证码被谷歌公司采用为 TLS 加密套件之一。目前，已有大量协议和软件实现了 ChaCha 密码算法。由于 ChaCha 的广泛应用，因此对其安全性进行系统、深入的分析非常重要。

在欧密会 2021 上，Coutinho 和 Neto 针对 ARX 密码算法提出了寻找线性逼近的新方法，并且使用该方法首次显式地推导出 3 轮和 4 轮 ChaCha 的线性逼近[6]。此外，他们声称发现了 3 和 3.5 轮 ChaCha 的新差分-线性逼近。利用这些发现，他们改进了 6 轮和 7 轮 ChaCha 的分析结果。但是，随后 Juan、V´asquez 指出其公开代码中的错误，该错误导致了 3.5 轮 ChaCha 上差分-线性（单比特掩码）的不正确结果。独立地，Dey 等人[7]也指出其 3.5 轮 ChaCha 上新差分-线性逼近（单比特掩码）的相关性是不准确的。

后续能否发现新的区分器和新型的分析方法进一步评估 ChaCha 的安全性仍然是个值得继续深入研究的问题。

GPRS（General Packet Radio Services）是基于 2G 技术的移动数据标准，在 21 世纪初在世界范围内被广泛部署，提供了初代的移动网络连接。尽管一些国家将要废止或已经废止了 2G 技术，但是仍有一些国家依赖 GPRS 作为守备方案，因此其安全性依然关系着大量用户群。在无线环境中，攻击者通过在受害者附近监听通信进行窃听敌手攻击。为了防止运行在电话和基站之间的 GPRS 遭受窃听敌手攻击，标准采用流密码进行加密，并且设计了两个保密的流密码算法 GEA-1 和 GEA-2。GEA-1 的设计要求采用 64 比特密钥，同时要求算法应当满足当前的欧洲国家密码产品出口条例，而所获得的出口条例文件中提到了密钥穷搜的复杂度不超过 2^{40}。GEA-2 是后续算法，设计要求中不包含出口条例的要求。

2011 年 Nohl 和 Melette 报告了通过逆向工程得到的 GEA-1 和 GEA-2 的一些性质，并给出了 GEA-1 的状态恢复攻击的现场展示，但并没有公开攻击细节。2013 年，GEA-1 被禁用，

但 GEA-2 仍然是使用标准。尽管有披露的 GEA-1 的安全弱点，但业界仍缺失对 GEA-1 和 GEA-2 的公开的密码分析，人们仍然无法确认这两个算法的安全性。在欧密会 2021 上 Beierle 等人[8]的工作弥补了这一缺失，从不愿公开身份的来源处得到了 GEA-1 和 GEA-2 两个算法的设计，给出了第一个公开的密码分析结果。

GEA-1 流密码算法有 3 个线性反馈移位寄存器 A、B、C，级数分别为 31 比特、32 比特和 33 比特，以 Galois 模式运转；输出函数为一个代数次数为 4、输入为 7 个比特的非线性过滤函数 f。A、B、C 寄存器的初始化利用一个 64 比特的非线性反馈移位寄存器 S 进行，将 32 比特初始向量 IV、64 比特密钥 K 等吸收。作者发现了该算法的弱点：初始化过程之后 A、C 两个寄存器的 64 比特联合初始状态只能取自 2^{40} 种状态，而不是期望的 2^{64} 种状态。这个性质直接导致了分而治之状态恢复攻击，能够在已知 65 个密钥流比特（其中至少 24 个比特在同一数据帧中）的情况下以 2^{40} 次 GEA-1 运算的时间复杂度恢复寄存器 S 的状态。攻击需要预计算一个大小为 44.5GiB 的表，建表的时间复杂度是 2^{37} 次 GEA-1 计算。一旦这个表建好了，可以对每个新的 64 比特会话密钥在 2^{40} 次 GEA-1 运算的时间复杂度下进行攻击。作者在 64 核的集群上实现了该攻击，预计算和状态恢复共需约 1 个小时。文章还通过实验发现，对随机选择的线性反馈移位寄存器不太可能出现这样的弱点，这意味着 GEA-1 的弱点不太可能是偶然出现的，推断 40 比特的安全界与出口条例规定有关。

GEA-2 算法比 GEA-1 算法多了一个寄存器 D，初始化过程使用了更长的寄存器。对 GEA-2 算法，针对密钥流生成本身而不是初始化阶段进行了状态恢复攻击，方法是采用混合列表合并算法，并与代数技术结合。该攻击的时间复杂度是 $2^{45.1}$ 次 GEA-2 计算，存储复杂度约为 32GiB，需要在知道每帧所有的 1600 字节密钥流的情况下进行，无法利用多个帧的信息。

Beierle 等人在欧密会 2021 上的分析结果表明，GPRS 加密算法标准 GEA-1 只有 40 比特而非理想的 64 比特的安全性[8]，隐藏在 64 比特之下的 40 比特的安全界是为了获取必要的行业认证而设置的，用今天的一般的计算硬件即可攻破。GEA-2 算法中虽然没有了类似的弱点，但仍然没有达到 64 比特的安全性。现在 GPRS 的运营商已经较少使用 GEA-1 和 GEA-2 了，而手机基带芯片仍然支持这两个算法，这会导致攻击者通过伪基站触发选用 GEA-1 算法，在不能抗重放攻击的 GSM 认证中调用该算法可以使攻击者获取以前会话的密钥，从而破解之前使用更强算法加密的会话的内容。作者向有关组织提交了停止支持 GEA-1 和 GEA-2 算法的建议，在相关产品认证工作中得到采纳。

这篇论文的成果基于作者知道了两个外界未知的标准算法的设计，并利用密码分析方法准确挖掘到了疑似设计者隐藏在算法中的降低安全级别的技法，对现实世界的安全性有一定影响。

分组密码和密码置换函数往往可以通过许多不同的方式设计。一种流行的设计方法是采用类似 AES 的思路，即将比特以字节为单位划分，而所有的运算均以字节为基本单元完成，

称为对齐的设计（Aligned Approach）。这种方法能够方便密码研究者发现密码算法中类似超级 S 盒的层次结构并推断密码算法的差分或线性传播特征。相反，另外一种设计方式打破了字节级（或其他分组形式）的设计思路，其运算以比特为基本单元完成，如 Keccak 算法，称为非对齐的设计（Unaligned Approach）。该类设计丧失了类似 AES 的层次结构，因此需要复杂的计算机程序来研究搜索并证明差分和线性传播特征的安全性。

在美密会 2021 上，Bordes 等人[9]讨论了上述两种设计方法，形式化地定义了对齐（Alignment）设计，并具体地研究了分属不同设计策略的 4 个典型密码算法。作者提出了一种方法来研究线性和非线性之间的相互作用，以及对差分和线性传播的影响。通过计算机实验验证，作者系统地比较了 Rijndael、Saturnin、Spongent 和 Xoodoo 4 个算法。作者的研究表明，对齐（Alignment）自然会导致不同形式的聚类（Clustering）。例如，S 盒中的活跃比特、活跃模式中的两轮路线，以及差分和线性近似中的轨迹。作者证明 Rijndael、Saturnin 及 Spongent 是对齐设计，而 Xoodoo 不是。

中间相遇（MITM）攻击最早由 Diffie 和 Hellman 在 1977 年针对使用两个独立密钥的双重 DES 分组密码提出。在 40 多年的发展历程中，密码学者提出了各种增强 MITM 攻击的新技术，包括剪切缝合技术、部分匹配技术和初始结构技术等。然而，这些技术过于复杂，仅仅把它们描述清楚都是困难的。这导致最终只有少数学者掌握了相关方法。另外，即使完全理解这些技术的原理，实施它们也是困难的。应用这些方法的效果，极大地依赖于分析者的个人能力。

在文献[10-11]中，我国学者 Bao、Dong 等人系统化地用基于约束规划的语言刻画了中间相遇攻击的实质，将中间相遇密钥恢复攻击、原像攻击和碰撞攻击统一为一个如图 1 所示的闭合计算路径上的信息传播问题，并给出了字级对称密码中间相遇特征的局部传播规则和“中性字”自由度的确切计算公式。

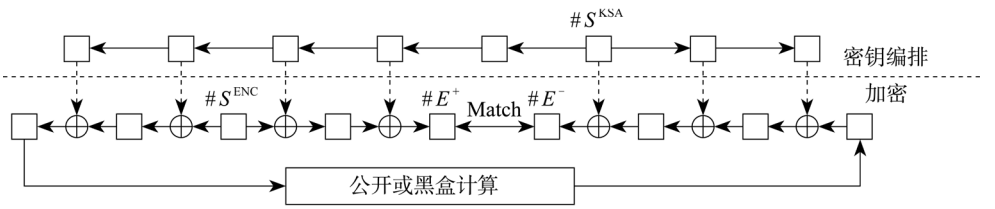

图 1　针对闭合计算路径的 MITM 攻击

这一工作彻底解决了字级中间相遇特征的自动化搜索问题。利用该方法对具体算法进行分析，发现了极为复杂的中间相遇特征。这些特征最大限度地利用了它所攻击目标的全部初始自由度，并巧妙地利用线性甚至非线性约束平衡了正向和反向计算路径的最终自由度，从而达到最优攻击复杂度。利用该方法，我国学者成规模地改进了一系列重要对称密码的攻击

结果，包括：①对 AES 杂凑模式的原像攻击突破了 7 轮的界限，打破了 20 年来对 AES 模式分析的记录；②第一个在经典计算模型下对 6 轮杂凑函数 Whirlpool（国际 ISO / IEC 标准）的攻击，这一结果打破了 10 年来学界对 Whirlpool 分析的记录；③破解了杂凑函数 Haraka-512 版本。

差分分析和线性分析均是两种最基本的密码分析方法。在欧密会 1993 上，Matsui 提出了线性分析方法，过去 30 年线性分析的有效性使得线性分析成为分组密码的一种基本分析方法，并且发展出了多种形式的线性分析方法。Kaliski 和 Robshaw 提出了利用多条独立线性逼近式的多线性分析方法。随后 Hermelin、Cho 和 Nyberg 提出了多维线性分析方法，还产生了利用非线性逼近表达式的线性分析方法。这些多样化的线性分析方法的模型通常基于独立性和马尔可夫链的假设下应用了堆积引理。另外，不同于利用堆积引理，通过 miss-in-the-middle 技术，Rijmen 和 Bogdanov 提出了零相关线性分析。1994 年，Langford 和 Hellman 将线性分析和差分分析结合提出了差分-线性分析方法。该方法将密码分为两部分，第一部分构造差分路径，第二部分构造线性路径。Leander 和 Nyberg 基于两部分独立的假设前提，提出了一种精确的相关度刻画方式。在欧密会 2019 上，Bar-On 等人指出两部分的相关性极大地可能影响攻击的复杂度，提出了一种通过差分-线性连接表（DLCT）的方法来考虑两部分的相关性。

我国学者 Liu 等人在文献[12]中提出了利用旋转差分特征替换差分-线性攻击中标准差分特征的方法，并证明这是差分-线性攻击的严格推广。这一“自然”的推广使得已有的复杂度估计方法失效。为此，作者考虑了线性特征部分的输出掩码为单位向量的特殊情形，并对这一情形给出了实用的复杂度评估方法，建立了自动化旋转差分-线性分析的模型。这篇文章利用该方法对以下算法进行了分析：①FRIET（由 AES 和 SHA3 设计者、著名密码学家 Joan Daeman 教授的团队设计，发表于 EUROCRYPT 2020）；②Xoodoo（由 AES 和 SHA3 设计者、著名密码学家 Joan Daeman 教授的团队设计， 发表于 FSE 2019，美国 NIST LWC 竞赛第 2 轮候选算法）；③Alzette（由著名密码学家、卢森堡大学密码团队带头人 Alex Biryukov 教授的团队设计，发表于美密会 2020 上，是美国 NIST LWC 竞赛第二轮候选算法 SPARKLE 的核心置换）；④SipHash（由美国著名密码学家 Bernstein 设计，是 OpenDNS 的核心安全算法，在工业界被广泛使用）。分析结果显示，旋转差分-线性攻击相对其他分析方法具有很好的优势，得到了针对 8 轮 FRIET、4 轮 Xoodoo 和 Alzette 的实际攻击。对 4 轮 Xoodoo 的攻击，获得了概率为 1 的区分器，而已有最好区分器的概率为 2^{-36}。

格式保留加密能够将特定格式的明文加密为相同格式的密文，如将 9 位整数的社会保险号加密为 9 位整数的密文。几种通用的方法可用来将可调分组密码转换为格式保留密码。但是当密码的作用域和格式保留的目标域差别较大时，这种转换方式是不高效的。因此一些定制的小版本上的 Feistel 结构的可调分组密码被提出，如美国 NIST 标准 FF3-1（NIST SP800-38G rev.1）和韩国标准 FEA-1 和 FEA-2（TTAK.KO-12.0275）。在美密会 2021 上，Beyne 发

现迭代使用 tweak 的小版本的 Feistel 结构的分组密码具有较强的线性相关度[13]。

在对 FF3-1 的线性分析中，Beyne 推广了传统的二元域上的线性逼近，给出了对 Z/N 群上的一般的线性分析理论。Beyne 利用线性逼近构造了χ^2区分器，并且提出了将χ^2区分器转化为消息恢复攻击的方法。利用多维线性分析方法约减了数据复杂度，最终给出了对 FF3-1、FEA-1 和 FEA-2 的区分攻击和消息恢复攻击的实际可实现的攻击。

线性分析的另一个研究方向是研究不同分析方法之间的关系。在 2021 年亚密会上，Beyne[14]还发现了多维零相关线性逼近和积分区分器的关系。一些分组密码抵抗不变子空间的能力较弱，Abdelraheem 等人发现了线性分析和不变子空间攻击之间的关系，研究发现不变子空间能够被描述为线性分析的相关矩阵的特征向量。

在亚密会 2021 上，Beyne 提出了线性分析的一种新解读，将线性逼近式推广为内积空间的子空间对。这些子空间可以对应到概率分布、集合、布尔函数等。在这种新的解读下，线性分析的很多扩展分析方法之间的关系能够更直观清晰地被阐明。例如，积分分析和零相关分析之间的关系，完美线性逼近和零相关线性逼近之间的差别，不变子空间和线性分析之间的关系。

在美密会 2021 上，Liu 等人[15]从代数角度给出了差分-线性分析的新解释，更好地研究了差分和线性分析两部分的结合相关性。Liu 等人通过引入差分代数过渡型（DATF）的新技术，发展了差分-线性分析的相关度评估的新理论和密钥恢复攻击新方法。Liu 等人将提出的新的方法应用在国际著名竞赛 CAESAR 的最终候选 Ascon 和 AES 竞赛的最终候选 Serpent 上，相对于 DLCT 连接表方法得到了更精确的相关度理论值。将此前 Serpent-128、Serpent-256 的差分-线性密钥恢复攻击结果进一步推进了一轮。从理论的角度重新理解或解读密码分析的统计模型是一个值得不断研究的问题，不仅能够检验假设的合理性，而且可能发现更好的密码学性质。

标准分组加密算法，如 DES 和 AES，是为了加密和解密固定长度的比特串设计的。但是在现实使用中，有时需要保持加密后数据的一些格式，如加密身份证号码或者银行卡号码时，我们希望密文能够和明文有相似的格式。否则，很多系统是无法处理密文的。1997 年，Brightwell 和 Smith 最早提出了保持格式加密（Format-Preserving Encryption，FPE）的概念。FPE 可以将一个空间（Domain）内的消息加密到同一个空间里。例如，16 位的银行卡卡号的空间是$[0,10^{16}-1]$，使用 FPE 加密一个卡号的密文一定是同一个空间$[0,10^{16}-1]$中的某个数字。FPE 在实际中非常有用，已经被一些大型跨国公司广泛使用。

从 FPE 的概念提出之后，一些 FPE 的具体算法也被相继提出。第一个支持 FPE 的算法是 AES 竞赛的候选算法之一“Hasty Pudding Cipher”。在 2002 年，Black 和 Rogaway 提出了提供 FPE 功能的 3 种方法：Cycle walking、prefix cipher 和一种基于 Feistel 的设计。Cycle walking 方法就是反复使用一个密钥迭代加密明文，直到某一次的密文落入预先定义的空间

内。2008 年，Spies 提出了一种基于 AES 算法、平衡 Feistel 模式（Balanced Feistel Network）和 Cycle walking 算法，称为 Feistel Finite Set Encryption Mode（FFSEM）。FFSEM 成为后来很多 FPE 算法的底层方法。

随后，许多研究组向 NIST 提交了 FPE 算法设计。Bellare 等人提出了 FFX 算法，后来被 NIST 称为 FF1；Vance 提出了 VAES3，后来被 NIST 称为 FF2；Brier、Peyrin 和 Stern 提出了 BPS 算法，其核心构件被称为 FF3。这些算法都是基于 Feistel 模式的分组密码算法。2016 年，美国标准技术局发布了 SP800-38G 文件，规定了 FF1 和 FF3 为两种保持格式加密算法。

对 FF3 算法的首个分析中提出了一种针对小明文空间（Small Domain）高效恢复明文的攻击方法。在美密会 2017 上，Durak 和 Vaudenay 对 FF3 提出了一种新的滑动攻击算法。这种攻击方法可以在不知道密钥情况下计算新的密文。它的攻击复杂度是 $O(N^{11/6})$（明文空间的大小假设为 N^2），时间复杂度为 $O(N^5)$。这个攻击使用了 FF3 的一个性质：FF3 的 tweak-key 扩展算法容易导致一个相关 tweak 攻击，使攻击者可以直接攻击 4 轮的 FF3 算法而不是 8 轮，这种攻击被称为第一代攻击。

为此，美国标准技术局修改了 FF3 的 tweak-key 生成算法，修改之后的算法被称为 FF3-1。尽管如此，学术界仍然对 FF3 的安全性保持兴趣，在欧密会 2019 上，Hoang、Miller 和 Trieu 使用了第二代滑动攻击，将之前的第一代攻击的时间复杂度大幅提升到 $O(N^{17/6})$，而数据复杂度保持不变($O(N^{11/6})$)。

Amon 等人在欧密会 2021 上提出了对 FF3 的第三代通用攻击[16]。使用对称滑动攻击（Symmetric Slide Attack），他们将数据和时间复杂度提升为 $O(N^{7/4})$和 $O(N^{5/2})$。使用非对称滑动攻击，数据复杂度被提升为 $O(N^{3/2})$（时间复杂度仍然为 $O(N^{5/2})$）。这是首次数据复杂度小于明文空间大小 N^2。此外，他们还利用循环结构的滑动攻击（Slide Attack Using the Cyclic Structure）给出了 FF3 和 FF3-1 的实际复杂度的区分攻击。

立方攻击（Cube Attack）是由 Dinur 和 Shamir 提出的一项强大的用于分析流密码算法的有效手段。立方攻击有两个主要阶段，在第一阶段，也称为线下阶段，攻击者需要找到合适的立方，保证这些立方的超级多项式（Superpoly）是关于密钥变量的线性多项式，并且把这些超级多项式的具体表达式恢复出来；在第二阶段，也称为线上阶段，攻击者对于真实的密钥是未知的，但他可以通过访问加密机，计算出在真实的密钥下超级多项式的值，从而建立一个关于密钥变量的线性方程组。通过对这个方程组求解，攻击者可以获取到一部分密钥比特的信息，而剩下的密钥比特可以通过穷搜的方法恢复出来。从立方攻击提出开始，提出了许多改进的攻击方案，如立方测试器（cube Testers）、动态立方攻击、条件立方攻击、基于 Division Property 的立方攻击和相关立方攻击等。

对于立方攻击中关键的一步——恢复给定 cube 的超级多项式（Superpoly），在早期的立方攻击中，分析人员使用实验的方法来恢复。他们首先通过一些线性测试，确定出超级多项

式是线性的，然后通过加密 cube 集合并加和输出的比特得到一些密钥的线性方程，最终通过解这些线性方程来得到密钥信息。这种方法只可以恢复非常简单的超级多项式，立方攻击能达到的轮数也是非常有限的。

在美密会 2017 上，Todo 等人提出了使用分离特性（Division Property）辅助超级多项式的恢复。通过搜索特定的分离特性，他们可以确定出一定不存在超级多项式中的密钥比特。之后，他们对可能存在超级多项式中的密钥比特建立真值表，从而使用真值表恢复出超级多项式的精确形式。这种方法可以用于恢复更加复杂的超级多项式。在美密会 2018 上，Wang 等人改良了 Todo 等人的方法，使得可以恢复更加复杂的超级多项式。在亚密会 2019 上，Wang 等人提出一种使用三子集的分离特性（Division Property with Three Subsets）精确恢复超级多项式的理论方法，并且验证了之前一些论文中的理论恢复的超级多项式（其实是区分器）。在欧密会 2020 上，Hao 等人提出了不带未知集合的分离特性的方法（Division Property Without Unknown Subsets），首次可以实际恢复出精确的超级多项式。在亚密会 2020 上，Hu 等人从多项式中单项式的传播对分离特性进行了分析，给出了不依赖多重集合就可以推导的易于理解的传播规则，称为单项式预测技术，并且使用分而治之的方法，恢复出了更复杂的超级多项式。这些方法都只可以恢复相对简单的超级多项式，对于复杂的超级多项式无能为力。

Hu 等人[17]在亚密会 2021 上给出了一种使用嵌套模式的单项式预测技术的方法，可以恢复非常复杂的超级多项式。利用这个方法，他们提升了 Trivium、Grain128A 和 Kreyvium 的立方攻击。

尽管基于立方攻击，已经提出了很多密钥恢复的攻击手段，但是这些攻击的重点是如何增加攻击的轮数，因此很少有攻击能够在实际可行的时间复杂度下恢复出 Trivium 的 80 比特密钥。上一次对 Trivium 密钥恢复实际可行的攻击，是在 2013 年的 FSE 上由 Fouque 和 Vannet 提出的对 784 轮 Trivium 的攻击。实际可行的密钥恢复攻击需要足够数量的低次数的超级多项式，但是如何构造一个有用的立方一直以来都是立方攻击中的一个难题。当 Shamir 提出 767 轮 Trivium 的立方攻击的时候，为了找有线性超级多项式的立方，采用了随机游走的策略。这个方法先是随机选取一个立方，如果超级多项式是常数，就随机移除一个 IV 变量；如果超级多项式是非线性的，就随机加入一个 IV 变量。这个过程一直持续到找到有线性超级多项式的立方。

Ye 和 Tian 在亚密会 2021 上首次给出了在实际可行复杂度下，对 805 轮 Trivium 的密钥恢复攻击[18]。他们提出了一系列启发式的算法，来寻找具有线性超级多项式的立方。第一，他们提出了一个算法来构造具有低次数的超级多项式的立方，基本思路是先选取一个小的立方，然后慢慢地往立方中添加变量，从而不停地降低超级多项式的代数次数。第二，由于 Trivium 的输出比特是由 6 个状态比特异或得到的，因此他们认为很有可能线性超级多项式仅仅是由其中一个状态比特产生的。他们通过实际实验验证了这个猜想，并且提出了一个算

法来预测这个关键的状态比特，预测的准确率达到了 75.3%。关键状态比特的成功预测，可以为第一步中小的立方的选取提供思路。第三，在立方攻击中，莫比乌斯变换可以用于同时检验一个大的立方的子立方是否具有线性超级多项式，但是这需要耗费巨大的内存。为了减少内存的消耗，他们发现在莫比乌斯变换的过程中，ANF 中一部分项的系数可以提前被计算出来。换言之，可以通过只进行一部分的莫比乌斯变换来进行线性检测。这个方法成功将原本 1024GB 的内存复杂度降低到了 9GB。第四，他们将这些技术运用于 805 轮 Trivium，成功地找到了 42 个线性无关的线性超级多项式，并恢复了 42 比特的密钥信息，而剩下的 38 比特密钥可以直接穷搜得到。用一台装载 GTX-1080 GPU 的电脑，在几个小时内他们成功地恢复了 805 轮 Trivium 的密钥。

SIMON 算法是美国国家安全局（NSA）提出的轻量级分组密码算法，由于其设计文档中没有给出安全性分析，且由于美国国家安全局的身份特殊性，学术界对 SIMON 算法的安全性一直存在顾虑。SIMECK 算法是 SIMON 算法的改进版本，这两个算法的轮函数相似。

综合目前对 SIMON 和 SIMECK 的分析工作来看，针对 SIMON 和 SIMECK 这类密码的最佳攻击为差分、线性类分析。差分和线性攻击作用在 SIMON 和 SIMECK 上时表现出了差分路径和线性路径的强聚集效应，也就是说，存在许多具有相同输入-输出的轨迹。聚集之后的差分概率（或线性路线的相关度）明显高于最佳差分特征的概率（最佳线性路线的相关度）。为了估算差分和线性路线聚集之后的概率和相关度，一个流行的方法是将尽可能多的具有相同的输入-输出路径堆叠在一起。

Leurent、Pernot 和 Schrottenloher 在亚密会 2021 上提出了一种更为通用的研究差分、线性聚集效应的方法[19]。他们没有用给定的输入-输出建立一个轨迹列表并把概率（相关度）堆叠，而是考虑一类高概率轨迹，其中活跃比特保持在固定的位置窗口（w 位）。特别是他们还观察到以前大多数攻击中使用的差分和线性壳都适用于此框架。

利用 SIMON 类密码轮函数的性质，他们高效地通过在这个空间上的差分转移矩阵的乘法计算了概率，同样通过线性相关矩阵计算了相关度。这提供了一个差分概率和线性相关度的更严格的下限，即对比在以前的工作中使用的区分器，他们找到了更高概率区分器。这是因为这种方法考虑了所有中间状态的活跃比特在窗口中的所有轨迹。Leurent、Pernot 和 Schrottenloher 利用这种方法得到了输入-输出单比特活跃的高概率差分路线和高相关度线性壳。利用这些更强大的区分器，他们把 SIMECK 和 SIMON 的最优攻击提升了 3~6 轮。

Shor 量子算法可以在量子计算模型下以多项式时间复杂度分解大整数和求解离散对数，从而威胁到目前主流公钥密码算法。近年来，密码学者开始研究如何利用量子计算算法分析对称密码，其中利用量子计算攻击哈希函数时无须访问带密钥的量子谕言，使其成为备受关注的研究课题。在经典模型下，对 n 比特输出的理想哈希函数的最优通用碰撞攻击（Generic Attacks）为 rho 算法，其时间复杂度为 $T=2^{n/2}$，并且不需要存储。当有 S 台计算机并行计算碰

撞时，时间复杂度为 $T=2^{n/2}/S$，因此当攻击小于 $T\cdot S=2^{n/2}$ 时，即认为该攻击优于通用碰撞攻击，攻击有效。在量子计算模型下，该攻击界（$T\cdot S=2^{n/2}$）仍然适用，即如果存在量子计算攻击优于该界，即认为攻击有效。该通用攻击界通常称为时空权衡（Time-Space Tradeoff）成本。

在美密会 2021 上，Hosoyamada 和 Sasaki 首次研究了针对 SHA-256 和 SHA-512 的专用量子碰撞攻击[20]。攻击分别达到 38 步和 39 步，显著改进了 31 步和 27 步的经典碰撞攻击。两种攻击都采用了先前工作的框架，将许多半自由起始碰撞（Semi-Free-Start Collision）攻击转换为两消息块碰撞攻击，并且在时空权衡的成本度量上比通用攻击更快。Hosoyamada 和 Sasaki 观察到在量子计算条件下，可以显著减少所需的半自由启动碰撞次数，这使其能够将之前经典的 38 步和 39 步半自由起始碰撞攻击转换为碰撞攻击。该攻击背后的想法很简单，也将适用于其他密码哈希函数的量子碰撞攻击。

在亚密会 2021 上，Dong 等人[21]利用相关密钥差分路线构造基于反弹技术（Rebound Technique）的经典和量子碰撞攻击，并构造了基于混合整数线性规划的自动化搜索模型，搜索适用于构造经典和量子碰撞攻击的特征。该攻击的想法是利用密钥自由度，提高差分路线的概率。在自动化模型建立过程中，有效地解决了某些哈希函数存在相关密钥差分路线多轮不兼容的情形，自动化地搜索出有效的攻击路线。基于上述想法，Dong 等人首次给出 ISO 标准哈希函数 Whirlpool 的 9 轮自由起始量子碰撞攻击（Free-Start Quantum Collision），将 Whirlpool 压缩函数（全轮为 10 轮）的抗碰撞攻击界降低至仅 1 轮。同时，该论文还给出了 Saturnin-hash 和 SKINNY 的多个经典和量子分析结果。

近来，量子周期寻找算法被广泛地应用于量子叠加态询问模式（Quantum Superpositi on State Interrogation Mode）下针对分组密码算法的量子攻击中，并破解了 Even-Mansour 结构以及多种消息认证码（MAC）与认证加密模式（AE）。这些攻击都是通过构造性的方法建立关于单块输入的周期函数，并利用 SIMON 算法恢复该隐藏的周期从而完成密钥恢复或区分攻击。

在亚密会 2021 上，Bonnetain 等人[22]提出了一种称为量子线性攻击的新型攻击方法，可以在量子叠加态询问模式下利用 SIMON 算法对多种 MAC 进行攻击。该方法通过将每块输入的两个特定取值映射为一个布尔变量，并利用多块输入将 MAC 算法的输出映射为一个隐含线性结构的周期函数。虽然该周期函数在每次询问时的表达式均不同，但是其周期保持不变。而利用 SIMON 算法只需要单次询问就能得到一个与该周期正交的向量。因此，通过多项式次的询问即能有效地恢复该周期，从而实现伪造攻击。

此外，文中还提出了基于上述量子线性攻击思想的几类变形的攻击方法，这些方法有效地将 Deutsh 算法、Bernstein-Vazirani 算法及 Shor 算法运用于攻击中，破解了 LightMac、PMAC 等可并行 MAC 算法，以及一些具有经典 BBB（Beyond-Birthday-Bound）安全强度（LightMAC+，PMAC+）或使用可调分组密码（ZMAC）的变形 MAC 算法。文中还指出构

造可并行的量子安全的伪随机函数可能会是一个十分具有挑战性的工作。

上述研究首次将 Deutsch 算法、Bernstein-Vazirani 算法及 Shor 算法应用于伪造攻击与密钥恢复攻击中，在量子叠加态询问模式下展现了很好的攻击效果。后续，能否在经典询问模式下，利用上述量子算法进行有效的量子攻击，仍然是对称密码分析领域的一个有待研究的问题。

超奇异同源密钥封装（Supersingular Isogeny Key Encapsulation，SIKE）是 NIST 后量子密码竞赛的 15 个密钥封装和数字签名优胜方案之一，其特点是在所有优胜方案中有最小的公钥大小，且是唯一的基于同源的方案。其安全性依赖于计算超奇异同源（Computational Supersingular Isogeny，CSSI）问题的困难性。2019 年及之后的研究表明，实际中求解 CSSI 问题的最好的经典算法是 vOW（van Oorschot-Wiener）并行碰撞搜索算法，应当利用这个算法建立 SIKE 的后量子安全性并选取参数。但是这些安全性估计的局限性是武断地限制了敌手可获得的存储空间，转化为经典门计数下的攻击开销时能否达到NIST 的要求还比较模糊。

密码方案的安全性评估往往直接从攻击开销得出，攻击开销通常通过其询问复杂度衡量，或者转化成攻击的时间复杂度、执行攻击的指令或周期数等。例如，SIKE 的安全界直接从 vOW 算法的询问复杂度或执行同源计算的指令数得出。这种方法的一个主要的缺点是忽略了存储开销，无法衡量在 SIKE 的利用 vOW 算法的攻击中大量共享存储器的开销。另外，怎样精确计数实际攻击所需的门数、指令数和周期数也不清楚。一种比较实际的开销模型是基于预算的开销模型，攻击者将固定的预算用于获取必要的计算和存储资源，使得攻击所需的时间最小化，方案的安全性即由攻击所需的时间给出。

Longa 等人[23]在美密会 2021 上对 SIKE 在基于预算的开销模型下进行了安全性分析，攻击过程考虑了实际的计算和存储开销。在基于预算的开销模型中对半导体和存储器的历史价格和未来价格都进行了分析。为了确定模型中实际的硬件花费，作者为 SIKE 攻击算法中最重要的两个操作——$\mathbb{F}_{p^2}$上的乘法运算和同源计算特别定制了 ASIC 友好的硬件加速器。通过 ASIC 综合获取了电路面积和时间，评估了 SIKE 上 vOW 算法攻击的实际花费和相应的运行时间，对 SIKE 第三轮参数和引入的新参数集进行了安全性评估。为了验证该设计的可靠性，作者在 FPGA 上实现了 vOW 算法的概念验证的硬件/软件设计。作者将这个开销模型扩展到 AES 和 SHA-3 算法，得到了这些原语与 NIST 后量子密码竞赛相关的更实际的安全性估计。这项分析和目前对 SIKE 的量子安全性的分析表明，目前 SIKE 方案的参数选取是保守的，实际上以较高的冗余提供了比想要达到的 NIST 安全级别要求更高的安全性，解决了 SIKE 参数实际安全性评估的公开问题。此外，作者还给出了三组新的参数，在满足安全要求的同时达到更高效的实现。

密码方案的安全性一般通过攻击的渐进复杂度给出，表示为安全参数的函数表达式。然而人们更想获得一个更直观的攻击难度的估计，基于预算的攻击开销就是给出这样一种分析

结果，得到的结论是“在预算为多少美元时攻破该方案需要几年”。这种更直观的安全性评估不仅需要密码分析者的智力贡献，还受攻击所需的硬件价格影响。Longa 等人从基于预算的攻击开销角度给出了 SIKE 后量子方案、AES 算法和 SHA-3 算法的安全性评估及其对比，明确反映了 NIST 后量子竞赛安全级别的要求。其开销模型和分析可以用于其他密码方案和原语，对 NIST 后量子密码标准化过程的其他候选算法产生影响。

在美密会 2019 上，Gohr 提出了一种基于机器学习算法的新型密码分析策略。针对分组密码 SPECK，Gohr 构建了一个基于神经网络的区分器。在相同攻击轮数上，该神经网络区分器超越了最先进的密码分析工作。虽然该工作为机器学习辅助密码分析开辟了新的可能性，但目前尚不清楚这种区分器实际上是如何工作的，以及机器学习算法推导出的信息是什么。攻击者留下了一个黑匣子，它并不能说明所测试算法可能存在的弱点的性质，因为深度神经网络的可解释性是一项众所周知的艰巨任务。

在欧密会 2021 上，Benamira 等人[24]对这种新的神经网络区分器的内在工作原理进行了详细的分析和透彻的解释。首先，作者研究了分类集，并试图找到一些可以指导我们更好地理解 Gohr 结果的模式。他们通过实验表明，神经区分器通常依赖于密文对的差分分布，但也依赖于倒数第二轮和倒数第三轮的差分分布。为了验证该发现，作者构建了一个基于纯密码分析的 SPECK 密码区分器，不使用任何神经网络，它达到了与 Gohr 的神经网络区分器基本相同的精度和相同的效率（因此改进了以前的基于非神经网络的区分器）。其次，作者提供了一种基于机器学习的新型区分器，它将 Gohr 的深度神经网络剥离到最低限度。使用简单的标准机器学习工具，能够保持 Gohr 区分器的准确性。该工作表明 Gohr 的神经网络区分器实际上在学习阶段内在地构建了一个非常好的密码差分分布表（DDT）的近似值，并使用该信息直接对密文对进行分类。该结果具备可解释性，并且其本身代表了对深度神经网络可解释性的有趣贡献。最后，作者提出了一些方法来改进 Gohr 的工作和可能的新神经网络区分器参数设置。

本节作者：乔珂欣（北京理工大学）、胡凯（南洋理工大学）、董晓阳（清华大学）、郝泳霖（北京信息科学技术研究院）、黄震宇（中国科学院信息工程研究所）、闫海伦（中国科学院大学）、史丹萍（中国科学院信息工程研究所）、孙思维（中国科学院大学）

参考文献

[1] LIU F, SARKAR S, MEIER W, et al. Algebraic Attacks on Rasta and Dasta Using Low-Degree Equations[C]. ASIACRYPT 2021(1): 214-240.

[2] BANIK S, BAROOTI K, VAUDENAY S, et al. New Attacks on LowMC Instances with a Single Plaintext/Ciphertext Pair[C]. ASIACRYPT 2021(1): 303-331.

[3] LIU F, ISOBE T, MEIER W. Cryptanalysis of Full LowMC and LowMC-M with Algebraic Techniques[C]. CRYPTO 2021(3): 368-401.

[4] PEYRIN T, WANG H. The MALICIOUS framework: embedding backdoors into tweakable block ciphers[C]. CRYPTO 2020(3): 249-278.

[5] SHEN Y, WANG L, GU D, et al. Revisiting the Security of DbHtS MACs: Beyond-Birthday-Bound in the Multi-user Setting[C]. CRYPTO 2021(3): 309-336.

[6] COUTINHO M, TERTULIANO C, NETO S. Improved Linear Approximations to ARX Ciphers and Attacks Against ChaCha[C]. EUROCRYPT 2021(1) : 711-740.

[7] DEY S, DEY C, SARKAR S, et al. Revisiting cryptanalysis on ChaCha from crypto 2020 and eurocrypt 2021. Cryptology ePrint Archive, Report 2021/1059 (2021).

[8] BEIERLE C, DERBEZ P, LEANDER G, et al. Cryptanalysis of the GPRS Encryption Algorithms GEA-1 and GEA-2[C]. EUROCRYPT 2021(2): 155-183.

[9] BORDES N, DAEMEN J, KUIJSTERS D, et al. Thinking Outside the Superbox[C]. CRYPTO 2021(3): 337-367.

[10] BAO Z, DONG X, GUO J, et al. Automatic Search of Meet-in-the-Middle Preimage Attacks on AES-like Hashing[C]. EUROCRYPT 2021(1): 771-804.

[11] DONG X, HUA J, SUN S, et al. Meet-in-the-Middle Attacks Revisited: Key-Recovery, Collision, and Preimage Attacks[C]. CRYPTO 2021(3): 278-308.

[12] LIU Y, SUN S, LI C. Rotational Cryptanalysis from a Differential-Linear Perspective - Practical Distinguishers for Round-Reduced FRIET, Xoodoo, and Alzette[C]. EUROCRYPT 2021(1): 741-770.

[13] BEYNE T. Linear Cryptanalysis of FF3-1 and FEA[C]. CRYPTO 2021(1) : 41-69.

[14] BEYNE T. A Geometric Approach to Linear Cryptanalysis[C]. ASIACRYPT 2021(1) : 36-66.

[15] Liu M, Lu X, Lin D. Differential-Linear Cryptanalysis from an Algebraic Perspective[C]. CRYPTO 2021(3) : 247-277.

[16] AMON O, DUNKELMAN O, KELLER N, et al. Three Third Generation Attacks on the Format Preserving Encryption Scheme FF3[C]. EUROCRYPT 2021 (2): 127-154.

[17] HU K, SUN S, TODO Y, et al. Massive Superpoly Recovery with Nested Monomial Predictions[C]. ASIACRYPT 2021(1): 392-421.

[18] Ye C, Tian T. A Practical Key-Recovery Attack on 805-Round Trivium[C]. ASIACRYPT 2021(1): 187-213.

[19] LEURENT G, PERNOT C, SCHROTTENLOHER A. Clustering Effect in Simon and

Simeck[C]. ASIACRYPT 2021(1): 272-302.

[20] HOSOYAMADA A, SASAKI Y. Quantum Collision Attacks on Reduced SHA-256 and SHA-512[C]. CRYPTO 2021 (1): 616-646.

[21] DONG X, ZHANG Z, SUN S, et al. Automatic Classical and Quantum Rebound Attacks on AES-Like Hashing by Exploiting Related-Key Differentials[C]. ASIACRYPT 2021(1): 241-271.

[22] BONNETAIN X, LEURENT G, NAYA-PLASENCIA M, et al. Quantum Linearization Attacks[C]. ASIACRYPT 2021(1) : 422-452.

[23] LONGA P, WANG W, SZEFER J. The Cost to Break SIKE: A Comparative Hardware-Based Analysis with AES and SHA-3[C]. CRYPTO 2021(3): 402-431.

[24] BENAMIRA A, GERAULT D, PEYRIN T, et al. A Deeper Look at Machine Learning-Based Cryptanalysis[C]. EUROCRYPT 2021(1): 805-835.

理论基础

1　安全模型

1.1　组合框架

在定义安全性时，一般考虑不诚实方不能做什么，但是对于某些特定的应用，需要赋予不诚实方特定的能力才能保证安全，如指定验证者签名（DVS），需要赋予不诚实方可以伪造与原始签名不可区分的签名的能力。对于这种应用，一般使用的可组合安全（CC）框架不能很好地定义其安全性。具体来说，CC 框架通过设计一个理想世界并证明与现实世界不可区分来定义安全性（其中理想世界与现实世界都以集合的形式表达），这就需要首先设计一个理想函数，然后将模拟器连接到这个理想函数上，包括对不诚实的一方可用的理想输入和输出，并向环境提供“真实的”输入和输出（与现实世界的输入和输出不可区分）。因为模拟器覆盖了理想函数所有与不诚实方的接口，它只能限制不诚实方的能力，而无法为不诚实方提供特定的能力，因为这些可能会被模拟器阻止。

Maurer 等人[1]给出了一种针对上述特定应用的基于 CC 框架的安全性建模。更具体地，Maurer 等人给出了针对多指定验证者签名（MDVS）的可组合安全性定义，其中 MDVS 是 DVS 的一般化方案。尽管传统的 CC 框架无法为不诚实方提供特定的能力，但是 Maurer 等人注意到之前的一些文献指出，在 CC 框架中模拟器不一定要覆盖理想函数所有与不诚实方的接口，再结合 CC 框架中方案安全的定义：“现实世界是理想世界的子集”，Maurer 等人针对 MDVS，将其理想世界定义为分别具有两种特定性质（认证性质与伪造性质）的理想世界的交集。若现实世界是这个交集的子集，则 MDVS 方案安全。在此之前，虽然 Damgård 等人[2]对 MDVS 定义了基于游戏的较强的安全性，但是相对于 Maurer 等人提出的可组合安全性，Damgård 等人定义的安全性有些过强，即强于实际所需要的安全性。具体来说，满足 Damgård 等人定义的安全性的 MDVS 方案可以用于构造 Maurer 等人提出的可组合安全性中的理想世界，但是满足可组合安全性的 MDVS 方案并不能满足 Damgård 等人研究工作中定义的安全性。

代数群模型（AGM）是通用群模型（GGM）更有表达性的一个版本，其针对的敌手与 GGM 相似，都是通过一系列“通用”群运算来计算群元素，但相较于 GGM，AGM 允许在

计算过程中使用实际的比特串来表达群元素。AGM 已被用来证明各种数论假设的等价性，以及证明了 SNARKs 和盲签名的安全性等。值得注意的是，上述结论中并没有提供任何当其与其他协议组合时的安全性保证，为此，Abdalla 等人[3]在通用组合（UC）框架下形式化了 AGM 的定义（UC-AGM），并证明了一个恰当的组合理论。首先，为了定义 UC-AGM，Abdalla 等人将敌手、环境以及模拟器都定义为代数的，其中代数的是指将某一群元素表达为一系列群元素（以基的形式）的连乘，即$h = \Pi g_1^{x_1} \cdots g_n^{x_n}$，这一定义将敌手实体（现实与理想的敌手、环境）与非敌手实体（诚实的参与方、理想函数）区分开来。此外，Abdalla 等人还要求在现实世界中敌手代数地向诚实参与方以及理想函数（用于带有现实世界的混合证明中）传输群元素，而且环境需要代数地向敌手传输群元素；反之不成立。Abdalla 等人提出的 UC-AGM 的组合理论与 UC 的组合理论也有不同点：只有当敌手和环境在现实世界中是代数的时，模拟器才会工作。值得注意的是，在理想世界中也必须对应地保证敌手与环境是代数的，不过这一点在具体证明中不算是一个限制，因为当交互过程中不存在群元素时，无须令敌手与环境是代数的。Abdalla 等人也基于 UC-AGM 给出了几个比较重要的协议的证明，这些协议之前是没有 UC 框架下的安全性证明的，包括：

（1）用于不经意传输的 Chou-Orlandi 协议；

（2）用于基于密码的认证密钥交换的 SPAKE2 和 CPace 协议。

1.2 基于游戏

基于游戏的安全性定义可分为搜索型与决定型。其中，搜索型的安全等级可以直观地由攻击者攻破方案所需的开销来衡量，即对于 λ 比特安全的密码方案，若攻击者的计算开销为 T，攻击成功的概率为ε，则方案安全当且仅当对于任意的攻击，有$T/\varepsilon \geq 2^{\lambda}$。但是对于决定型游戏，这种衡量方法并不能直接被代入，因为决定型游戏是否安全取决于攻击者成功猜测选择比特的概率是否接近于$1/2$。

Micciancio 和 Walter 的工作[4]为搜索型和决定型游戏设计了一个统一的安全性框架，但是他们的理论结果并不能被具体实验验证。基于此，Watanabe 和 Yasunaga[5]重新回顾了比特安全的理论概念并给出了一种具有实际操作意义的比特安全定义，该定义同时考虑了搜索型游戏和决定型游戏。具体来说，Watanabe 和 Yasunaga 采用了两种敌手：内部敌手和外部敌手。其中内部敌手是在常规的游戏框架下定义的，外部敌手通过多次调用内部敌手来提高赢得游戏的概率。对于搜索型游戏，Watanabe 和 Yasunaga 沿用了一般的衡量规则；对于决定型游戏，Watanabe 和 Yasunaga 通过多次调用内部敌手，对内部敌手计算阶为 12 的 Rényi 散度（一种信息度量）来衡量赢得游戏的概率。需要注意，在结构上，搜索型与决定型有所不同：对于搜索型，在重复的游戏中敌手每次收到相互独立的挑战，并且只要敌手发现任一可行的解法，敌手就会赢得游戏；对于决定型，在重复的游戏中敌手需要对同一个选择比特进行猜

测。为了验证该安全性定义的可行性，Watanabe 和 Yasunaga 给出了几种在此定义下的安全归约，包括伪随机生成器归约到单向函数、Goldreich-Levin 核心谓词归约到单向函数、DDH 归约到 CDH。这些归约的结论与 Micciancio 和 Walter 的理论结论大部分相符，除了 Goldreich-Levin 核心谓词归约到单向函数有一点偏差：Watanabe 和 Yasunaga 的工作表明最优归约只针对“平衡”的敌手。因此，Watanabe 和 Yasunaga 工作中的比特安全框架在 Micciancio 和 Walter 的基础上额外赋予了可操作性。此外，Watanabe 和 Yasunaga 还给出了 IND-CPA 安全性与加密方案单向性之间的归约关系：当一个加密方案是 λ 比特安全的 IND-CPA 且消息空间大小为 2λ 时，它具有$s(\lambda - O(1))$比特安全的单向性。

1.3　模块化安全规范

现有可证明安全领域，密码方案的安全分析通常分为基于模拟和基于游戏两种类型。由于在实际安全分析的过程中，需要将协议要求、环境和假设等不同方面同时进行考虑，因此基于这两种类型的安全分析方式，其分析过程会很复杂，且容易出错。文献[24]提出了一种称为模块化安全规范（Modular Security Specifications，MoSS）的安全分析框架，该框架将协议应实现的安全要求（目标）与确保每个要求的模型（假设）清晰地分开。这种模块化使我们能够跨不同的协议和任务重用单个模型与需求，并比较同一任务的协议，无论是在不同的假设下还是满足不同的需求集。具体而言，在 MoSS 中，安全规范包括一组模型（假设）和特定需求（目标），模型和需求是使用谓词和概率函数定义的，通过分别定义每个模型和需求，允许模块化、标准化和重用；此外，MoSS 还包括一个定义明确的执行过程。由于应用协议通常包含大量的需求和模型，因此这种模块化对于应用协议特别有利。MoSS 是灵活且可扩展的。例如，它可以支持渐进和具体的安全定义。本文利用 MoSS 分别对安全广播协议和 PKI 方案进行了安全分析。

2　陷门函数

陷门函数（TDF）的引入奠定了传统公钥密码学的基础，一直以来人们习惯性地将 TDF 的应用场景限制在传统公钥密码中，直到最近的一些工作结果表明：TDF 的应用并不只限于传统公钥密码。这是因为 TDF 可以还原全部的输入，包括随机数，而公钥密码方案还原的输入不包括随机数。

基于此，Garg 等人[6]研究了如何针对更先进的公钥密码构造 TDF，如 ABE 和 IBE，分别称为 AB-TDF 和 IB-TDF。AB-TDF 主要针对一些不能单独由 ABE 或 TDF 得到的应用，如指定验证者非交互零知识证明（DV-NIZK）；IB-TDF/AB-TDF 可以直接得到安全的确定性加密方案，以及可以用于构造低级的 IBE/ABE 方案。目前的 TDF 的构造大多是使用了特定的构

造技术。因此，如何使用这些构造方法得到更先进的原语的 TDF 并不明确。Garg 等人探索了新的构造方法，给出了一种构造针对先进原语的 TDF 的通用编译方法：给定任意的 IBE/ABE/PE 和伪随机密文，返回一个安全的 IB/AB/P-TDF。同时，该方法延续了原有加密方案的安全性。例如，如果给定的 ABE 的安全性是单密钥的，那么得到的 AB-TDF 也是单密钥的。这一点使得可以在加密方案构造完成后再加入 TDF，不需要重新分配密钥，只需要添加一些额外的公共参数即可。在技术层面上，该方法只额外用到了暗示伪随机生成器（Hinting PRG）。此外，Garg 等人开创了对于陷门投影混淆电路的研究，即还原混淆电路所用的随机数。Garg 等人给出的陷门投影混淆电路方案基于 DDH 或 LWE，利用了密钥相关消息（KDM）和随机数相关消息（RDM）技术的相互作用，并且可以直接得到一个 AB-TDF 方案，该 AB-TDF 方案相对于上述构造的 AB-TDF 方案而言，不仅满足完善正确性，还拥有投影密钥且满足适应性安全。

3 数域筛法

虽然后量子的研究一直在发展，但离散对数问题仍然是当前许多已经应用的公钥协议安全性的基础。离散对数问题的困难性与所在的群相关，所在群的选择一般有两种：有限域中的可逆元群与椭圆曲线上的点群。对于第一种选择，一般根据有限域的特征分类进行具体讨论，即根据所使用算法的计算复杂度将有限域的特征分为小、中、大 3 种情况。当有限域的特征比较小时，由于拟多项式时间算法的存在，一般不被讨论。当有限域的特征为中和大的情况时，数域筛法（NFS）及其变体是目前已知解决离散对数问题最快的算法。对于某些扩张的度为合数的域，变体中的塔数域筛法（TNFS）有比 NFS 更高的效率，其主要是充分利用了与 NFS 不同的目标域表达式。

Kim、Barbulescu[7]和 Jeong[8]对 TNFS 进行了进一步优化，得到了扩展的塔数域筛法（exTNFS）。Micheli 等人[9]发现与 NFS 不同的是，exTNFS 对于中特征的有限域的计算复杂度比大特征的有限域的计算复杂度要低，这意味着 exTNFS 可能在计算记录上会有优势（计算记录可以反映有限域的安全性）。衡量一个有限域的安全性并不是一件容易的事情，因为并不知道具体用哪种算法会快一些，不过，一个固定的扩张的度渐进地定义了有限域比较大的特征，在这种情况下，一般会选择 NFS 算法。同时，有限域特征的大小又是由算法的复杂度定义的，如果确定了要研究的有限域，那么无法衡量其特征的大小。Micheli 等人通过改进 exTNFS 算法给出了针对这一问题的实用见解。具体来说，Micheli 等人给出了第一个用改进的 exTNFS 记录 512 比特有限域$\mathbb{F}_{p^6}$的计算记录。其具体的改进包括以下 2 点。

（1）算法是在高维球体中而不是在棱正交的多胞形中进行筛选的，因为当维数大于等于 3 时，高维球体比棱正交的多胞形更准确。

（2）算法使用了格枚举法来进一步适应新的搜索空间，即高维球体。

此外，Micheli 等人也对 exTNFS 中的关系集合和重复关系进行了详细的分析，并就如何定义和删除 exTNFS 中出现的重复关系提供了新的见解。

4　RSA 密码分析

RSA 密码分析是一个历史较长的研究领域，由于 RSA 应用的广泛性，RSA 密码分析也具有重要的意义。其中针对小私钥d的攻击方面，已经有一些比较好的成果[10]，即$d \leqslant N^{0.284}$和$d \leqslant N^{0.292}$；对基于最高有效位（MSB）的部分密钥泄露攻击，Takayasu 和 Kunihiro[11]的工作结果似乎已经是最优了，即$N^{0.292} \leqslant d \leqslant N$，因为他们的区间是从文献[10]中的上界 $N^{0.292}$ 光滑扩展到d的全尺寸 N。同时，Takayasu 和 Kunihiro 的工作也留下一个公开问题：是否存在一种基于最低有效位（LSB）的部分密钥泄露攻击？除了一般 RSA，还有一种基于中国剩余定理的 RSA 体制，即 CRT-RSA，其有两个私钥d_p和d_q。相对于一般的小私钥 RSA，小私钥的 CRT-RSA 更难分析。目前针对小私钥 CRT-RSA 的最优分析结果为$d_p, d_q \leqslant N^{0.122}$（称为 TLP 攻击）[12]，不过对此仍留有以下一些问题。

（1）TLP 攻击是否是最优的仍不清楚。

（2）小私钥 CRT-RSA 是否存在部分密钥泄露攻击仍不清楚。

（3）即使针对小私钥 CRT-RSA 的部分密钥泄露攻击存在，攻击是否能覆盖全尺寸$d_p, d_q \leqslant N^{0.5}$？

为此，May 等人给出了第一个全尺寸$N^{0.083} \leqslant d_p, d_q \leqslant N^{0.5}$的小私钥 CRT-RSA 部分密钥泄露攻击[13]。同样，由于该结果是从 TLP 攻击的上界 $N^{0.122}$ 光滑扩展到 $N^{0.5}$ 的，因此该结果被认为是最优的。具体来说，May 等人用牛顿多面体对 TLP 攻击进行了几何解释，从而构造了部分密钥泄露攻击。除此之外，May 等人还给出了一种在区间$N^{0.083} \leqslant d_p, d_q \leqslant N^{0.5}$的基于 LSB 的部分密钥泄露。值得注意的是，May 等人给出的 LSB 部分密钥泄露攻击覆盖了全尺寸，但是已知的针对 RSA 的 LSB 部分密钥泄露攻击并没有覆盖全尺寸。May 等人认为在实际的侧信道攻击中，d_p,d_q的信息比d的信息更容易获得，因为出于效率方面的考虑，实际的 RSA 应用中一般都使用 CRT-RSA。

5　超定可解线性方程组中的随机不均匀性

元生日悖论指的是在$[0,\ p-1]$（其中p是素数）中选取$O(\sqrt{p})$个随机元素，其中大概率会存在两个元素相等。Wagner 提出了一般化的生日悖论问题，即给定l个列表，从每个列表中选择一个元素x_i，使得$x_0 + \cdots + x_{l+1} = 0 (\mathrm{mod}\ m)$。在 Wagner 理论的应用中，最重要的是

超定可解线性方程组中的随机不均匀性（ROS）问题的次指数解算法。ROS 问题的定义如下：

给定一个素数p和一个在$\mathbb{Z}_p$范围内的随机谕言机H_{ros}，对于$\text{aux}_i \in \{0,1\}^*$找出$(l+1)$个函数$\rho_i$（$i \in [0,1]$）和向量$\boldsymbol{c} = (c_0, \cdots, c_{l-1})$，满足对于所有$i \in [0,1]$，有$H_{\text{ros}}(\rho_i, \text{aux}_i) = \rho_i(\boldsymbol{c})$。Wagner 的方案可以在次指数的时间内求解该问题，这个问题最早是由 Schnorr 在盲签名方案中提出的[15]。通过解决 ROS 问题，Wanger 证明当发出超过$O(\log_2 p)$个签名时，Schnorr 和 Okam oto-Schnorr 盲签名方案的不可伪造性可以在次指数时间内被攻击。

Fabrice 等人[14]重新回顾了 ROS 问题及其应用。Fabrice 等人首先给出了第一个维度为$l > \log_2 p$的 ROS 问题在多项式时间内的破解方案。并且，Fabrice 等人还将该方案同 Wagner 的方案相结合，提出了一个在维度$l < \log_2 p$时的优化版的次指数时间内的破解方案。最后，Fabrice 等人给出了一些他们的攻击可以应用的方案，包括盲签名、门限签名、多签名、部分盲签名、条件盲签名，以及在$l > \log_2 p$次并行设置下的匿名凭证。注意，尽管 Fabrice 等人的攻击并没有影响以上方案的理论安全性证明，但是 Fabrice 等人的攻击证明了上述方案在一些现实应用中并不实用。

6 Window τNAF 方法

Koblitz 曲线族$E_a: y^2 + zy = x^3 + ax^2 + 1$是椭圆曲线密码中具有重要理论和实践意义的一类，其中 4 宽度的 Koblitz 曲线直到 2020 年还在 NIST 的特殊发布中的密钥管理章节被提及。在 Koblitz 曲线上，可以使用 Frobenius 映射$\tau(x,y) = (x^2, y^2)$代替点倍增计算，提高标量乘法，即nP的速度。Koblitz 在 1991 年提出了一种通过将n表示成离散求和形式来快速计算nP的方法，其中P属于一个 Koblitz 曲线的主子群[16]。Solinas 在 2000 年将其进一步发展成了 Window τNAF 方法[17]，此后还有一些跟进工作对 Window τNAF 进行了进一步优化。在 Window τNAF 的程序中，预计算步骤对于计算标量乘法的效率有着重要影响。Yu 和 Xu 针对 Window τNAF 中标量乘法的预计算步骤做了进一步优化[18]。Yu 和 Xu 首先提出了新$\mu\bar{\tau}P$的操作，其中$\mu = (-1)^{1-a}$，$\bar{\tau}$是 τ 的复共轭，P是 Koblitz 曲线上的有理点，操作被用于代替 Window τ NAF 标量乘法预计算中的点加法运算或混合加法，其与目前的全加法和混合加法相比，减少了大量域运算。为了进一步发挥这一操作的优势，Yu 和 Xu 还提出了一种用于生成R_i（出现在 Window τNAF 程序的第二步）的平面搜索法，从而进一步减少了更多的域运算。在保持高效的同时，Yu 和 Xu 将 Window τNAF 的最优宽度从 6 扩展到 7，宽度的扩展意味着更稀疏 τ 扩展，从而有利于标量乘法的灵活性，即计算不固定的点。Yu 和 Xu 的研究结果使基于 Koblitz 曲线的标量乘法带入一个崭新的阶段。

7　单向函数的存在性

在密码学中，单向函数是否存在一直是最重要的公开问题。尽管目前已有许多单向函数的候选构造，但是大多数都是基于因数分解、离散对数或格上的困难问题假设的，能否构造基于更“标准”的假设的单向函数，如 NP≠P 或 NP≠BPP，仍然是一个公开的问题。到目前为止，对于这个问题的大多数结果都是否定的，但是这些结果是在有限制的情况下得到的，即要么讨论有限制的单向函数，要么用到了有限制的黑盒归约。

Liu 和 Pass[19]突破了这些限制，进一步讨论是否存在基于非常弱的假设 EXP≠BPP 的单向函数。他们注意到决定性 Levin-Kolmogorov 复杂度（MKtP）问题的困难性与是否 EXP≠BPP 问题相关，从而基于 Liu 和 Pass 的研究结果[20]，将单向函数的存在性与 MKtP 问题联系起来。具体来说：

（1）MKtP 是无限频繁双边误差中等平均困难的当且仅当无限频繁的单向函数存在；

（2）MKtP 是无限频繁无误差中等平均困难的当且仅当 EXP≠BPP。

因此，如果能解决 MKtP 的双边误差和无误差平均困难之间的技术空白，那么就可以得到基于 EXP≠BPP 的（无限频繁）的单向函数。此外，任何从 MKtP 无误差平均困难到 MKtP 双边误差平均困难的归约都可以推出 NP≠P。Liu 和 Pass 还考虑了有界空间 Kolmogorov 复杂度（MKSP）和有界空间条件 Kolmogorov 复杂度（McKSP）问题与单向函数的关系。具体来说：

（1）对数空间中无限频繁单向函数存在当且仅当$\mathrm{MKSP}[O(\log_2 n)]$是无限频繁双边误差中等平均困难的；

（2）当 McKSP 是几乎处处平均困难的，无限频繁单向函数存在当且仅当存在某个多项式时间图灵机 F，使得$\mathrm{McKSP}[F, O(\log_2 n)]$是无限频繁双边误差中等平均困难的。

尽管 Liu 和 Pass 的工作对于基于 EXP≠BPP 的单向函数是否存在这个问题并没有给出一个完整肯定的答案，但是 Liu 和 Pass 将这个问题归结于 Levin-Kolmogorov complexity 问题的各种平均情况之间的技术问题上，为解决基于 EXP≠BPP 的单向函数是否存在问题指明了一个方向。

8　叛徒追踪

叛徒追踪通过在用户各自的解密密钥中嵌入身份识别信息来阻止盗版情况的发生。它要求从一个已泄露的密钥中追踪提取出叛徒的身份信息，从而进一步对其进行罚款、控诉或撤销权限等。因为一般来讲分析程序的代码是困难的，所以大多数叛徒追踪算法通常采用黑盒

解码的形式，即只向解码器问询密文，然后观察返回的结果。

黑盒追踪存在一些限制。具体来说，首先相较于秘密追踪，一般更倾向于公开追踪，这是因为：

（1）公开追踪允许任何人进行追踪；

（2）秘密追踪没有一个比较自然的机制用于呈递证据；

（3）秘密追踪花费的时间更长，并且无法保证追踪者一定能发现解码器。

对于黑盒公开追踪，存在两个问题：隐私问题和一致性问题。隐私问题是指远程攻击者可以通过在互联网上运行叛徒追踪算法得到其他用户的隐私信息，因为算法是公开的，且其黑盒性，攻击者只需要问询密文。一致性是指不同用户解密密钥的功能应该是不同的，即对于同一个密文，不同用户用各自解密密钥得到的结果应该是不同的。这一点是固有的，因为如果功能是相同的，那么叛徒可以对其使用不可区分混淆器，从而不存在高效的叛徒追踪算法；反之，如果存在高效的叛徒追踪算法，那么会得出不可区分混淆器不存在的结论，但目前这似乎是遥不可及的。对于黑盒公开追踪，找到上述的密文一定是件容易的事情，因为追踪者只能用黑盒访问解码器，所以为了能有效识别叛徒，他必须找到这样的密文。将这种一致性应用到 MPC 中会破坏“广播信道”这个假设，因为“广播信道”要求接收方收到的信息是一致的，但敌手如果利用追踪算法中的一致性问题，就会导致接收方最终得到的信息不一致。以上两种问题可以通过白盒追踪解决。首先，针对隐私问题，使用白盒追踪会使得远程攻击者无法接触到解码器，从而无法进行攻击；其次，针对一致性问题，由于白盒追踪不限制仅用黑盒问询解码器，因此攻击者有可能找不到上述的密文。

Zhandry[21]给出了一个新的针对隐私问题和/或一致性问题的白盒公开追踪算法，也是第一个考虑在一般解码器情况下的白盒追踪算法。具体来讲，首先，Zhandry 给出了叛徒追踪算法中隐私问题和一致性问题的定义，并证明了对于黑盒公开追踪算法，不可能解决隐私问题和一致性问题。其次，Zhandry 基于通用公钥加密和非交互零知识证明或不可区分混淆器构造了只解决隐私问题的白盒公开追踪系统，该系统给出了不同合谋数和参数大小之间的权衡结果。最后，Zhandry 基于全同态加密、计算-比较混淆器和非交互零知识证明（基于 LWE 或者不可区分混淆器）构造了常数级合谋的白盒公开追踪算法，该算法同时解决了隐私问题和一致性问题，但是 Zhandry 并没有完全解决一致性问题。Zhandry 还总结了关于白盒叛徒追踪未来几个可以研究的方向：首先，他们的方案只能对恒定数量的合谋构建一致性的叛徒追踪，是否可以在极强的假设下实现对任意数量合谋构建一致性追踪？其次，由于他们的结构利用了多重机制，在多个层面上使用了非黑盒技术，导致其构造效率是较低的，未来是否能实现真正高效的白盒叛徒追踪？叛徒追踪可以看作是软件水印问题的一个特例，在水印环境中的大多数工作是使用黑盒标记检测和提取的，隐私和一致性问题自然会转化为水印，那么白盒技术可以用于克服水印中的类似问题吗？白盒标记检测和提取可以用于弥补这种差距吗？最

后，用于追踪叛徒的选择密文攻击以前也被考虑过，本文强调了 CCA 攻击的进一步后果，是否还有其他可能的后果？

9　Fiat-Shamir 转换

Fiat-Shamir 转换是一种通过哈希函数将公共投掷（Public-Coin）交互协议转换为具有相同功能的非交互协议的一般方法。由于 Fiat-Shamir 转换的一般性以及实际高效性，在其提出之后的 30 多年来，一直在理论和应用密码学中扮演重要的角色。到目前为止，Fiat-Shamir 转换的大部分应用仅仅是启发式健全的。证明其健全性的方法主要有随机谕言模型和相关难处理性两种。第一种方法试图高度概括地证明 Fiat-Shamir 转换的可用性，第二种方法对于特定的协议和哈希函数给出了 Fiat-Shamir 转换的完整的证明。这两种方法都要求所用哈希函数具有某种形式的密码困难性，具体来说，第一种方法要求哈希函数与一个真正的随机函数不可区分，第二种方法中的相关难处理哈希函数族的安全性是基于标准密码学假设的。此外，基于之前相关工作中 Fiat-Shamir 转换所用的哈希函数都比较复杂，Chen 等人[22]提出并探讨：是否可以用一个非密码的哈希函数实例化启发式安全的 Fiat-Shamir 转换，以及是否可以用一个简单的哈希函数实例化 Fiat-Shamir 转换。Chen 等人关注如下两个常规且重要的 Fiat-Shamir 转换的用例：

（1）压缩轮数 3 消息识别方案；

（2）用压缩轮数 3 消息诚实验证者论述系统（Argument System）获得对 NP 的非交互零知识证明论述（NIZK Arguments）。

对于这两个用例，Chen 等人给出了几个可以应用 Fiat-Shamir 转换的常规消息协议：

（1）Schnorr 的识别方案；

（2）对于 Diffie-Hellman 语言的 Chaum-Pedersen 交互式证明系统；

（3）Lyubashevsky 的基于格的识别方案；

（4）σ-协议。

对于以上协议，Chen 等人首先给出了可以用非密码哈希函数实例化健全的 Fiat-Shamir 转换的正面结果。

定理 1　考虑在n维$\mathbb{Z}_q$空间上的 Lyubashevsky 验证方案。定义哈希函数$h_{ae}\mathbb{Z}_q^n \to \mathbb{Z}_2^{n\log q}$作为比特分解函数$h(v) = G^{-1}(v)$。那么，在短整数解（SIS）假设下，应使用$h$的 Fiat-Shamir 的 Lyubashevsky 方案，在随机实例中是健全的。

定理 2（非正式）　使用比特分解 Fiat-Shamir 哈希函数的基于格的 Lyubashevsky 签名等价于基于格的 Hash-and-Sign 签名。

定理 3（使用$\mathbb{Z}_q$-线性哈希函数的 Schnorr 签名）　考虑定义在素数阶p的群G上的 Schnorr

签名方案，其明文空间 $\mathcal{M}$ 是$\mathbb{Z}_p$的稀疏子集，即$\mathcal{M} \subset \mathbb{Z}_p$且$|\mathcal{M}|/\mathbb{Z}_p \leqslant negl(\lambda)$。令$\ell$表示任意群元素的最大比特表示长度，使得对于任意的$g \in G$可以被看作$g \in \{0,1\}^{\ell} = [2^{\ell}]$。定义哈希族$h_k(g,m) \coloneqq g + m + k(\mathrm{mod}\, p)$，其中，等式右侧的$g$是用二进制表示的整数，即$g \in \{0,1\}^{\ell}$。

定理 4 令 Π^{CP} 表示定义在阶为p的群 G 上的改进的 Chaum-Pedersen 协议。令ℓ表示任意群元素的最大比特表示长度，使得对于任意的 $g \in G$可以被看作$g \in \{0,1\}^{\ell} = [2^{\ell}]$。定义哈希函数$h(g_1, g_2, g_3, g_4) = g_1 + g_2 + g_3 + g_4(\mathrm{mod}\, p)fi$，其中，等式右侧的$g_i$是用二进制表示的整数，即$g_i \in \{0,1\}^{\ell}$。

此外，Chen 等人还给出了一些负面结果，即对于一些协议，Fiat-Shamir 转换确实需要密码学哈希函数。

定理 5（非正式） 对于在通用群模型（GGM）中实例化的$\Pi = \Pi^{\mathrm{bit-sch}}$，如果$\mathcal{H}$是一个不调用群谕言机的哈希族，那么$\Pi^{t}_{\mathrm{FS},\mathcal{H}}$在 GGM 中是不健全的。对于在随机谕言模型下 Blum's Hamiltonicity 协议的任意实例化，如果$\mathcal{H}$是一个不依赖于谕言机$\mathcal{O}$的哈希族，那么$\Pi_{\mathrm{FS},\mathcal{H}}$是不健全的。

相对于谕言机$\mathcal{O}$构造的任意$\Pi \in \mathcal{C}$，如果$\mathcal{H}$不依赖于$\mathcal{O}$，那么$\Pi_{\mathrm{FS},\mathcal{H}}$是不健全的。

定理 6（非正式） 如果$\mathcal{H}$不是抗 mix−and−match 的，那么对于任意的$\Pi \in \mathcal{C}$，存在对$\Pi_{\mathrm{FS},\mathcal{H}}$健全性的多项式时间攻击。

此外，Chen 等人留下了一个公开的问题：哪些交互协议允许“简单”的 Fiat-Shamir 转换编译器？

总的来说，Chen 等人认为他们的工作为更进一步理解“什么时候 Fiat-Shamir 转换的哈希函数需要是（密码）困难的”提供了一个出发点。

10 有损陷门函数

有损陷门函数是一种密码学的基础工具，其在密码学中有很多应用，包括证明 CCA 安全、带有许多核心比特的陷门函数、抗碰撞哈希函数、不经意传输等。有损陷门函数包括一族由公共函数密钥索引的函数，其中公共函数密钥有两种生成模式：单射模式和有损模式。单射模式下生成单射函数族，并带有一个陷门使得可以求出原像；有损模式下生成原像损失的函数族，两种模式生成的函数族在计算意义下不可区分。目前的有损陷门函数的构造都是基于密码性质的假设，如 DDH、LWE、QR 等。

为了基于更弱的假设，扩大应用的广泛性，Quach 等人[23]基于单射伪随机生成器或放宽单射条件的单向函数，提出了相对于有损陷门函数有所放宽的目标有损函数（TLF）。TLF 同样分为单射模式和有损模式，其在单射模式下不需要存在可逆陷门，在有损模式下只需指定损失特定目标x^*的l位信息。TLF 函数族$F_{\mathrm{fk}}(x) = h(G(x))$由压缩

一位的通用哈希函数h_a和单射的伪随机生成器$G(x)$复合而成。在单射模式下，随机选取h_a作为函数密钥fk，从而获得了极大概率的单射性。在信息有损模式下，函数初始化指定丢失的特定目标位，并通过调节哈希的参数a以获得伪随机数生成的碰撞，而这样的调节和随机取值不可区分，从而保证了两个模式下的函数不可区分。通过实现单射模式和有损模式，从而成功构造了一位有损陷门函数，通过将 1-TLF 并行重复，得到l-TLF。该文还提出了带分支的 TLF:$F_{\mathrm{fk,tag}}(\cdot)$，根据目标分支模式的不同，分为 T-AIBO 和 T-ALBO 两类。其中 T-AIBO 只有一个目标分支tag为有损模式，其余分支为单射模式，可以通过修改无分支 TLF 的fk得到；T-ALBO 只有一个目标分支tag为单射模式，其余分支为有损模式。TLF 单射的要求也可被进一步放宽，在单射模式下，不需要$F_{\mathrm{fk}}(x)$唯一确定x，只需确定x的某些属性；在有损模式下，只需损失x^*属性的l位信息。这些目标有损函数上的研究新进展也带动了若干应用，包括伪熵函数、抗泄露密码、高熵源均匀随机数提取、对称加密方案中的选择打开安全、单射陷门函数的 CCA 加密。

本节作者：陈洁（华东师范大学），陈荣茂、王毅（国防科技大学）

参考文献

[1] MAURER U, PORTMANN C, RITO G. Giving an Adver- sary Guarantees (Or: How to Model Designated Verifier Signatures in a Composable Framework)[C]. ASIACRYPT 2021(3): 189-219.

[2] DAMGÅRD I, HAAGH H, MERCER R, et al. Stronger Security and Constructions of Multi-designated Verifier Signatures[C]. TCC 2020(2): 229-260.

[3] ABDALLA M, BARBOSA M, KATZ J, et al. Alge- braic Adversaries in the Universal Composability Framework[C]. ASIACRYPT 2021(3): 311-341.

[4] MICCIANCIO D, WALTER M. On the Bit Security of Crypto-graphic Primitives[C]. EUROCRYPT 2018(1): 3-28.

[5] WATANABE S, YASUNAGA K. Bit Security as Computational Cost for Winning Games with High Probability[C]. ASIACRYPT 2021(3): 161-188.

[6] GARG S, HAJIABADI M, MALAVOLTA G, et al. How to Build a Trapdoor Function from an Encryption Scheme[C]. ASIACRYPT 2021(3): 220-249.

[7] KIM T, BARBULESCU R. Extended Tower Number Field Sieve: A New Complexity for the Medium Prime Case[C]. CRYPTO 2016(1): 543-571.

[8] KIM T, JEONG J. Extended Tower Number Field Sieve with Application to Finite Fields of Arbitrary Composite Extension Degree[C]. PKC 2017(1): 388-408.

[9] DE MICHELI G, GAUDRY P, PIERROT C. Lattice Enumer- ation for Tower NFS: A 521-Bit Discrete Logarithm Computation[C]. ASIACRYPT 2021(1): 67-96.

[10] BONEH D, DURFEE G. Cryptanalysis of RSA with private key d less than N0.292[C]. EUROCRYPT 1999: 1-11.

[11] TAKAYASU A, KUNIHIRO N. Partial Key Exposure Attacks on RSA: Achieving the Boneh-Durfee Bound[C]. SAC 2014: 345-362.

[12] TAKAYASU A, LU Y, PENG L. Small CRT-Exponent RSA Revisited[C]. EUROCRYPT 2017(2): 130-159.

[13] MAY A, NOWAKOWSKI J, SARKAR S. Partial Key Expo- sure Attack on Short Secret Exponent CRT-RSA[C]. ASIACRYPT 2021(1): 99-129.

[14] BENHAMOUDA F, LEPOINT T, LOSS J, et al. On the (in)security of ROS[C]. EUROCRYPT 2021(1): 33-53.

[15] SCHNORR C-P. Security of Blind Discrete Log Signatures againstInteractive Attacks[C]. ICICS 2001: 1-12.

[16] KOBLITZ N. CM-Curves with Good Cryptographic Properties[C]. CRYPTO 1991: 279-287.

[17] SOLINAS J A. Efficient Arithmetic on Koblitz Curves[J].DCC 2000,19(2/3): 195-249.

[18] YU W, XU G. Pre-computation Scheme of Window τNAF for Koblitz Curves Revisited[C]. EUROCRYPT 2021 (2): 187-218.

[19] LIU Y, PASS R. On the Possibility of Basing Cryptography on EXP≠ BPP[C]. CRYPTO 2021 (1): 11-40.

[20] LIU Y, PASS R. On One-way Functions and Kolmogorov Complexity[C]. FOCS 2020: 1243-1254.

[21] ZHANDRY M. White Box Traitor Tracing[C]. CRYPTO 2021 (4): 303-333.

[22] CHEN Y, LOMBARDI A, MA F, et al. Does Fiat-Shamir Require a Cryptographic Hash Function?[C]. CRYPTO 2021(4): 334-363.

[23] QUACH W, WATERS B, WICHS D. Targeted Lossy Functions and Applications [C]. CRYPTO 2021(4): 424-453.

[24] HERZBERG A, LEIBOWITZ H, SYTA E, et al. MoSS: Modular Security Specifica-tions Framework. CRYPTO 2021(3): 33-63.

公钥密码

1　公钥加密

CCA 安全性一直以来被认为是公钥加密（Public-Key Encryption，PKE）机制的标准安全性。随着密码应用场景的不断丰富，研究人员发现需要重新审视 CCA 安全性的定义并根据实际应用场景做出调整和改进，以满足现实世界对密码系统的需求。

可重放的 CCA（Replayable CCA，RCCA）安全性非常接近于公钥加密方案的标准安全性——CCA 安全，且支持密文的可重随机化（该性质与 CCA 安全性互斥）。而具有可重随机性的 RCCA 安全公钥加密方案可用于构造密码逆向防火墙、混淆网络、具有可控延展性的 NIZK 协议。Canetti 等人在 2003 年美密会上首次提出 RCCA 安全性。随后，Prabhakaran 和 Rosulek 在 2007 年美密会上给出了第一个标准模型下的可重随机 RCCA 安全方案（简称 PR 方案）。然而，可重随机 RCCA 安全性下如何实现匿名性一直没有得到解决。匿名性确保密文隐藏了接收者信息，对混淆网络等隐私保护应用而言不可或缺，也被认为是充分释放 RCCA 安全性价值的关键。因此，Prabhakaran 和 Rosulek 将如何构造具有匿名性的可重随机 RCCA 安全公钥加密方案列为公开问题在 2007 年美密会上提出。

Yi Wang 等人[1]首次给出了匿名的可重随机 RCCA 安全公钥加密方案的构造框架。该框架沿用了 PR 方案的双股架构，确保了密文的普适可重随机性。为了实现匿名性，该框架对密文的重随机方式进行进一步的限制，使得敌手无法使用猜测的公钥对密文进行重随机，并利用解密谕言机返回的结果对猜测进行验证。该框架的核心组件是一种新的密码原语——可重随机的平滑投影哈希函数。该原语可以被视为 Cramer 和 Shoup 在 2002 年欧密会上提出的平滑投影哈希函数的一种变体。利用该哈希函数的平滑性可以大大简化 PR 方案原有的复杂证明，并为 RCCA 安全性和匿名性提供一种简洁的模块化分析。另外，该工作使用 k-Lin 困难性假设对该框架进行实例化，得到了匿名的可重随机 RCCA 安全加密的具体方案。与 PR 方案相比，该方案具有更为一般的困难性假设。

随后，Yi Wang 等人[2]将上述工作拓展到标识加密的背景下，首次定义了匿名的基于标识的 RCCA（Anonymous Identity-Based RCCA，ANON-ID-RCCA）安全性，给出了可重随机的 ANON-ID-RCCA 安全标识加密方案的具体构造，并将其用于构造基于标识的通用混淆网络（Identity-Based Universal Mixnet）。ID-RCCA 安全性可以被视为标识加密方案的标准安全性，

即 ID-CCA 安全性的宽松版本，因而允许对密文进行重随机操作。具体标识加密方案构造的核心思路则是将双股架构应用于著名的 Gentry 标识加密方案，并对密文的重随机操作进行限制，使得密文仅有的重随机方式是自身重随机。基于标识的通用混淆网络则实现了可追踪的匿名通信应用。一方面，标识加密使得发送者不需要请求接收者的公钥信息，并且密文的匿名性使得敌手无法从消息获得接收者信息。另一方面，密钥生成中心由于拥有主私钥，能够对滥用匿名通信服务的恶意用户进行追踪，从而实现了公平的匿名性（Fair Anonymity）。

除对 CCA 安全性进行合理弱化以适应特定应用的需求外，有些工作也尝试针对现实密码系统中公钥加密方案的 CCA 安全性进行了分析。公钥加密方案通常用于多用户的开放性系统。这导致某些参与方遭受敌手的选择性开放攻击（Selective Opening Attacks，SOA），即敌手自适应地攻陷部分参与方来尝试学习一些未开放密文的消息。然而，之前关于 SOA 安全的公钥加密方案仅提供发送方选择性开放安全（Sender Selective Opening，SSO，部分发送方被攻陷，消息以及加密的随机性被泄露）或接收方选择性开放安全（Receiver Selective Opening，RSO，部分接收方被攻陷，消息以及解密的密钥被泄露）。这在实践中很难预测。此外，以往关于 RSO 安全性的研究大多只关注“单个挑战”的情况，即对于每个公钥敌手只能获取一个对应的挑战密文。对于一个具有多个公钥且每个公钥将被多次使用的多用户系统，Junzuo Lai 等人[3]给出一种新的通用选择性开放安全，即敌手可以同时自适应地攻陷发送方和接收方的一部分，并获得明文消息以及用于加密的内部随机性和用于解密的密钥，同时维持未被攻陷的参与方的消息仍然受到保护。Lai 等人[3]首先用基于模拟（Simulation-based，SIM）的方式形式化了 Bi-SO（Bi-Selective Opening）安全性，并证明一些实用的公钥加密方案在随机谕言模型中达到了 SIM-Bi-SO-CCA 安全性。其次，Junzuo Lai 等人[3]提出了一个 Bi-SO 安全的弱模型，称其为 SIM-wBi-SOk-CCA 安全，其中敌手必须明确在收到公钥后并且接收挑战密文前是否攻陷发送方或接收方，同时只允许攻陷其中公钥最多 k 次被用于加密的接收方。弱模型仍然是有意义的，因为它同时提供了原始的 SIM-SSO-CCA 安全和 SIM-RSO-CCA 安全，并被证明其安全性更强。最后，基于新设计的哈希证明系统的变体（满足密钥等价性），Junzuo Lai 等人[3]提出了一个公钥加密方案的通用构造，在标准模型中实现 SIM-wBi-SOk-CCA 安全性，并从各种标准假设中对其进行了实例化。

与传统公钥加密机制不同，时间锁加密（Time-Lock Encryption，TLE）是一种特殊的加密机制。在 TLE 中，消息的密文需要在经过一定时间之后才能被解密。此前已有的 TLE 方案均很难被证明是通用可组合（Universal Composability，UC）安全的。其原因与 UC 框架本身有关。直观上，为了刻画语义安全性，在理想世界中，任何与明文相关的信息都不应被泄露。然而，所有密文最终都会打开，从而导致现实世界与理想世界存在细微的差别。Myrto Arapinis 等人[4]提出了首个 UC 安全的 TLE 方案 Astrolabous。其核心思想是采用 Nielsen 在 2002 年美密会上提出的技术对安全的 TLE 方案在随机谕言模型下进行扩展，使得模拟器能

够进行模拟。该扩展方式可以被应用于任何满足独立安全性的 TLE 方案，尤其是 Astrolabous 方案。具体来讲，Myrto Arapinis 等人[4]所提出方案通过采用 UC 模型对 TLE 的理想功能（Ideal Functionality）进行了抽象，定义了基于游戏的 TLE 的独立安全性定义，并设计了一个基于 TLE 的混合协议。如果底层的 TLE 满足独立安全性，那么该混合协议在随机谕言模型下 UC 实现了 TLE 功能。与 Myrto Arapinis 等人提出的方案从安全模型的角度关注 TLE 安全性不同，文献[5]针对时间锁密码学相关困难性假设开展了研究，具体研究了未知阶有限阿贝尔群以及群上的困难性假设，并且将代数群模型和强代数群模型从循环群推广到任意的未知阶有限阿贝尔群。此外，Myrto Arapinis 等人提出了归约到这种困难性假设的方式和这种假设在时间锁密码学中的应用。

混合公钥加密（HPKE）是一种新兴的公钥加密标准，已被用于组密钥建立协议 MLS（Messaging Layer Security）。HPKE 包括 4 种模式。作为 HPKE 中最具创新性的部分，认证模式 $HPKE_{Auth}$ 受到广泛关注。认证模式下的 HPKE 被称为认证公钥加密（APKE）。在 MLS 协议中，与使用基本模式 $HPKE_{Base}$ 相比，使用 $HPKE_{Auth}$ 具有更低的计算和通信开销，但它降低了协议的安全性，容易遭受密钥泄露伪装攻击（KCI），且在接收方的私钥泄露的情况下，无法对发送方进行认证。针对上述问题，Joël Alwen 等人[6]研究 HPKE 的认证模式，并对 APKE 方案和相关的认证密钥封装（AKEM）进行准确的安全定义，证明了对每个 AKEM 的安全定义，其在单用户场景下的安全可以直接保证多用户场景下的安全。此外，Joël Alwen 等人研究了 RFC 中通用的 $HPKE_{Auth}$ 方案 DH-AKEM 的安全性，提出了通用的 AKEM/DEM 组合理论，用于实现 $HPKE_{Auth}$ 安全。针对 HPKE 突出的实用性，Joël Alwen 等人对 $HPKE_{Auth}$ 的所有分析都是基于多用户场景的，且所有的证明结果中的安全损失系数都是各参数的具体函数，有助于在实际应用 HPKE 时设置可靠的安全参数。

本节作者：陈荣茂、王毅（国防科技大学）

参考文献

[1] WANG Y, CHEN R, YANG G, et al. Receiver-Anonymity in Rerandomizable RCC A-Secure Cryptosystems Resolved[C]. CRYPTO 2021(4): 270-300.

[2] WANG Y, CHEN R, HUANG X, et al. Identity-Based Encryption for Fair Anonymity Applications: Defining, Implementing, and Applying Rerandomizable RCCA-Secure IBE[C]. ASIACRYPT 2021 (2): 427-455.

[3] LAI J, YANG R, HUANG Z, et al. Simulation-Based Bi-Selective Opening Security for Public Key Encryption[C]. ASIACRYPT 2021 (2): 456-482.

[4] ARAPINIS M, LAMPROU N, ZACHARIAS T. Astrolabous: A Universally Compos

able Time-Lock Encryption Scheme[C]. ASIACRYPT 2021(2): 398-426.

[5] VAN BAARSEN A, STEVENS M. On Time-Lock Cryptographic Assumptions in Abelian Hidden-Order Groups[C]. ASIACRYPT 2021 (2): 367-397.

[6] ALWEN J, BLANCHET B, HAUCK E, et al. Analysing the HPKE Standard[C]. EUROCRYPT (1) 2021: 87-116.

2 数字签名

环签名允许用户临时组织若干用户形成一个用户环，并以该环的名义匿名地签署消息。环签名所能提供的匿名性保障程度与环的大小成正比。因此，设计环签名方案使其签名大小最小化到关于群成员数量的某个函数，成为密码学研究中的一个重要目标。文献[1]基于 LWE 问题提出了第一个紧凑环签名方案（签名大小与环大小的对数相关）。该方案具备标准模型下的安全性，不依赖于公共引用串模型或者启发式的随机谕言模型。与 Backes 等人在欧密会 2019 上的工作相比，该方案能够基于 LWE 困难问题被证明是后量子安全的。该方案的核心是对 NP∩coNP 问题中紧致且统计证据不可区分的 ZAP 论证的一个新构造，该构造可在 LWE 假设下被证明安全。在这项工作之前，（对于任意的 NP 语言的）统计 ZAPs 只有在 LWE 的亚指数难度假设下才存在。

文献[2]引入了层次集成签名和加密（HISE）的概念，其中单个公钥被用于签名和加密；签名密钥起着主密钥的作用，可以根据签名密钥获得解密密钥，但不能从解密密钥获得签名密钥，这种两级的密钥派生结构使得密钥隔离和密钥重用达到了很好的平衡。HISE 有密钥重用以及允许个人密钥托管的双重优势。文献[2]提出了 HISE 的两个通用结构。一种是（受限的）基于身份的加密；另一种则来自均匀单向函数、公钥加密和通用公共硬币零知识证明。为了进一步实现全局密钥托管，文献[2]回顾了全局托管 PKE，对全局托管 PKE 的算法和安全模型进行了形式化定义，并提出了两种通用的构造。第一种，通过将任意 PKE 转化成具有全局托管属性 PKE。第二种，通过三方非交互式密钥交换和全局托管 PKE 之间建立连接。结合以上研究结果，文献[2]得到了一个既支持个人密钥托管又支持全局密钥托管的 HISE 方案。进一步地通过实例化（全局托管）HISE 的通用构造获得了 128 比特安全性的 HISE 方案。该方案具有与最佳笛卡尔乘积组合公钥方案相当的性能，并且在功能以及公钥重用方面都具有优势。

数字签名的标准安全概念是单挑战（Single-Challenge）EUF-CMA 安全，其中敌手输出单个消息及签名对；若该签名对是伪造的，则敌手赢得游戏。Auerbach 等人在美密会 2017 上介绍了内存紧致（Memory-Tightness）归约，并认为在这种情况下，正确的安全目标实际上是一个更强的多挑战（Multi-Challenge）定义，其中敌手可输出多个消息签名对；如果至少有一

个签名对是伪造的，那么敌手赢得游戏。到目前为止，还不存在简单标准假设下的构造方案可以同时实现时间、成功概率和完全的内存紧致性。Auerbach 等人在美密会 2017 及 Wang 等人在欧密会 2018 上发表的研究表明内存紧致签名无法通过一般的归约来实现。这些不可能性结果可能会给人一种内存紧致签名实现起来很困难或不可能实现的印象。文献[3]通过给出第一个多挑战环境下所有方面均具有完全紧致的签名方案构造，证明了这种印象是错误的。为了规避已知的不可能性结果，文献[3]首先在单挑战环境下引入了正则归约（Canonical Reductions）的概念，同时证明了一个一般性定理：如果签名是强不可伪造的，敌手对于每个消息只获得一个签名，并且紧致安全的伪随机函数存在，那么每个具有正则归约的签名方案在多挑战环境下都是内存紧致安全的。其随后通过一个简单的通用转换，在多挑战环境下实现了内存紧致的每消息多签名（Many-Signatures-per-Message）安全性。这是多挑战环境下第一个满足内存紧致且具有强存在性不可伪造安全性的签名方案。文献[3]还说明了标准的安全性证明通常已经可以被视为正则归约，并利用 Abdalla 等人在欧密会 2012 上提出的有损认证（Lossy Identification）方案、Bellare 和 Rogaway 在欧密会 1996 上提出的 RSA 全域哈希的两种变体以及 Boneh 等人在亚密会 2001 上提出的 BLS 签名的两种变体对签名方案进行了实例化。

由于 Schnorr 签名的线性的验证等式容易应用于门限的场景。因为实际部署中的随机源通常不可靠，或者因为软件错误、恶意操作者或停电等使得状态的连续性（State Continuity）难以实现，Schnorr 签名中的随机数往往成为现实中的攻击的对象。尽管作为 Schnorr 签名变体的 EdDSA 方案利用消息和密钥的函数来确定地产生随机数，并不存在需要状态的连续性这一问题，但是当 Schnorr 变体方案应用到门限环境下时，这一优点便不再存在。构造一个既不要求各方使用新的随机数，也不需要更新长期秘密值的门限 Schnorr 方案成为新的挑战。文献[4]构造了一个不诚实者大多数的门限 Schnorr 协议，该协议使用标准化的分组密码实现了产生无状态确定性的随机数。其核心思想是利用混淆电路（ZKGC）范例中实现零知识的一些新技术，以辅助验证所有正确推导出的随机数，包括基于 UC 承诺的新技术；允许证明者对证据只承诺一次，并仅使用简单对称密钥操作就可以证明无限数量的在线描述，以及可将中间混淆电路的导线标签转换为算术编码的小工具。

文献[5]提出了由多个标准的识别方案构造的一个通用的环签名 DualRing 方案。传统的环签名方案的构造基于累加器或基于签名者索引的零知识证明。DualRing 方案的构造与这些主流的环签名方案不同，它是一个由承诺环和挑战环构成的双环结构。由于 DualRing 方案由 n 个挑战和一个回应组成，因此基于 DL 和基于格的 DualRing 方案可以获得更短的签名。具体地说，利用 Schnorr 识别协议基于 DL 的 DualRing 方案通过一个知识系统的证明（如 Bulletproofs）使得签名的长度可以被压缩成与环中用户的数量对数复杂性相关，并进一步地通过求和论证（Sum Argument），在保持相同证明大小的基础上能减少几乎一半的计算代价，

因此提高了 Bulletproofs 的效率。采用具有求和论证的 Schnorr 识别方案构造基于 DL 的 DualRing-EC，是目前没有使用可信设置阶段的最短环签名。基于格的 DualRing 方案通过实例化基于 M-LWE 和 M-SIS 假设的正则识别方案，构造出当环的大小在 4~2000 范围内时基于格的最短环签名 DualRing-LB，并且表明在签名和验证方面比最先进最快的构造（在签名和验证的运行时间方面）至少快 5 倍。

盲签名允许签名者在不获得消息的前提下对消息进行签名。到目前为止，所有轮优化的盲签名均需要可信的设置阶段，或者交互式的假定，或者复杂的 Leveraging。文献[6]构造了第一个轮数最优的、在 Plain 模型下不需要可信设置阶段基于标准假设的盲签名。该方案的构造利用了各种标准的密码原语，以及通过实例化经典的和后量子标准多项式时间假设得到的新原语。其主要构建块是一个符合盲签名的零知识论证系统，其显著的特点是用量子多项式时间模拟器来对抗非均匀、传统的多项式时间敌手。

文献[7]首次提出了在随机谕言模型下分别基于 RSA、因式分解和离散对数假设的盲签名方案。该核心思想是对 Pointcheval 转换（欧密会 1998）进行扩展和推广，将特定对（并发）发布对数级多签名安全的盲签名方案，转换为对（并发）发布多项式级多签名安全的盲签名方案，从而增强了特定盲签名方案的安全性。与 Pointcheval 转换相比，采用文献[7]的转换有下列优势：首先，转换后的方案（CCBS）无须要求签名者在检测到作弊时停止进行签名；其次，如果原始签名方案（BS）是并发安全的，那么 CCBS 也是并发安全的；最后，该变换可以应用于任何以线性函数族的某种方式构造的盲签名方案 BS，如 Fiat-Shamir、Okamoto-Guillou-Quisquater 和 Okamoto-Schnorr 盲签名方案，并且可以在随机谕言模型的标准假设下证明所有这些方案对于对数级执行都是安全的。

本节作者：黄琼（华南农业大学）、马莎（华南农业大学）、黄建业（澳大利亚伍伦贡大学）

参考文献

[1] CHATTERJEE R, GARG S, HAJIABADI M, et al. Compact Ring Signatures from Learning with Errors[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology-CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

[2] YU C, TANG Q, WANG Y. Hierarchical Integrated Signature and Encryption[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13091. Springer, Cham.

[3] DENIS D, GELLERT K, JAGER T, et al. Digital Signatures with Memory-Tight Security in the Multi-challenge Setting[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology -

ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093. Springer, Cham.

[4] GARILLOT F, KONDI Y, MOHASSEL P, et al. Threshold Schnorr with Stateless Deterministic Signing from Standard Assumptions[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

[5] YUEN T, ESGIN M, LIU J, et al. DualRing: Generic Construction of Ring Signatures with Efficient Instantiations[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

[6] KATSUMATA S, NISHIMAKI R, YAMADA S, et al. Round-Optimal Blind Signatures in the Plain Model from Classical and Quantum Standard Assumptions[C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology - EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12696. Springer, Cham.

[7] KATZ J, LOSS J, ROSENBERY M. Boosting the Security of Blind Signature Schemes[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093. Springer, Cham.

3　认证/密钥协商/密钥封装/密钥交换/密钥生成

在安全归约中，我们总是将敌手攻破密码方案的能力转化为解决相关的底层困难问题的能力。在这个归约过程中，通常都会存在安全损失，即敌手攻破方案的能力不能完全转化为解决相关的底层困难问题的能力。认证密钥交换（AKE）协议的安全证明更是如此。目前，各种 AKE 协议的安全证明几乎都存在“承诺问题”（Commitment Problem）。该问题使得安全损失总是与用户数量和会话数量呈线性关系或平方关系，从而无法实现“紧致安全”（Tight Security）。

在 2020 年的亚密会上，Liu 等人提出了一种具有紧致安全和前向安全（Forward Security）的 AKE 协议。该工作考虑不标准的多比特猜测（Multi-Bit-Guess，MBG）AKE 安全模型，且其攻击模型中没有考虑状态泄露攻击。文献[1]提出具有紧致安全的 AKE 协议，该协议可实现前向安全。它基于标准的单比特猜测（Single-Bit-Guess，SBG）AKE 安全模型实现随机谕言模型（ROM）下的紧致安全。文献[1]的核心贡献是提出了一种通用的方法来构造紧致安全的 AKE 协议，同时避免了“承诺问题”。首先，文献[1]基于哈希证明系统构造在 ROM 下的非承诺密钥封装（NCKE）方案，该方案具有紧致安全。然后，文献[1]基于 NCKE 方案构造具有紧致安全的 AKE 协议，该协议满足弱前向安全（wPFS）。在此基础上，通过添加一个紧致安全的签名方案，文献[1]构造出具有紧致安全和前向安全的 AKE 协议。

和文献[1]采用随机谕言模型不同，文献[2]首次在标准模型下提出了 SBG 安全概念下紧

致安全的 AKE 协议。文献[2]的构造分两步进行，首先构造一个安全的签名方案。签名方案经常被用于构造 AKE 协议。目前所有紧致安全的 AKE 协议都基于选择消息攻击下存在性不可伪造（EUF-CMA）的签名方案进行构造。但在多用户的场景中，这类签名方案并不是紧致安全的，存在一个与用户数量有关的线性安全损失。文献[2]利用紧致安全的分层身份加密（HIBE）方案来构造新的消息认证码（MAC）协议。该 MAC 协议在适应性破坏下仍然是紧致安全的。文献[2]基于该 MAC 协议设计了第一个多用户场景下具备 EUF-CMA 安全（MU-EUF-CMA）的紧致签名方案。基于该签名方案，文献[2]共提出了 3 种 AKE 协议。它们的安全性依次递增。第一种 AKE 协议设计遵循经典的“KEM+2×SIG”设计范式。为实现抗重放攻击，在第一种协议的基础上，文献[2]增加了一个随机数消息。此外，为支持会话状态泄露询问，文献[2]增加了一个对称加密算法用于加密临时私钥，与文献[1]的做法类似。

在 AKE 协议中，如果两方各发送一条消息就完成密钥协商，且发送消息前都不需要等待对方的消息，则称为“一轮密钥交换”（One-Round Authenticated Key Exchange，ORKE）。一些经典的 AKE 协议都是一轮密钥交换协议，包括 MQV、HMQV 和 NAXOS 等协议。这些协议对“会话标识”和“匹配会话”等重要概念均有不同的定义。然而，模型中对会话的不同定义对评估 AKE 协议安全性的影响尚不明确。文献[3]对 ORKE 协议的安全模型进行了定义，将 CK、eCK 等安全模型都统一到 ORKE 中。在保证各模型的安全性同时，ORKE 模型极大简化了已有方案的安全证明。其次，利用逐密钥的可恢复函数（Key-wise Recoverable Function，KRF）和被动安全的密钥交换方案，文献[3]实现了 ORKE 的模块化构造，所生成的 ORKE 协议的安全性依赖于 KRF 的安全性。这种模块化构造思想源于 NAXOS、HMQV 和 BJS 等经典 AKE 协议的构造。最后，文献[3]证明了通过对上述模块化构造的实例化可以得到一些满足经典安全模型（CK、CK+和 eCK 等）定义的 AKE 协议。

与基于公钥方法的 AKE 协议相比，基于预共享密钥的 AKE（PSK-AKE）协议更加高效。然而 PSK-AKE 协议需要设计额外的机制满足协议的安全性。前向安全作为 AKE 协议的重要安全属性，要求在用户长期私钥泄露的情况下，仍然保证泄露前所有会话密钥的安全性。我们可以利用公钥密码算法生成临时密钥，长期私钥仅用于认证，由此实现前向安全。然而，在对称密码中，当前实现前向安全的唯一办法是通过初始的长期密钥派生出会话密钥，然后“演化”长期密钥，会话密钥也随之“演化”。文献[4]基于预共享密钥构造轻量级的 AKE 协议。在仅依赖于对称加密原语的情况下，该 AKE 协议能够实现前向安全。文献[4]定义了线性密钥演化（Linear Key Evolution），利用线性密钥演化构造了 3 种协议（LP1、LP2 和 LP3），它们分别具有不同的效率和安全性。其中 LP1 和 LP2 非常轻量，仅需传输一个 MAC 和一个计数器值。文献[4]提出“同步鲁棒性”（Synchronization Robustness，SR）这一安全目标，衡量通信方重新同步的效率。如果诚实执行协议的双方最终计算出的会话密钥不一致，就称该协议不具备 SR 性质。此外，因为线性密钥演化不能保证所有的会话成功实例化，所以并发

的协议会话执行的正确性很难保证。文献[4]利用穿刺伪随机函数（Puncturable PRFs）提出了两种新的构造（PP1 和 PP2），同时实现了前向安全、同步鲁棒性和并发正确性（Concurrent Correctness）。穿刺伪随机函数可以通过哈希函数进行高效地实例化，因此 PP1 和 PP2 都非常轻量。

密钥封装机制（KEM）是一种重要的公钥密码原语。对于部署在多用户场景下的 KEM，其安全模型与单用户场景下的安全模型有较大区别。该模型可以允许敌手询问部分用户的私钥或生成的封装密钥，最后敌手需要对未被腐蚀用户的密钥进行随机性的判定。对应到使用 KEM 的 AKE 协议中，则要求底层的 KEM 支持腐蚀用户（User Corruption）和暴露密钥（Key Reveal）。多用户场景下的 KEM 安全定义（ECPA、ECCA）更加实用，也更加严格。与此同时，判断一个 KEM 是否能达到紧致的 ECPA/ECCA 安全也至关重要。文献[6]研究了一些经典的 KEM 方案是否满足紧致的 ECPA/ECCA 安全，以及如何判断一个 KEM 是否能达到紧致的 ECPA/ECCA 安全。Bader 在 2016 年的欧密会利用元归约的办法证明某些 KEM 方案不可能达到 mCPA/mCCA 下的紧致安全。显然，由于 ECPA（ECCA）安全的 KEM 方案一定是 mCPA（mCCA）安全的，这类 KEM 方案也不可能达到 ECPA（ECCA）下的紧致安全。文献[5]定义 KEM 的“秩”用于判断 KEM 方案是否可能达到紧归约，证明了一个 KEM 方案的秩只要是多项式有界的，该方案就无法达到紧致安全，且归约的安全损失系数为 $\Omega(n)$，其中 n 为用户数量。目前，对于大部分具有 mCPA 下紧致安全的经典 KEM 方案来说，其秩都是多项式有界的，很难找到 ECPA/ECCA 下紧致安全的 KEM 协议。文献[5]还证明，任何具有 mCPA（mCCA）下紧致安全的 KEM 都能达到 ECPA（ECCA）下的安全，损失系数为 $O(n)$。在使用了 KEM 的 AKE 协议中，KEM 的公私钥通常是用户长期公私钥的一部分，AKE 协议的紧致安全与底层 KEM 方案的紧致安全相关。因此，对于这类 AKE 协议，如果想在标准模型下实现紧致安全，KEM 的私钥就不能作为用户长期私钥的一部分。

SIDH 是基于在超奇异椭圆曲线之间寻找同胚的困难性假设的一种后量子 AKE 协议。然而，SIDH 和相关密码系统揭示了额外的信息：将秘密同源性限制为曲线的子群（扭力点信息）。Petit 在 2017 年亚密会上首次证明扭力点信息可以显著降低发现秘密同源的难度。特别地，证明了 SIDH 的“过度拉伸”参数化可以在多项式时间内被打破。然而，这并没有影响文献中提出的任何密码系统的安全性。文献[6]通过利用来自对偶和 Frobenius 同源的额外信息来加强 Petit 的工作中证明扭力点信息可以显著降低发现秘密同源的难度的技术，这大大扩展了扭转点攻击的影响。特别是，作者还提出了一种经典攻击，它完全破坏了 n 方之间的密钥交换，并且给出了该攻击适用的全部参数。此外，文献[6]还构造了抵御该攻击的 SIDH 变体（包括起始曲线的后门选择等），但是该结果不会降低 NIST 第三轮中公钥加密候选算法 SIKE 的安全性或揭示其中的任何弱点。

互联网中用户通常以输入口令的方式进行身份认证，在输入口令前，用户与服务器已建

立一条机密的 TLS 信道。服务器存储用户口令对应的 salt 值以及哈希值，对用户发送的口令进行验证。但这种认证方式导致用户登录时，口令以明文的方式出现在服务器中。基于口令认证的密钥交换协议（PAKE）考虑双方共享一个口令，利用该口令进行密钥交换，所生成的会话密钥需抵抗离线字典攻击。为消除基于 TLS 的口令认证方式的缺陷，Bellovin 提出非对称的 PAKE 协议（aPAKE），但大多数的 aPKE 协议都无法抵抗预计算攻击。Jarecki 在 2018 年欧密会上提出一种抗预计算攻击的 aPAKE 协议 OPAQUE，该协议被考虑纳入 TLS1.3 协议标准。Jarecki 提出加密凭证（Credentials）的思想，将口令输入不经意伪随机函数（OPRF），产生用于加密凭证的密钥，并将加密后的凭证存储在服务器端，认证时用户再从服务器端获取并用口令解密凭证，使得用户和服务器可以运行常规的 AKE 协议做密钥协商。

文献[7]指出，OPAQUE 协议的安全性严重依赖 OPRF，一旦 OPRF 被攻破，用户的口令就会受到离线字典攻击。为避免对 OPRF 的依赖，文献[7]提出 KHAPE 协议。KHAPE 协议使用了一种“黑盒”机制，允许直接用口令加密凭证的同时抵抗字典攻击。KHAPE 协议的核心思想有两个：一是抛弃了用户凭证的认证，采用非承诺加密（Non- Committing Encryption），使得敌手对解密机的询问无助于判断是哪个密钥产生了挑战密文；二是使用了密钥隐藏（Key-Hiding）的 AKE，使得敌手不能识别出用于密钥交换的长期密钥。值得注意的是，文献[7]中所有的安全分析均是在通用可组合（Universal Composability，UC）框架下完成的，且基于理想密码模型（Ideal Cipher Model）。KHAPE 协议具有通用性，能够将任何具有密钥隐藏性质的 AKE 协议“编译”为一个 aPAKE 协议。最后用经典的 HMQV 协议做了一个实例化。先在 UC 模型中证明了 2005 年 CK+安全模型下的 HMQV 协议的密钥隐藏性，然后用 KHAPE 将 HMQV 协议转化为 aPAKE 协议。

IRTF 工作组 CFRG 在 2019 年发起了口令认证密钥交换协议（Password-Authenticated Key Exchange，PAKE）标准化方案征集活动。在各方面平衡考虑下，CPace 协议脱颖而出。在该协议中各方共享同一口令。在随后的标准化工作中，CPace 协议得到了进一步发展，产生了一系列新的协议，但是这些协议在实际环境中的安全性质未能得到充分的认识和分析。针对这个问题，文献[8]在 UC 模型下对 CPace 协议进行了全面的安全分析。该工作考虑了现实场景下的自适应破坏（Adaptive Corruption），并避免了将 CPace 协议中的 Map2Pt 函数视为理想函数。为了将安全证明扩展到针对特定椭圆曲线的不同 CPace 协议变体，该工作采用了一种新方法，将证明所需的假设表示为模拟器可以访问的库。通过允许对证明中使用的假设进行模块化替换，这种新方法避免了对未更改协议部分的重复分析，大大提高了对所有不同 CPace 协议变体进行安全的分析的效率。最终，该工作表明，当前所有 CPace 协议变体都具有自适应 UC 安全性。

除密钥协商之外，分布式密钥生成也是密码学应用研究相关的热点问题，主要用于保护密码系统的密钥安全。第一，文献[9]提出了一个可聚合和公开验证的分布式密钥生成

（Distributed Key Generation，DKG）协议。这种基于阈值加密的方法通常在分布式系统的建立过程中发挥重要作用，包括拜占庭共识、时间戳服务、公共随机信标和数据存档系统。DKG 本质上是将各方变成可验证秘密共享（Verifiable Secret Sharing，VSS）方案的份额分发方。聚合可以由任何一方完成，也可以增量式完成。与之前的公开验证方法相比，文献[9]的 DKG 将最终记录的大小和验证时间从 $O(n^2)$减少到 $O(n\log n)$，其中 n 表示参与方的数量。与之前的非公开验证方法相比，文献[9]的 DKG 利用 Gossip 传播算法来降低验证和通信的复杂性，用户可以暂时离线并且消除“申诉轮”。第二，相对于保密性，文献[9]提出了新的 DKG 安全定义，即健壮性和安全保护，以降低复杂性和克服低效性。在新定义下，文献不仅证明了可聚合 DKG 以及几个现有 DKG（包括 Pedersen 变体）的安全性，还证明了这些现有的 DKG 可以用于产生常用密码系统的安全阈值变体，如 El Gamal 加密和 BLS 签名。第三，提出一个新的高效的可验证不可预测函数（Verifiable Unpredictable Function，VUF），并证明在随机谕言模型下的安全性和对 DKG 的高度适用性。实验表明，每方的开销是线性的。

FIDO2 是快速身份在线（Fast Identity Online，FIDO）联盟提出的无密码用户身份验证（Passwordless User Authentication）标准提案，其目标是启用用户友好的无密码身份验证，以防止网络钓鱼和身份欺诈。其核心思想是依靠安全设备（通过生物识别和/或 PIN 控制），然后可以使用这些设备进行注册，最后无缝地对在线服务进行身份验证。文献[10]对 FIDO2 协议进行了首次可证明的安全分析，涵盖了 FIDO2 的核心组件：W3C 的网络身份验证(Web Authentication，WebAuthn) 规范和新的客户端到身份验证器协议（Clinet-to-Authenticator Protocol，CTAP2）。分析以模块化方式进行，即首先分别分析 WebAuthn 和 CTAP2 组件，然后推导出 FIDO2 的一个典型使用的整体安全性。特别是，本文分析了 CTAP2 存在的安全缺陷，并提出了具有更高安全属性的 sPACA 协议作为代替，而且新提出的协议具有更高的效率。

本节作者：陈荣茂、王毅（国防科技大学）

参考文献

[1] HAN S, LIU S, GU D. Key Encapsulation Mechanism with Tight Enhanced Security in the Multi-user Setting: Impossibility Result and Optimal Tightness[C]. ASIACRYPT 2021(2): 483-513.

[2] HAN S, JAGER T, KILTZ E, et al. Authenticated Key Exchange and Signatures with Tight Security in the Standard Model[C]. CRYPTO 2021(4): 670-700.

[3] XIAO Y, ZHANG R, MA H. Modular Design of Role-Symmetric Authenticated Key Exchange Protocols[C]. ASIACRYPT 2021(4): 742-772.

[4] BOYD C, DAVIES G-T, DE KOCK B, et al. Symmetric Key Exchange with Full Forward

Security and Robust Synchronization[C]. ASIACRYPT 2021(4): 681-710.

[5] JAGER T, KILTZ E, RIEPEL D, et al. Tightly-Secure Authenticated Key Exchange, Revisited[C]. EUROCRYPT 2021(1): 117-146.

[6] DE QUEHEN V, KUTAS P, LEONARDI C, et al. Improved Torsion-Point Attacks on SIDH Variants[C]. CRYPTO 2021(3): 432-470.

[7] GU Y, JARECKI S, KRAWCZYK H. KHAPE: Asymmetric PAKE from Key-Hiding Key Exchange[C]. CRYPTO 2021(4): 701-730.

[8] ABDALLA M, HAASE B, HESSE J. Security Analysis of Cpace[C]. ASIACRYPT 2021(4): 711-741.

[9] GURKAN K, JOVANOVIC P, MALLER M, et al. Aggregatable Distributed Key Generation[C]. EUROCRYPT 2021(1): 147-176.

[10] BARBOSA M, BOLDYREVA A, CHEN S, et al. Provable Security Analysis of FIDO2[C]. CRYPTO 2021(3): 125-156.

4 身份基加密/属性基加密/函数加密/广播加密

属性基加密和函数加密可对访问权限与解密结果进行细粒度控制，在云计算保证安全性中，是近些年的一个研究热点。2021 年，国际三大密码年会上的相关论文共有 9 篇，在属性基加密、函数加密与身份基加密的转换关系，以及不同访问控制策略下的方案构造、更高的安全目标、去权限中心化、基于最小假设的细粒度密码学等方面有所突破，具体总结如下。

属性基加密是实现细粒度访问控制的一种加密方式，分为密文策略属性基加密（Ciphertext Policy Attribute Based Encryption，CP-ABE）和密钥策略属性基加密两种，文献[11]针对 KP-ABE 进行研究。初始阶段属性基加密的研究针对布尔公式或电路等策略，其特点是属性长度固定，这类策略称为非一致策略，不适用属性长度可变的场合。2012 年，Waters 开启了针对一致策略的属性基加密研究，这些策略包括 DFA、NFA、RAM、TM 等，基于双线性映射、LWE 问题及不可区分混淆（iO）等具体数学问题进行方案构造。文献[11]研究了由身份基加密向图灵机策略的受限属性基加密的一般性构造。该构造表明，一致性策略的受限属性基加密存在的充分条件是 IBE 的存在性，即由分解因子、CDH、LWE 等假设均可构造。

文献[11]的核心构造为单次选择密钥安全（只进行一次密钥询问，且该询问在挑战属性之前给出）的 ABE。该核心构造利用了混淆 RAM（Garbled RAM）的构造技术，引入了迭代可模拟安全并进行了安全性证明以适应图灵机策略安全性。在此基础上，该方法利用弱非承诺加密（wNCE）处理挑战密文之后的密钥询问，从而将单次选择密钥安全 ABE 转为单次自适

应安全的 ABE。这种 wNCE 在随机谕言（RO）模型下可由 PKE 实现。进一步地，单次自适应安全的 ABE 可由组合技术提升为受限自适应安全 ABE。文中留下了几个公开问题，如是否可将此构造进一步推广，得到受限询问下安全的函数加密（Functional Encryption，FE）？是否可在标准模型下实现由 IBE 向受限 ABE 的转换，而不依赖理想的 RO 模型？

属性基加密领域的一个挑战是实现其自适应安全性，此前的主要实现方式为双系统加密，适用于基于双线性对群结构。但这些实现均依赖合数阶群上的判定子群假设或判定线性假设。与此相对应，较弱的选择安全属性基加密则可依赖计算假设实现。文献[9]的研究目标即为是否可基于双线性群上的计算假设实现自适应安全的属性基加密。

2019 年，Tsabary 首次基于 LWE 假设提出了自适应安全的属性基加密方案，其构造思路为通过自适应安全的受限伪随机函数（Constrained PseudoRandom Function，CPRF）将选择性安全密钥策略属性基加密（KP-ABE）转为自适应安全密文策略属性基加密（CP-ABE），其中 CP-ABE 的策略与 CPRF 的策略相同。但该方案要求基于 KP-ABE 满足部分明文可计算性，而目前绝大部分 KP-ABE 并不满足该性质。文献[9]提出了一个新的构造框架，要求基础 KP-ABE 方案满足属性可删除性质，即任何人可将关于属性$x \in \{0.1\}^n$的密文转化为关于属性$x' \in \{0,1,\perp\}^n$的密文，其中x'为将x中某些比特替换为$\perp$，即删除了某些比特，在密钥策略−电路 C 并未读到$\perp$时，解密结果不变。同时，该文要求 CPRF 满足“删除一致”性，即计算 CPRF 输出结果的过程与 KP-ABE 的策略−属性保持一致，受限密钥的计算可通过对主密钥的删除操作完成。在 KP-ABE 和 CPRF 满足上述性质时，文献[9]给出了构造自适应安全 CP-ABE 的框架。该框架对 Tsabary 的结果进行了简化和扩展，因此可涵盖更多具体构造，如基于搜索 Diffie-Helleman 和$\mathbf{NC}^1$中伪随机函数的 GPSW 方案[10]和基于 LWE 假设的 Boyen 方案[6]。从而，文献[9]首次基于双线性对上的搜索类假设给出了自适应安全的属性基加密方案。

由于最终 CP-ABE 支持的策略与 CPRF 支持的策略一致，而目前“删除一致”CPRF 仅能支持子集关系策略，因此如何实现表达更丰富的策略是一个公开问题，这可能需要新 CPRF 的构造或新的实现框架。

在多权限属性基加密（Multi-Authority Attribute Based Encryption，MA-ABE）中，系统中存在多个权限中心控制不同的属性集合，每个权限中心可以独立为它所控制的属性分发用户私钥。MA-ABE 是 CP-ABE 的扩展，根据访问控制策略加密消息，拥有满足该策略的属性集合的用户通过与控制这些属性的权限中心进行交互得到解密密钥。在 MA-ABE 中，敌手不仅可能攻破拥有用户私钥的解密方，还有可能攻破权限中心，抗合谋性需要考虑这两者的结合。MA-ABE 相对于 CP-ABE 的优势在于其去中心化特点，因此“真正”去中心化的 MA-ABE 应满足如下特点：系统建立后，任何参与方均可成为权限中心，发布其公钥并为其控制的属性生成私钥。此前“真正”去中心化的 MA-ABE 均在随机谕言模型中的双线性群下构造，受量子计算机发展的威胁。文献[7]研究了抗量子攻击的 LWE 假设下“真正”去中心化 MA-ABE

的实现。

文献[7]首先基于 LWE 假设针对NC^1策略给出了 CP-ABE 方案。此前 LWE 假设下的属性基加密构造思路均为通过同态加密等工具实现 KP-ABE，然后由普适电路将属性和策略对换，从而实现 CP-ABE。文献[7]的构造首次摆脱了同态加密等工具，通过类比双线性群上线性秘密共享的思路实现了 LWE 假设下的 CP-ABE 方案。尽管该方案在参数、效率方面并未优于以往构造，但是它提供了一种新的构造思路，可扩展到 MA-ABE，有其他的潜在优势。

文献[7]将构造 CP-ABE 的方法进一步扩展，得到了基于 LWE 假设针对析取范式策略 DNF 的“真正”去中心化的 MA-ABE。与以往的构造相同，该方案的安全性依赖随机谕言模型。

细粒度密码学是指在敌手资源事先限定，并且诚实参与方拥有资源少于敌手的情况下构造密码系统。在现代密码学中，方案的安全性基于底层问题的平均情况困难性，如基础的单向函数存在性或拥有具体数学结构的分解因子、离散对数或格上困难问题。而细粒度安全性基于困难问题的最坏情况复杂性，如文献[14]中用到的$\mathrm{NC}^1 \subset \oplus L/\mathrm{poly}$假设，其中$\oplus L/\mathrm{poly}$指由所有多项式大小的分支程序构成的语言类，而$\mathrm{NC}^1$中的语言类可以由固定宽度的多项式大小的分支程序表达，因此该假设指存在只能由非固定宽度的多项式尺寸构成的分支结构，更有可能成立，是目前广泛被认可的假设。

细粒度密码学由 Merkle 开启，一个重要的研究目标是探索在更弱的假设下不同密码学原语的存在性。近年来，细粒度安全的单向函数、对称加密、公钥加密、哈希证明系统和非交互零知识证明已被构造，但许多重要的密码学原语如签名的存在性仍旧未知。文献[14]基于$\mathrm{NC}^1 \subset \oplus L/\mathrm{poly}$首次构造了自适应安全的属性基加密。这就意味着即使不存在单向函数，只要上述假设成立，属性基加密就仍旧存在。该构造使用了一般的谓词编码框架和新的证明系统，不再依赖具体的 Diffie-Hellman 假设。根据对谓词编码的具体实例化，可以得到细粒度安全的 IBE、内积策略 ABE、非零内积加密、广播加密、模糊 IBE 等方案。利用类似的技术，文献[14]还给出了高效近似自适应的非交互零知识证明系统（Quasi-Adaptive Non-Interactive Zero Knowledge，QANIZK）。

广播加密是一个活跃的研究领域，其中一个主要目标就是构造参数更小的方案，这里的参数包括公钥、用户私钥和密文尺寸。此前基于双线性对的广播加密参数尺寸只能达到$O(\sqrt{N})$，其中N为用户个数。文献[13]打破了这个限制，基于双线性对上的双边线性假设（Bilateral k-Lin Assumption）实现了参数尺寸为$O(N^{\frac{1}{3}})$的广播加密方案，该方案在非受限合谋下达到了自适应安全性。

在文献[13]中第一次基于双线性对给出了 3 次多项式策略的参数紧凑的密文策略属性基加密（CP-ABE），这里参数紧凑是指函数输入长度为n时，公钥、密文、私钥长度之和为$O(n)$。

然后利用广播加密与 CP-ABE 之间的关系，将用户属于广播列表关系表达为三次函数，从而基于双线性对得到参数尺寸为$O(N^{\frac{1}{3}})$的广播加密。此前，基于双线性对的参数紧凑的 CP-ABE 只能实现二次多项式策略，为打破这个障碍，文献[13]借鉴了 FE 中的二次项重建技术，从而降低参数尺寸。

近年来，在参数更为紧凑的广播加密方面有不少尝试，如 Boneh 等基于多线性映射、Agrawal 和 Yamada 基于双线性对加 LWE、Brakerski 和 Vaikuntanathan 基于 LWE 给出了$\mathrm{poly}(\log N)$参数尺寸的广播加密，但这些构造要么依赖的假设公信力不足，要么需要两个不同类别的假设，要么缺少严格的安全性证明。在格密码领域，如何仅仅基于 LWE 假设构造参数尺寸为$O(N)$的广播加密尚是一个公开问题。

由于其解密结果只揭示消息运算后的结果，函数加密近年来广受关注。文献[3]首次给出了多输入的二次函数加密，基于双线性群上的标准假设达到了不可区分选择安全性。多输入的函数加密（Multi-Input Functional Encryption，MIFE）是函数加密（Functional Encryption，FE）的扩展。在 MIFE 中，n 个用户独立加密消息$x_1,\cdots,x_n$得到$c_1,\cdots,c_n$，拥有函数密钥sk_f的解密者只能得到$f(x_1,\cdots,x_n)$而无其他信息。与公钥加密和对称加密类似，根据加密算法是否需要秘密密钥，MIFE 分为公钥 MIFE 和私钥 MIFE。公钥 MIFE 不能蕴含私钥 MIFE，这是由于在公钥MIFE中，给定x_1的密文c_1，sk_f的拥有者可以对任意$x_2',\cdots,x_n'$计算出$f(x_1,x_2',\cdots,x_n')$，比在私钥 MIFE 中泄露的信息多。由于不同用户加密使用的随机数不同，MIFE 的实现相对于 FE 要困难得多。利用内在的信息泄露，文献[3]证明了公钥二次函数的 MIFE 可由公钥内积函数的 MIFE 构造。对称二次函数的 MIFE 实现难度则大得多，下面简单介绍对称二次函数的 MIF 实现思路，提到的 MIFE 均为对称 MIFE。

文献[3]借鉴了单输入二次函数加密的实现思路，利用函数隐藏的内积函数加密对输入分别产生密文和函数密钥，这样只需线性个数的密文就可使解密结果产生全部二次项，进而利用另一个内积加密方案对这些二次项进行干扰和消干，从而保证解密结果的正确性。在多输入场景下，敌手可利用不同用户掌握的信息进行组合攻击（Mix and Match）。为处理这种情况，文献[3]提出并构造了函数隐藏的谓词内积函数加密（pIPFE，内积加密和内积函数加密的结合）。通过在方案构造中额外引入 pIPFE 和混合群的内积函数 MIFE，最终实现了二次 MIFE。该文献留下了以下公开问题：更大函数族如何用来实现不可区分混淆（iO）的 2.5 次函数的 MIFE；更好的自适应安全性，缩小密文和密钥尺寸等。

函数加密（FE）的一个主要安全性要求为抗合谋，即敌手不能通过合谋获得额外信息。根据合谋方个数限制可分为非受限抗合谋和受限抗合谋。非受限抗合谋的通用 FE 功能强大，与 iO 关系密切，到目前为止的构造需要综合双线性群上假设与格上假设，结构复杂。受限抗合谋 FE 具有以下缺陷：①需要在系统建立之初便确定合谋上限Q，这就要求所有加密数据处

于同一安全等级，部署跨安全等级要求的系统时只能依从最高要求，影响执行效率；②在公钥领域，目前的构造只能支持电路策略，电路策略为非一致性策略，要求输入为固定长度，限制了适用场景，因此需寻求适配度更高的一致性策略下的方案，如图灵机（Turing Machine，TM）、非确定性对数空间（Non-deterministic Logarithm space，NL）等；③目前的构造均针对密钥策略函数加密（KP-FE），而在许多使用场景下更自然的密文策略函数加密（CP-FE）均需通过普适电路（Universal Circuit，UC）进行转换，受到颇多限制。

针对以上研究背景，文献[5]提出了动态受限抗合谋的函数加密，即在系统建立之初无须确定合谋个数上限，而加密执行者可根据数据（策略）的安全等级由合谋个数自适应地调整加密算法，从而只有密文尺寸依赖合谋个数，公共参数尺寸不被合谋个数影响。此外，文献[5]针对电路、TM 等策略在 CP-FE 和 KP-FE 方面均给出了一系列研究成果，具体如下。

（1）CP-FE 方面：给出了第一个动态受限的电路策略的 CP-FE 方案。基于 IBE 假设，为尺寸、输出长度、深度不受限的电路策略构造了非自适应模拟安全的 CP-FE 方案（这里非自适应指限制敌手在得到挑战密文之前进行密钥提取询问）。将假设进一步增强，基于接收者可打开（RSO）安全的 IBE，为尺寸、输出长度、深度受限的电路策略构造了自适应模拟安全的 CP-FE 方案。基于具体的 LWE 假设，为电路构造了简洁的自适应安全的方案（此处简洁指电路尺寸不受限，但输出长度和深度固定）。

（2）KP-FE 方面：首次为电路策略给出了简洁的 KP-FE 构造。该构造基于 LWE 假设，达到了自适应模拟安全性，相对之前的简洁构造提高了安全性。

（3）针对一致性策略 TM/NL 的 KP/CP-FE 方面：在公钥场景下，基于 LWE 假设首次为 TM 策略给出了受限合谋下安全的 KP-FE 和 CP-FE。该构造可达到非自适应模拟安全性，其输入和机器尺寸均不受限，但密文尺寸随 TM 在给定输入下运行时间的增长而增长。在较小的 NL 计算类下，给出的构造可达到自适应模拟安全性。

属性加权和的函数加密由 Abdalla 等人提出[4]。在这种 FE 方案中，加密算法的输入为N对属性$\{x_i, z_i\}_{i\in[N]}$，解密密钥的输入为权重函数f，解密结果为$\sum_{i\in[N]} f(x_i)z_i$。在上述定义中，指标i为槽，$\{x_i\}$为公开属性，$\{z_i\}$为秘密属性。在$N = 1$时，称函数加密为单槽方案，当公共参数设定后加密算法仍可随意选择 N 时，称 FE 为槽数不受限的方案。属性加权和包含了现实中的一系列函数，如特定族群的平均工资、烟民患癌的平均年龄等数据统计，受到了广泛关注。Abdalla 等人基于 k-Lin 假设给出了槽数不受限的属性加权和函数加密，但只达到了较弱的半自适应安全性，即只能在挑战密文之后进行私钥提取询问。文献[8]第一次给出了自适应安全的属性加权和 FE，即在挑战密文前后均可进行私钥提取询问。与文献[4]相同，这里的权重函数为算数分支程序。

文献[8]首先给出了基于 k-Lin 假设的单槽方案，在进行受限次挑战密文询问和不受限次私钥提取询问时达到了自适应模拟安全性。由于模拟安全性蕴含不可区分安全性，在不可区

分条件下，单次挑战密文安全性蕴含多次挑战密文安全性，上述方案可在非受限次数挑战密文和私钥提取询问下达到自适应不可区分安全性。该构造主要运用了算数密钥混淆（Arithmetic Branching Programs，ABP）和满足函数隐藏性质的带槽内积函数加密。

在单槽方案的基础上，文献[8]改进了文献[4]提出的由单槽方案向槽数不受限方案的自举框架，引入扩展的单槽 FE 和三槽结构，将 k-Lin 假设改为双边 k-Lin 假设，从而处理自适应安全性游戏中在挑战密文之前进行私钥提取询问带来的障碍，最终实现了槽数不受限的属性加权和算数分支程序函数加密。该方案在受限次挑战密文和挑战前私钥提取询问，不受限次挑战后私钥提取询问下达到了自适应安全性。该方案的公共参数和私钥长度均不依赖槽数 N，但与挑战前的私钥提取次数相关。

函数加密的属性和策略可用二元方式表达，即将密文输入写作(x,μ)，将私钥提取输入写作(P,g),当$P(x)=1$时解密可恢复$g(\mu)$，否则解密中止。当 P 是布尔电路，g保持恒等函数即可表达所有布尔函数加密，如文献[1]中构造的函数加密中所使用；当g为内积函数，P 为相等性测试或属性策略时，即为文献[2]中身份基内积函数加密和属性基内积函数加密。在这种二元表达下，安全游戏中进行私钥提取询问时满足$P(x^*)=1$的询问称为 1 询问；否则，称为 0 询问。进一步，（3,5）-SIM 安全性表示在进行 3 次 1 询问、5 次 0 询问时，方案可达到模拟安全性。

文献[1]基于 LWE 构造的函数加密具有深刻的理论价值，但有以下限制：所有 1 询问只能在进行挑战密文询问之前一次性给出。为打破这个限制，文献[12]给出了对满足关系$(A|AR)y=u$的四元组(A,AR,u,y)的新的两步抽样算法，可在无陷门时首先确定(A,u)后确定(R,y)。通过在方案设计和证明中使用新提出的抽样算法，文献[12]实现了（Q，poly）半自适应模拟安全性，可在挑战密文前、后进行任意次私钥提取询问，并且该方案满足密文简洁性，仅需增加$O(Q)$开销。

通过在私钥提取和安全性证明中应用该二步抽样算法，文献[12]将文献[2]中的身份基内积函数加密由（1，poly）选择安全性提高到自适应安全性，并实现了（Q，poly）半自适应安全的属性基内积函数加密。

本节作者：贾仃仃（中国科学院信息工程研究所）

参考文献

[1]AGRAWAL S. Stronger Security for Reusable Garbled Circuits, General Definitions and Attacks[C]. CRYPTO 2017 (1): 3-35.

[2] ABDALLA M, CATALANO D, GAY R, et al. Inner-Product Functional Encryption with Fine-Grained Access Control[C]. ASIACRYPT 2020 (3): 467-497.

[3] AGRAWAL S, GOYAL R, TOMIDA J. Multi-input Quadratic Functional Encryption from Pairings[C]. CRYPTO 2021 (4): 208-238.

[4] ABDALLA M, GONG J, WEE H. Functional Encryption for Attribute-Weighted Sums from k-Lin[C]. CRYPTO 2020 (1): 685-716.

[5] AGRAWAL S, MAITRA M, Narasimha Sai Vempati, et al. Functional Encryption for Turing Machines with Dynamic Bounded Collusion from LWE[C]. CRYPTO 2021 (4): 239-269.

[6] BOYEN X. Attribute-Based Functional Encryption on Lattices[C]. TCC 2013: 122-142.

[7] DATTA P, KOMARGODSKI I, WATERS B. Decentralized Multi-authority ABE for DNFs from LWE[C]. EUROCRYPT 2021 (1): 177-209.

[8] DATTA P, PAL T. (Compact) Adaptively Secure FE for Attribute-Weighted Sums from k-Lin[C]. ASIACRYPT 2021 (4): 434-467.

[9] GOYAL R, LIU J, WATERS B. Adaptive Security via Deletion in Attribute-Based Encryption: Solutions from Search Assumptions in Bilinear Groups[C]. ASIACRYPT 2021 (4): 311-341.

[10] GOYAL V, PANDEY O, SAHAI A,et al. Attribute-based encryption for fine-grained access control of encrypted data[C]. CCS 2006: 89-98.

[11] GOYAL R, SYED R, WATERS B. Bounded Collusion ABE for TMs from IBE[C]. ASIACRYPT 2021 (4): 371-402.

[12] LAI Q, LIU F, WANG Z. New Lattice Two-Stage Sampling Technique and Its Applications to Functional Encryption - Stronger Security and Smaller Ciphertexts[C]. EUROCRYPT 2021 (1): 498-527.

[13] WEE H. Broadcast Encryption with Size N1/3 and More from k-Lin[C]. CRYPTO 2021(4): 155-178.

[14] WANG Y, PAN J, CHEN Y. Fine-Grained Secure Attribute-Based Encryption[C]. CRYPTO 2021(4): 179-207.

5 同态加密

同态加密（Homomorphic Encryption，HE）支持无须解密的情况下在加密数据上进行计算，从而适用于将计算安全地外包给不可信的云。其中，能够支持任意次数加法和乘法运算的同态加密方案被称为全同态加密（Fully Homomorphic Encryption，FHE）。出于安全考虑，所有已知的全同态加密方案构造都需要产生有噪声的密文。密文中的噪声会随着同态运算不

断累加，可能导致噪声超过上界而无法正确解密。自举（Bootstrapping）是用于刷新密文的通用技术，也是目前实现真正的全同态加密方案的唯一途径。

文献[1]：在同态标量乘法中，控制噪声增长的一个标准技巧是根据小基数B分解标量$k \doteq \sum_{i=0}^{n} k_i' B^i$，然后利用预先计算的密文$\mathrm{Enc}(B^i x)$得到$c \leftarrow \sum_{i=0}^{n} k_i' \mathrm{Enc}\left(B^i x\right)$。当$-(B-1) \leqslant k_i' \leqslant B-1$时，$(k_n', \cdots, k_0')_B$被称为$k$的 modified radix-B form。由于上述密文中噪声的方差与$(k_n', \cdots, k_0')_B$的欧几里得重量$\sum_{i=0}^{n} k_i'^2$成正比，因此寻找能够最小化该值的 modified radix-B form 就可以有效控制密文中的噪声传播。

为了使得上述技术达到最优的噪声控制，文献[1]定义了一种新的 modified radix-B form，称为平衡的非邻接形式（Balanced Non-Adjacent Form，BNAF）。该方案证明每个整数有且仅有一个 BNAF，且在给定整数k的所有 modified radix-B form 中，该 BNAF 能够使$\sum_{i=0}^{n} k_i'^2$最小化。同时，文献[1]中提出了简单的评判标准来检验一个 modified radix-B form 是否为 BNAF，进一步提出了生成 BNAF 的有效算法，并研究了 BNAF 在渐近和精确情况下的统计特性。该定义还被扩展到模整数情况下，并拥有类似的最优性结果。

文献[2]：Brakerski-Gentry-Vaikuntanathan（BGV）和 Brakerski-Fan-Vercauteren（BFV）方案是两种主要的执行有限域和整数上精确运算的 FHE 方案。虽然它们在相同的明文空间$\mathbb{Z}_p$（其中正整数p表示明文模数）下工作，但是使用了不同的策略对$\mathbb{Z}_p$中由整数组成的消息进行编码和噪声控制。BGV 方案将消息编码在$\mathbb{Z}_q$（其中正整数q表示密文模数）的最低有效位，并采用模切换技术使噪声量级保持不变；BFV 方案则将消息编码在$\mathbb{Z}_q$的最高有效位，密文乘法时不使用模切换技术，但需要进行q/p倍的缩放。

BGV 和 BFV 在噪声管理策略上的差异导致了两种方案的效率和噪声增长上的不同。文献[2]通过优化噪声增长、核心算法的计算复杂性和可用性，对 BGV 和 BFV 方案进行了改进来缩小两种方案之间的差距。同时比较了两种方案的理论复杂性，并在 PALISADE 库中实现了改进的 BGV 和 BFV 方案以评估它们的实验性能。实验结果表明，改进后的 BGV 实现对于中、大型的明文模数而言更快，这类模数经常在需要密文打包的实际场景中使用，而改进后的 BFV 实现在小型的明文模数情况下更快。

文献[3]：Chillotti-Gama-Georgieva-Izabachène（CGGI）方案，也称为 TFHE 方案，是一种支持快速自举的 FHE 方案，其自举技术还可以在降低噪声的同时对一个表示为查找表的单变量函数求值，称为可编程自举（Programmable Bootstrapping，PBS）。然而，在 TFHE 中明文的最高有效位需要被设置为零或被预先知晓，导致存储消息的空间会有 1 比特的浪费，因此在许多用例中，这意味着更大的计算开销。

为了克服上述限制，文献[3]提出了一种无须填充的可编程自举（Programm-able Bootstrapping Without Padding）。其主要技术在于将 BFV 型 FHE 方案的同态乘法整合进 TFHE 中，即先进行张量积再进行重线性化操作，并对 TFHE 的 PBS 进行了推广。后者提供了在自举时选择加密明文中任意比特块的灵活性，且可以在无须额外计算和噪声的同时计算多个查表函数。这两种技术可以用于改进 TFHE 的门自举和电路自举。例如，在计算布尔电路时无须在每次同态门运算后进行自举，甚至可以支持大整数模运算和精确运算的高效计算。最后，作者介绍了两种新的方法用来自举高精度密文。

文献[4]：Cheon-Kim-Kim-Song（CKKS）方案是一种 IND-CPA 安全的支持定点数运算的近似同态加密方案。文献[4]在 IND-CPA 模型的基础上，给敌手提供了严格限制的解密谕言，首次提出了针对 CKKS 方案的被动攻击模型 IND-CPA$^{\mathrm{D}}$，并且从理论和实验角度证明了 CKKS 方案不是 IND-CPA$^{\mathrm{D}}$ 安全的。文献[4]还证明了对于精确加密方案，IND-CPA 安全与 IND-CPA$^{\mathrm{D}}$ 安全等价；对于近似加密方案，IND-CPA 安全不一定意味着 IND-CPA$^{\mathrm{D}}$ 安全。

具体来说，当用 IND-CPA$^{\mathrm{D}}$ 模型攻击 CKKS 方案时，敌手询问$m = 0$的密文，得到$ct = (a, b = \langle s, a\rangle + e)(\mathrm{mod}\ q)$（其中，$a$是公开参数，$s$是私钥，$e$是加密噪声，$q$是密文模数），再询问$ct$的解密结果，得到$e$。由于$b - e = \langle s, a\rangle(\mathrm{mod}\ q)$，敌手通过重复上述步骤获取多个这样的线性等式，用高斯消去法以高概率恢复私钥s。通过以上方式，文献[4]对目前主流的同态计算开源代码库 HEAAN、RNS-HEAAN、SEAL、HElib 和 PALISADE 中的 CKKS 方案进行了攻击。当维数取 2^{16}、q取 350 比特、f为机器学习常见函数时，攻击 256 比特安全的 CKKS 方案仅需小于一分钟甚至几秒。另外，文献[4]也提供了一些可能抵御该攻击的方法，如对解密结果添加高斯噪声、利用噪声泛洪技术添加确定性噪声等。

文献[5]：针对 CKKS 方案的自举与一般同态加密方案的自举不同，前者是通过增大密文模数；而后者是通过降低噪声来更新密文。CKKS 方案自举的主要步骤包括提高模数、同态编码、同态计算模函数和同态译码。提高 CKKS 方案的自举效率也是现在的研究热点之一，目前的 CKKS 方案及其自举的实现，都使用了稀疏私钥（小汉明重量）来保证高效性，但是最近出现的针对稀疏私钥的（R）LWE 密文的攻击，使大多 CKKS 方案的实现变得不够可靠。另外，CKKS 方案的完全剩余数系统（Residue Number System）变体——RNS-CKKS 方案支持全程在较小模数上并行使用数论变换（Number Theory Transform，NTT）加速，来减少多项式乘法代价。

文献[5]提供了首个能够抵抗稀疏私钥攻击的高效 RNS-CKKS 方案的自举算法。该算法采用混合型密钥切换和提升（Hoisting）技术，改进同态明文槽循环置换过程，显著减少了 RNS 形式密文的重建次数。然后，通过对矩阵-向量乘法的 Baby-Step Giant-Step（BSGS）优化算法使用 Double-Hoisting 技术，进一步降低了同态编（译）码的计算量。至于同态模约简过程，文献[5]沿用前人的方法逼近模约简函数，还通过仔细分析多项式求值过程中明文的缩放因子

变化，利用树状计算结构，改进了对使用 Chebyshev 基的多项式进行求值的 BSGS 算法，减少了对 RNS 形式密文进行近似模切换导致的误差。

最后，文献[5]基于 Lattigo 库提供了首个 RNS-CKKS 方案自举算法的开源实现代码。实验结果表明，在 128 比特安全强度下使用稠密私钥（汉明重量$h = 192$），对加密了 32768 个复数的密文进行自举仅需 18 秒，输出密文模数达 505 比特，对应解密结果的平均精度达 19.1 比特，自举失败概率仅为 $2^{-15.58}$，综合性能比先前最优结果（$h = 64$）好得多。

文献[6]：从提高精度的角度，改进了 RNS-CKKS 方案的自举算法。首先，文献[6]改进了多区间 Remez 算法，该算法是基于 Chebyshev 交替定理、迭代逼近任意函数的算法，文献[6]在此基础上修改了迭代参考点的选取准则，使之能够更高效地获得有限个任意区间并集上任意函数的最优近似多项式。其次，由于模函数可以用正/余弦函数与逆正弦函数的复合近似，文献[6]将该复合关系与改进的多区间 Remez 算法结合，再利用倍角公式技巧，缩小近似函数与模函数之间的差距，进一步缩小了自举误差。

文献[6]基于 SEAL 库的实验结果表明，自举误差与（逆）正弦函数的近似多项式次数、自举过程的缩放因子、该因子与非自举过程的缩放因子的比值、密文打包的消息个数等有关，自举后消息精度可以达 32.6~40.5 比特，使得 RNS-CKKS 方案能够支持训练深度神经网络等更广泛的应用。另外，文献[6]与文献[5]的技术兼容，能够抵抗针对稀疏私钥的 RLWE 密文的攻击。

文献[7]：在客户端-服务器模型中，为了解决 FHE 方案密文扩张大和客户端计算过载的问题，一种称为 transciphering framework 的用于精确计算的混合框架被提出。在该框架中，客户端使用对称密码E在密钥k下加密消息m，并使用 HE 算法$\mathrm{Enc}^{\mathrm{HE}}$加密密钥$k$，得到的密文$c = E_k(m)$和$\mathrm{Enc}^{\mathrm{HE}}(k)$被存储在服务器中。当服务器需要执行同态操作时，首先计算同态加密密文c得$\mathrm{Enc}^{\mathrm{HE}}(c)$，然后利用$\mathrm{Enc}^{\mathrm{HE}}(k)$对$\mathrm{Enc}^{\mathrm{HE}}(c)$执行同态的$E^{-1}$操作。这样可以安全地得到$\mathrm{Enc}^{\mathrm{HE}}(m)$，便于进行同态计算。只要对称密码$E$具有低乘法深度和复杂度，客户端在时间和内存方面就可以显著节省计算资源，并且客户端和服务器之间的通信负载将显著低于只使用了同态加密方案的情况。

文献[7]提出了一种新的支持 CKKS 方案的 transciphering framework，称为 RtF（Real-to-Finite-field）框架。该框架结合了 CKKS 和 FV 方案以支持客户端加密实数，并设计了一种新的流密码 HERA 作为中间的转换方案。HERA 使用了简单的随机化密钥编排算法，相比于现有 transciphering framework 中使用的对称密码需要更少的随机比特。最后，作者评估了 RtF 框架与 HERA 密码相结合的性能，相比于只使用 CKKS 的方案，该方案的密文膨胀率缩小了原来的 1/23，在延迟和吞吐量方面分别提速 9085 倍和 17.8 倍。

文献[8]：可否认加密适用于用户想用假明文消息伪装成被要求解密的明文消息的场景，可应用于避免电子投票行贿、保护私密证据等实际需求。文献[8]首次将 FHE 与可否认加密

结合，构造了可否认全同态加密方案。具体来说，先将 FHE 解密算法修改如下：计算内积$\langle ct, sk\rangle \bmod q$后，判断是否满足$|\langle ct, sk\rangle \bmod q| > B$（其中，$ct$为密文，$sk$为私钥，$q$为密文模数，$B$为能够正确解密的噪声大小上界），若是则输出 0，否则正常解密。因此，通过适当设置明文模数p与密文模数q，密文空间$\mathcal{C}$中任意元素能以极高概率被自举刷新成明文为 0 的小噪声密文。

接着，文献[8]修改 FHE 加密算法为：对于明文m，先从$\{0,1,\cdots,\delta-1\}$中随机选取i，若$m=0$且i为偶数，若$m=1$且i为奇数，计算i个明文为 1 的密文，否则重新选取i；再从密文空间$\mathcal{C}$中均匀随机取$\delta-i$个元素作为明文为 0 的密文（于是敌手发现用户对明文造假的概率为$O(1/\delta)$）；然后将这些密文自举，再同态计算模 2 求和电路，得到m的密文ct。用户当被敌手胁迫揭露明文m时，可以将加密时所用的某个明文为 1 的密文谎称是从密文空间中均匀随机选取的，从而让敌手相信ct是假消息$m^* = m \text{ XOR } 1$的密文。文献[8]还构造了允许对加密算法撒谎的、比上述可否认全同态加密方案更高效的弱可否认全同态加密方案，以及支持整数加密的可否认全同态加密方案，并且论证了无法构造接收方可否认的全同态加密方案。

本节作者：王丽萍（中国科学院信息工程研究所）、徐妍、徐科鑫

将学校名称挪到最后一个人名的后边

参考文献

[1] JOYE M. Balanced Non-adjacent Forms[C]. ASIACRYPT (3). Springer, 2021: 553-576.

[2] KIM A, POLYAKOV Y, ZUCCA V. Revisiting Homomorphic Encryption Schemes for Finite Fields[C]. ASIACRYPT (3). Springer, 2021: 608-639.

[3] CHILLOTTI I, LIGIER D, ORFILA J-B, et al. Improved Programmable Bootstrapping with Larger Precision and Efficient Arithmetic Circuits for TFHE[C]. ASIACRYPT (3). Springer, 2021: 670-699.

[4] LI B, MICCIANCIO D. On the Security of Homomorphic Encryption on Approximate Numbers[C]. EUROCRYPT (1). Springer, 2021: 648-677.

[5] BOSSYAT J-P, MOUCHET C, TRONCOSO-PASTORIZA J, et al. Efficient Bootstrapping for Approximate Homomorphic Encryption with Non-sparse Keys[C]. EUROCRYPT (1). Springer, 2021: 587-617.

[6] LEE J-W, LEE E, LEE Y, et al. High-Precision Bootstrapping of RNS-CKKS Homomorphic Encryption Using Optimal Minimax Polynomial Approximation and Inverse Sine Function[C]. EUROCRYPT (1). Springer, 2021: 618-647.

[7] CHO J, HA J, KIM S, et al. Transciphering Framework for Approximate Homomorphic Encryption[C]. ASIACRYPT (3). Springer, 2021: 640-669.

[8] AGRAWAL S, GOLDWASSER S, MOSSEL S. Deniable Fully Homomorphic Encryption from Learning with Errors[C]. CRYPTO (2). Springer, 2021: 641-670.

（后）量子密码

1 量子密码/量子安全

Fujisaki-Okamoto（FO）是一类将选择明文安全性转化为选择密文安全性的通用范式，NIST 第三轮的后量子密钥封装机制（Key Encapsulation Mechanism，KEM）提案均采用 FO 范式来实现选择密文安全性。为了评估这些 NIST 后量子 KEM 提案的抗量子安全性，一系列的工作（如 Jiang Haodong 等人 2018 年的美密会工作）开展 FO 范式在量子随机谕言模型（证明密码方案抗量子安全性的一种典型模型）下的安全性证明研究。然而，在标准安全假设下，当前工作中已知的（黑盒）安全性证明普遍存在二次的归约损失，关于该归约损失能否进一步提升是 NIST 后量子密码标准化过程中一个重要的公开问题。针对该问题，Jiang Haodong 等人[1]在 2021 年的亚密会上给出了黑盒归约的紧性下界，证明了当前已知的黑盒归约不可避免会引入二次的归约损失，进而解释了为什么当前的黑盒归约在归约损失次数方面没有进展。此外，Jiang Haodong 等人进一步将归约紧性下界推广至更为一般的 OW2H（抗量子安全性证明中应用最广泛的一类基础引理），证明了利用随机谕言机将区分问题转化为搜索问题的过程中黑盒归约不可避免会引入二次的归约损失。

不可区分性一般指以下两种情形：不可区分性（Indistinguishability）和无差别性（Indifferentiability）。在不可区分性场景下，敌手被赋予访问黑盒置换的权限，并试图区分它是使用随机密钥的分组密码还是一个随机置换。与不可区分性的传统概念不同，无差别性假设可能的对手可以访问有关所涉及系统的内部状态的附加信息，如从哈希函数族中选择成员的公共参数。

无差别性用于分析理想化对象结构的安全性，如随机谕言机或理想的密码。它在密码的安全性分析中扮演着十分重要的角色，是可证明安全理论的基石。对于无差别性，Ristenpart 等人提出了一种强化的概念，称为重置无差别性（Reset Indifferentiability）。

重置无差别性适用于更多场景，但由于缺乏有效的结果而在很大程度上被密码学家们放弃。在亚密会 2021 上，Zhandry Mark 重新关注了重置不可区分性[2]，并将其应用到后量子安全中。他的结果包含 3 个方面：第一，他发现在弱重置无差别性下，理想的密码意味着固定大小随机谕言机，并且域收缩是可能的，因此重置无差别性可能会比原先预料的结果还要有用；第二，Zhandry Mark 将经典环境下的分析迁移到量子环境中，表明理想密码就意味着量

子无差别性场景下的随机谕言机；第三，尽管存在 Shor 算法，但 Zhandry Mark 观察到通用群在量子上仍然有意义，即它们与理想密码（重置）存在量子无差别性，其中通用群是密码群的理想化。通过结合上述内容，Zhandry Mark 发现密码群能产生后量子对称密钥密码加密方案。特别地，他通过使用有限域乘法群的子集积与两个模约化操作，获得了一个较为合理的后量子随机谕言机。

现金的一个关键性质是任何人都能在不与银行通信的条件下本地验证货币。由于任意经典比特串都能被复制，因此经典数字货币不能真正实现这一性质。而量子世界存在不可克隆定理，所以可能实现量子货币的不能伪造性。鉴于公钥量子货币难以实现，文献[3]提出一个容易实现的量子货币的替代物——特许量子货币（Franchised Quantum Money）。在特许量子货币中，银行产生货币和验证密钥、管理货币系统。系统中其他参与者都是不可信用户，相互之间可以发送和接收货币。每个用户都能从银行获得一个独一无二的保密的验证密钥。密钥帮助用户验证接收到的货币，只有有效货币才能通过验证并被接受。恶意用户试图欺骗其他用户并让其接受无效货币。特许量子货币具有量子货币的局域可验证性，即任何人都可在不与银行通信的条件下本地验证接受到的货币。此特许量子货币在单向函数假设下是安全的，可抵抗伪造攻击和蓄意破坏攻击。恶意用户得不到其他用户的验证密钥，伪造让其他用户接收的无效货币是困难的，所以此特许量子货币可抵抗伪造攻击。另外，恶意用户不能产生一个让某用户接受而另外用户拒绝的货币，所以此特许量子货币可抵抗蓄意破坏攻击。

“删除可验证的量子加密”是量子情形下特有的一种密码技术，最早由 Broadbent 和 Islam 在 2020 年提出，它具有以下特性：密文都是量子态，当密文被接收方删除时，接收方能产生一串经典比特，作为该密文已经被其删除的证据。虽然他们所设计的方案是信息论安全的，但是存在 3 点局限：要求预先共享密钥且密钥是一次性使用的；要求使用量子通信信道；用于验证该证据的密钥是不可公开的，即该证据不可公开验证。在文献[4]中，Hiroka T 等人针对这些局限性重新开展研究，提出了多个具有“删除可验证”属性的量子加密方案，部分方案仍然要求使用量子信道，部分方案仅要求经典信道。在使用量子信道的情况下，他们设计了密钥可重用的公钥加密和属性加密方案，但这两个方案仍然不具有公开可验证性。在仅使用经典信道的情况下，他们设计了两个交互式的加密方案，其中一个具有公开可验证性。由于收发双方只能使用经典信道，发送方只能通过与接收方进行交互来实现量子密文的生成。与 Broadbent 和 Islam 的方案相比，Hiroka 等人的所有方案都依赖于计算困难假设，不可能达到信息论安全性。

无论是在现实世界中还是在理论上，随机谕言模型（ROM）都已经成为证明密码系统合理性的关键工具。自从引入 ROM 的概念以来，密码学家可以证实有效实用的密码系统是安全的，而在标准模型中，针对密码的可证明安全对人们来说是难以捉摸的。一般来说，ROM 允许更严格且概念上比标准模型安全性证明更简单的证明。随着后量子密码学的出现和量子

敌手的引入，ROM 在量子环境中被推广为可访问的量子随机谕言模型（QROM）。然而，QROM 迄今为止未能在许多环境中发挥出其特定优势。在亚密会 2021 上，Grilo 等人[5]专注于 QROM 的自适应可重编程性。

经典 ROM 的一个理想特性是，当敌手第一次用查询谕言机时，可以选择任意的谕言值。在不知道某些秘密信息的情况下，该事实经常被简化的模拟安全游戏所利用。只要新的值均匀分布且与敌手 A 的其余观点一致，敌手 A 就不会意识到重新编程。这一性质称为自适应可编程性。然而，在量子环境中，以叠加态方式查询谕言机的能力使以前这种简单的方法更加复杂。叠加态查询可以被视为一次可能包含所有输入值的查询。因此，第一个答案可能已经包含有关每个值的信息，这些信息可能需要随着游戏的进行而重新编程。尚不清楚是否有可能在不改变敌手观点的情况下自适应地重新编程量子随机谕言机。我们能否严格证明自适应重新编程也可以在量子随机谕言模型中完成？

Grilo 等人[5]的工作回答了这个问题。他们通过证明一个在区分随机谕言机是否已被重新编程时的抵抗性优势的界限，从而表明将自适应重新编程直接量子化是可行的。在文献[5]中，他们通过提供一个匹配攻击来证明他们的攻击界是紧的。此外，在以下的 3 个 QROM 的应用程序中，他们证明了 QROM 会保持和 ROM 同样的优势：①他们对使用消息压缩程序的 XMSS 的安全性提供了更严格的证明；②他们的工作表明针对 Fiat-Shamir 签名的选择消息安全性的标准 ROM 证明可以直接应用到 QROM 上，从而实现更严格的归约；③针对模糊 Fiat-Shamir 变换，他们给出了第一个针对错误注入和 nonce 攻击的 QROM 安全性证明。

密码的形式化定义对保护软件版权是一个重要但没引起足够重视的课题。由于经典程序易被复制，因此经典密码学的本质定义很难保证，由此引发利用量子计算技术保护软件版权的想法。Aaronson 将软件模型化为布尔函数，利用量子不可克隆定理，创造性地引入量子复制保护的概念，从而解决软件反盗版问题。量子复制保护的定义大致为：给定一个计算函数f的量子态，敌手不能产生两个都可计算函数f的量子态。该定义防止敌手根据软件产生盗版并分发盗版的行为。尽管量子复制保护防止了软件盗版，但是直到目前仍然没有对任意函数类的可证明安全的量子复制保护。现有的量子复制保护只在谕言模型下是安全的，或者对于点函数这类简单函数是启发式的。针对一个相对弱的应用环境，文献[6]形式化定义另一可替换概念——安全软件租赁（Secure Software Leasing，SSL），防止敌手产生能通过认证的盗版软件。SSL 的定义为：授权方给用户提供一个经典线路C对应的量子态ρ_C，用户对量子态ρ_C操作可计算出任意输入下经典线路C的输出；当租赁到期后，用户需根据租赁合同在规定时间内返回量子态ρ_C；量子态返回后，用户就不能再计算线路C的输出。针对线路集$\mathcal{C}$的 SSL 是一个量子多项式时间算法$(Gen, Lessor, Run, Check)$：对$Gen(1^\lambda)$输入安全参数$\lambda$，输出为保密密钥$sk$，可用于授权方在租赁到期后验证返回态的有效性；对任意$\mathcal{C}$中的线路C:$\{0,1\}^n \to \{0,1\}^m$，$Lessor(sk, C)$输出量子态ρ_C，Run可利用ρ_C估计C；对任意$x \in \{0,1\}^n$，用户执行算法

$Run(\rho_C, x) = C(x)$。$Check(sk, \rho_C)$检验返回态ρ_C是否有效。任意由授权方产生的态都是有效态，能通过检测。文献[1]指出在密码学假设下，存在一类无法学习函数类，对于这类函数不存在 SSL，侧面验证了对任意无法学习类函数量子复制保护不存在的结论。另外，针对可搜索线路类$\mathcal{C}$，在量子安全隐输入混淆器、量子安全子空间混淆器以及次指数量子算法攻击下安全带错学习存在的假设下，文献[3]构建了用户可永久保留量子态的无限期安全的 SSL 协议。

不经意传输（Oblivious Transfer，OT）是密码学的一类重要协议。Bennett 等人基于理想比特承诺协议和量子通信设计了 OT 协议，但是有学者指出，量子资源下不存在无条件安全的比特承诺和无条件安全的 OT。随后有学者设计了基于单向函数（One-Way Functions，OWF）的比特承诺协议，那么是否存在基于 OWF 和量子通信的 OT 就成了专家们关注的问题。Impagliazzo 提出密码学的 5 个世界理论，其中 Minicrypt 世界里存在后量子（Post-Quantum）安全 OWF、量子计算和量子通信。Minicrypt 具有鲁棒性且高效性。鲁棒性在于可从多种困难问题中提取出大量能抵抗量子攻击的 OWF，如对于无结构的 AES 和 SHA，不存在任何次指数攻击。Minicrypt 的高效性在于操作可组合且能快速实现，鉴于 OWF 可抵抗基本的暴力搜索攻击，所以密钥长度相对较短。显然，如果没有量子能力，只利用 OWF 不能构建 OT 协议。量子通信的发展使得利用 Minicrypt 的鲁棒性和高效性构建 OT 和安全计算成为可能。文献[7]在存在后量子 OWF 和量子通信的假设下，提出公共随机串模型下的常数轮 OT 协议，此协议在恶意量子多项式时间敌手攻击下是通用可组合模拟安全的。结合 OT 协议和之前的工作，文献[7]又在 Minicrypt 世界里构建了安全两方和多方计算协议。

数字货币的密码协议应满足两个条件：①不可信用户的验证，任何不可信的用户，甚至是企图伪造钞票的敌手，都可以区分真钞票和伪钞票；②无法伪造，只有可信的造币厂才能生产真钞票。一个经典的比特串可以很容易地被复制，因此不满足“无法伪造”的条件。然而，任意的量子比特串不能被复制，所以量子信息具备“无法伪造”的特性。因此，利用量子状态制造“无法伪造”的货币（记为公钥量子货币）成为一个研究话题。在欧密会 2019 上，Zhandry 定义了一个新的密码对象，称为量子闪电（Quantum Lighting），它提供了一种强大的公钥量子货币形式，使得造币厂都无法生产两张相同的钞票。Zhandry 还提出了一个具体的量子闪电的方案，但该方案是否安全尚不清楚。相反，Zhandry 提出了一个貌似合理的计算困难性假设，并证明如果该假设是真的，那么这个量子闪电方案就是安全的。

在欧密会 2021 上，Bhaskar Roberts 证明 Zhandry 的假设是错误的[8]，从而表明 Zhandry 的量子闪电方案的安全性证明是不成立的。但是 Bhaskar Roberts 并不能证明 Zhandry 的方案是不安全的，并且 Zhandry 提出的计算困难性假设有可能是可以被修复的。Bhaskar Roberts 的工作是 Zhandry 量子货币方案安全性研究上第一次的尝试，具有重要意义。

随机谕言模型（ROM）是密码学中广泛使用的启发式方法，其中哈希函数被建模为随机

函数，即只能作为谕言机访问。该 ROM 用于构建实用的密码方案，包括数字签名、选择密文攻击（CCA）安全的公钥加密、基于身份的加密（IBE）等。在亚密会 2011 上，Boneh 等人提出了量子随机谕言模型（QROM），该模型中敌手可以用量子叠加态访问量子实现的谕言机。Boneh 等人观察到经典的 ROM 中的许多证明技术不能直接转化到 QROM 模型中。因此，QROM 模型中需要新的证明技术来证明随机谕言模型系统的后量子安全性。幸运的是，最近证明技术的进步已经表明，大多数 ROM 模型中被证明安全的密码结构在 QROM 模型中仍然是安全的。鉴于这种情况，人们很自然地要问是否存在一个一般理论将任何经典 ROM 模型中的证明提升为 QROM 模型中的证明，记为提升理论（Lifting Theorem）。

这篇发表在欧密会 2021 上的论文即围绕该问题展开[9]。该论文给出了一些密码方案，能够将 ROM 模型和 QROM 模型分开，这表明完全通用的将 ROM 中的证明提升为 QROM 的证明的一般理论是不可能存在的。为了展示 ROM 和 QROM 的区别，该论文首先引入一个原语，称为随机谕言机的量子访问证明（PoQRO）。粗略地说，PoQRO 是一种协议，在该协议中，量子证明者向经典验证者证明他有能力对随机谕言机进行量子访问，而经典验证者只能获得对随机谕言的经典访问权限。这与量子性证明的概念密切相关，不同之处在于，量子性证明仅需要针对完全经典对手的合理性，而 PoQRO 需要针对具有经典访问的量子对手的合理性。该论文在容错学习（LWE）问题的量子困难性假设下获得了一个 PoQRO。该构造是非交互的，即在验证者生成一对公钥和私钥并发布公钥后，证明者无须任何交互即可生成证明。但是，该证明不可公开验证，因为验证依赖于私钥。该论文还研究了可公开验证的 PoQRO，并相对于经典谕言机（可以以叠加态访问）构建了一个可公开验证的 PoQRO。在此基础上，该论文构造了在 ROM 中安全但在 QROM 中不安全的数字签名和公钥加密方案。特别是该论文得到了在标准密码假设下分离 ROM 和 QROM 的第一个合理的密码方案示例。

另外，对于某些类型的密码方案和安全概念，该论文给出了从 ROM 模型证明提升到 QROM 模型证明的通用提升理论。例如，该提升定理适用于 Fiat-Shamir 非交互式参数、Fiat-Shamir 签名和全域哈希签名等。该论文还讨论了提升定理在量子查询复杂性中的应用。

Chung Kai-Min 等人在 2021 年欧密会上[10]重新检查了压缩谕言机技术（Compressed-Oracle Technique），该技术由 Zhandry 引入，用于分析量子随机谕言模型（QROM）中的量子算法。事实证明，这种技术对于验证已知下限结果非常强大，而且能够有效证明以前似乎不可能获得的结果。尽管该技术非常有用，但实际使用压缩谕言机技术是很麻烦琐碎的。

首先，该论文对压缩谕言机技术进行了简洁且在数学上严谨的描述，采用了比学术界中的其他描述更抽象的视角，这样能够将注意力集中在具体相关方面。该描述很容易扩展到并行查询 QROM，在每个查询轮次中，所考虑的量子谕言机算法可以并行地对 QROM 进行多个查询。QROM 的这种变体允许对量子谕言机算法进行更细粒度的查询复杂度分析。该论文主要技术贡献是一个框架，其简化了压缩谕言机技术（并行查询的一般化）的使用以证明查

询复杂度。在该框架可适用时就可以通过纯粹的经典推理来证明量子查询复杂度的下界。不仅如此，结果表明对于典型的例子，对经典下界的仔细观察足以得出相应的量子下界。在几个例子中也证明了这一点，不但可以重新获得已知结果（如并行 Grover 的最优复杂度），而且获得了新的结果（如并行 BHT 碰撞搜索的最优复杂度）。该技术的主要应用是证明找到 q-chain 的难度，即在少于 q 次并行查询下，找到一个序列 $X_0,X_1,\cdots,X_q$,对于 $1\leqslant i\leqslant q$，有 $X_i=H(X_{i-1})$。

在连续工作证明（Proofs of Sequential Work）的背景下，上述产生 q 链的问题具有基础的重要性。事实上，作为一个具体的密码应用，文中证明了 Cohen 和 Pietrzak 提出的“简单的连续工作证明”对于量子攻击仍然是安全的。该证明不仅仅是插入新的界的问题；整个协议都需要根据量子攻击进行分析，并且需要大量的额外工作。然而，归功于文中的框架，该证明现在可以基于纯粹的经典推理来完成。

知识的证明协议是现代密码学理论研究中的一类基础性原语。在量子密码学中，通常研究量子知识的量子证明协议，其中证明者和验证者都是量子的，他们之间可执行量子信息交互。在文献[11]中，Vidick T 等人研究了一种新型的知识证明协议——量子知识的经典证明，其中，证明者是量子的但验证者是经典的，此时证明者的证据通常是一个量子态，他想让经典验证者相信其拥有或知道该量子态。该文献中给出了这种协议的两条性质：①对于某个量子态，若存在一个非破坏性的经典证明（交互式证明过程中不会对量子态证据造成损坏），则存在一个不限计算能力的攻击者能够克隆该量子态；②当参数满足一定条件时，一个难克隆的量子态的知识证明协议可用作量子货币验证协议。他们还提出两个协议实例，其设计都类似于量子货币方案的经典验证协议。Vidick T 等人还研究了 Mahadev 在 FOCS’2018 会议上所提出的量子计算的经典验证协议，并证明该协议可视为一个关于 QMA（QMA 是 NP 复杂性类的量子版）的量子知识的经典证明协议，其中证明者只具有多项式时间量子计算能力。

本节作者：罗宜元（惠州学院）、宋婷婷（暨南大学）、董晓阳（清华大学）、邹剑、邹宏楷（福州大学）、江浩东（信息工程大学）

参考文献

[1] JIANG H, ZHANG Z, MA Z. On the Non-tightness of Measurement-Based Reductions for Key Encapsulation Mechanism in the Quantum Random Oracle Model [C]. ASIACRYPT 2021 (1): 487-517.

[2] ZHANDRY M. Redeeming Reset Indifferentiability and Applications to Post-quantum Security [C]. ASIACRYPT 2021 (1): 518-548.

[3] ROBERT B, ZHANDRY M. Franchised Quantum Money [C]. ASIACRYPT 2021:549-574.

[4] HIROKA T, MORIMAE T, NISHIMAKI R, et al. Quantum Encryption with Certified Deletion, Revisited: Public Key, Attribute-Based, and Classical Communication [C]. In: Tibouchi M., Wang H. (eds) Advances in Cryptology-ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13090. Springer, Cham.

[5] GRILO A B, HÖVELMANNS K, HÜLSING A, et al. Tight Adaptive Reprogramming in the QROM [C]. ASIACRYPT 2021 (1): 637-667.

[6] ANANTH P, LA PLACA R L. Secure Software Leasing [C]. EUROCRYPT 2021: 501-530.

[7] GRILO A B, LIN H, SONG F, et al. Oblivious Transfer Is in MiniQCrypt [C]. EUROCRYPT 2021: 531-561.

[8] ROBERTS B. Security Analysis of Quantum Lightning [C]. EUROCRYPT 2021 (2): 562-567.

[9] YAMAKAWA T, ZHANDRY M. Classical vs Quantum Random Oracles [C]. EUROCRYPT 2021 (2): 568-597.

[10] CHUNG K-M, FEHR S, HUANG Y-H, et al. On the Compressed-Oracle Technique, and Post-Quantum Security of Proofs of Sequential Work [C]. EUROCRYPT 2021 (2): 598-629.

[11] VIDICK T, ZHANG T. Classical Proofs of Quantum Knowledge [C]. In: Canteaut A., Standaert FX. (eds) Advances in Cryptology - EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12697. Springer, Cham.

2 格理论与格密码

格中最短非零向量问题（SVP）是一个著名的计算困难问题，长期受到密码学家和数学家的关注，人们一直在探索这个问题的更好的解决算法。δ-SVP 问题是 SVP 问题的逼近版本：允许找到的非零向量长度最多为 $\delta \cdot \lambda_1(L)$，其中 $\lambda_1(L)$ 为最短非零向量的长度。δ-SVP 问题是一些密码算法和协议的安全基础。δ-HSVP 问题（δ-Hermite SVP）与 δ-SVP 问题类似，要求找到一个长度不超过 $\delta \cdot \det(L)^{1/n}$ 的非零向量。在文献[1]中，设计了一个时间复杂度为 $2^{n/2+O(n)}$ 的算法，当 $\delta \leqslant \tilde{O}(\sqrt{n})$ 时，可以解决 δ-SVP 问题和 δ-HSVP 问题。此外，还设计了一个时间-逼近程度权衡（Time-Approximation Tradeoff）算法，具体为时间复杂度为 $2^{k/2+O(k)} \cdot \mathrm{poly}(n)$ 的算法解决 δ-HSVP 问题（$\delta \approx k^{n/(2k)}, k \leqslant 0.99n$）和 δ-SVP 问题（$\delta \approx k^{(n/k)-0.62}, k \leqslant n/1.63$）。结合这两个结果，文献[1]给出了如下两个问题的目前最快的时间复杂度可以得到证明的算法，如 n^c -HSVP（其中 $c > 1/2$）、n^c -SVP（其中 $c > 1$ 或

$c \in (0.5, 0.802)$）。文献[1]的工作主要是建立在如下几个工作之上的：2015 年 Aggarwal、Dadush、Regev 和 Stephens-Davidowitz 关于格塔（Tower of Lattices）的工作；2013 年 Micciancio 和 Peikert 关于格上高斯分布的工作、2016 年 Dadush 提出的逆 Minkowski 猜想与 2017 年 Regev 和 Stephens-Davidowitz 对逆 Minkowski 猜想的证明。此外，2014 年 Becker、Gama 和 Joux 的工作也独立地使用了格塔的想法。

1982 年，A K Lenstra、H W Lenstra 和 Lovász 提出了 LLL 算法。从那时至今，LLL 算法对密码学产生了非常大的影响，是密码分析的一个基本工具。1994 年，Schnorr 和 Euchner 提出了 BKZ 算法。在 BKZ 算法中，通过引入分块参数 k 来控制算法时间与算法输出质量之间的折中，k 越大，输出的约化基的质量越好，但算法时间随着 k 至少成指数增加。在 δ-HSVP 问题中，要求算法找到的非零向量长度不超过 $\delta \cdot \det(L)^{1/n}$。人们用 $\delta^{1/(n-1)}$ 来刻画解决 δ-HSVP 问题的算法的输出质量，这个值也被称为 Hermite 平方根因子（Root Hermite Factor，RHF）。2020 年，Li 和 Nguyen 发现，如果我们把搜索半径放大到 $\alpha > 1$ 倍，那么使用极端圆柱剪枝技术（Extreme Cylinder Pruning）的枚举算法在渐进情况下可以得到指数级的加速。2020 年，Albrecht、Bai、Fouque 等人设计了一个 BKZ 变形算法，可以在时间复杂度 $k^{k/8+O(k)}$ 内达到的 RHF 为 $GH(k)^{1/(k-1)}$，其中 $GH(k)^{1/(k-1)}$ 是 n 维体积为 1 的球的半径。与 1983 年 Kannan 的算法相比（2007 年 Hanrot 和 Stehlé 对该算法给出了更好的分析），Albrecht、Bai、Fouque 等人的时间复杂度具有超指数级的加速。在文献[2]中，主要综合运用以上两个工作的新想法，Martin R Albrecht、Shi Bai、Jianwei Li 等人提出了一个新的 BKZ 变形算法，包含两个可选参数 (α, c)。由于新自由度的引入，这个 BKZ 变形算法可以在时间复杂度 $2^{\frac{k \log k}{8} - 0.654k + 25.84}$ 内达到的 RHF 为 $GH(k)^{1/(k-1)}$，这是目前最好的结果。

在对 DSA 算法、ECDSA 算法、Diffie-Hellman 问题做侧信道攻击时，需要解决的最终问题可以抽象为 HNP 问题（Hidden Number Problem，隐藏数问题）。在 HNP 问题中，可以获得一个秘密整数的随机倍数模另一个公开整数的余数的某些最高位比特，目标是恢复这个秘密整数。HNP 问题可以转化为与格有关的 BDD 问题（Bounded Distance Decoding），然后使用格基约化算法来解决。在 BDD 问题中，给定一组格基和一个目标点 t，要求找到距离 t 最近的唯一的格中向量。如果直接使用格基约化算法来解决 BDD 问题，唯一性的要求导致在某些参数情况下 BDD 问题是不可解的，这个现象被称为格障碍（Lattice Barrier）。例如，当采样较少时，存在太多的格点与目标点 t 之间的距离比真正要找到的那个点更近。造成格障碍的原因是，将原始的密码分析问题抽象为 HNP 问题时，丢掉了一些有用的信息：原始的密码分析问题其实是可解的。在文献[3]中，为了更好地刻画我们面临的密码分析问题，Martin R

Albrecht、Nadia Heninger 定义了带谓词（Predicate）的 BDD 问题（$BDD_{\alpha,f(\cdot)}$），并设计了两个解决该问题的算法。具体如下：除了定义 $BDD_{\alpha,f(\cdot)}$ 问题，他们还定义了带谓词的唯一最短向量问题 $uSVP_{f(\cdot)}$，并将利用 Kannan 的嵌入方法，将 $BDD_{\alpha,f(\cdot)}$ 问题约化为 $uSVP_{f(\cdot)}$ 问题。进一步，基于枚举算法和筛法，设计了两个解决 $uSVP_{f(\cdot)}$ 问题的算法。与 2020 年 Dachman-Soled、Ducas、Gong 和 Rossi 设计的解决 LWE 问题的算法相比，其算法也有缺点，那就是适用场景有局限，不适用于任意的格基约化算法。

寻找有效的算法来解决 SVP 问题，在格理论和格密码领域中，一直堪称最重要的问题。从 1980 年以来，人们不断地探寻更好的算法来解决 SVP 问题，包括量子算法。到目前为止，这些算法主要可以分为枚举算法（Enumeration）和筛法（Sieving）两大类。其中，枚举算法速度较慢但需要的存储空间很少，而筛法速度更快但需要用到的存储空间很大。实际上，在 SVP 问题的挑战赛中，前 10 个最好的纪录都是用筛法创造的。2008 年，Nguyen 和 Vidick 设计了一个启发性的筛法，时间复杂度为 $2^{0.415d+O(d)}$，空间复杂度为 $2^{0.2075d}$，其中 d 是格的维数。2010 年，Micciancio 和 Voulgaris 又改进了这个算法，虽然渐进复杂度仍然一样，但是实际执行效率更高。此后，人们不断设计出更好的算法。2016 年，使用局部敏感滤波技术（Locality-Sensitive Filtering），Becker、Ducas、Gama 和 Laarhoven 给出了一个时间复杂度为 $2^{0.292d+O(d)}$ 的算法，这也是到目前为止最好的经典算法。对于量子算法，2016 年，Laarhoven 给出了一个时间复杂度为 $2^{0.265d+O(d)}$ 的量子算法，这可能是文献[4]之前最好的量子算法。2020 年，基于已有的量子算法，Albrecht、Gheorghiu、Postlethwaite 等人专门研究了格筛法的量子加速问题。在文献[4]中，利用量子随机游走算法，André Chailloux、Johanna Loyer 改进了解决 SVP 问题的量子筛法，得到了目前最好的结果。文献[4]中新算法的具体参数为：时间复杂度为 $2^{0.2570d+O(d)}$、QRAM 规模为 $2^{0.0767d}$、量子存储空间为 $2^{0.0495d}$，以及经典存储空间为 $\mathrm{poly}(d)\cdot 2^{0.2075d}$。此外，在量子存储空间给定和 QRAM 给定的两种情况下，André Chailloux、Johanna Loyer 还分别设计了时间-空间折中（Time-Memory Tradeoff）的算法。

公钥密码算法除提供加密和签名两个功能之外，还可以提供其他功能，如群签名、基于身份的加密、密钥的知识证明（Proof of Knowledge）。在文献[5]中，Kelong Cong 等人考虑了 CCA 安全混合公钥加密算法的分布式解密问题。在 KEM-DEM 框架（Key Encapsulation Mechanism，KEM；Data Encryption Mechanism，DEM）下的标准混合公钥加密算法很难实现分布式的解密，因为这样会破坏 CCA 安全性。为了解决这个问题，在文献[5]中，Kelong Cong 等人提出了两个一般性的变换 $Hybrid_1$ 和 $Hybrid_2$，不仅可以支持分布式解密，还可以保持

CCA 安全。在这两个变换中，$Hybrid_1$ 在 ROM 模型下是安全的，而 $Hybrid_2$ 在 QROM 模型下也是安全的。这两个变换与以下几个已知的构造方法关系极为密切：2001 年 Okamoto 和 Pointcheval 提出的 REACT 变换；2008 年 Abe、Gennaro 和 Kurosawa 提出的 Tag-KEM 框架；2013 年 Fujisaki 和 Okamoto 提出的 Fujisaki-Okamoto 变换的第二种混合变形。此外，在文献[5]中，基于 LWR 问题，几位作者设计了一个新的确定性后量子公钥加密算法（该算法的真正安全基础是作者新定义的 LVP 问题），并进一步设计了分布式的密钥生成算法和解密算法。这个新设计被命名为 Gladius。Gladius 与 NIST 第 3 轮候选算法 Saber 类似。

迄今为止，人们还没有找到有效的算法来计算超奇异椭圆曲线之间的同源，包括有效的量子算法。于是，超奇异椭圆曲线之间的同源这个计算困难问题成为构造后量子密码算法的另一种可能选择。在这个领域中，人们已经取得了不少进展，构造了哈希函数、公钥加密算法、密钥封装机制（KEM）、认证协议和签名算法。特别地，在 NIST 的后量子密码算法标准化活动中，进入第 3 轮的 SIKE 正是基于同源问题构造的。其中，SIKE.PKE 是一个 CPA 安全的公钥加密算法，而 SIKE.KEM 是一个 CCA 安全的 KEM。2011 年，受到如下两个工作的启发：Couveignes、Rostovtsev、Stolbunov 关于同源问题与密码算法设计的工作，以及 Charles、Goren 和 Lauter 关于扩张图（Expander Graph）与哈希函数构造的工作。Jao 和 De Feo 设计了一个基于超奇异椭圆曲线同源问题的密钥协商协议，称为 SIDH。2020 年，利用扭点（Torsion Point）给出的额外信息，de Quehen、Kutas、Leonardi 等人给出了针对 SIDH 协议的攻击算法，在某些参数情况下他们的攻击是有效的。同时，他们给出了一个算法来构造后门曲线（Backdoor Curve），使得在后门信息的帮助下，可以快速计算同源。在文献[6]中，Luca De Feo 等人利用这个构造后门曲线的算法，设计了一个 CPA 安全的公钥加密算法。他们再利用通用的 OAEP 变换，将这个公钥加密算法进一步地改进为 CCA 安全的。

对于格中最短非零向量问题，由于筛法比枚举算法渐进更好，近几年人们对筛法格外重视，无论是从理论的角度还是从实际效率的角度。2015 年，受到 May 和 Ozerov 搜索最近邻（Nearest Neighbor）码字算法的启发，使用次二次最近邻搜索（Sub-Quadratic Nearest Neighbor Search）技术，Becker、Gama 和 Joux 在不增加空间复杂度的情况下，提升了筛法的速度。对于他们的算法，时间复杂度为 $2^{0.3112n+O(n)}$，空间复杂度为 $2^{0.2075n+O(n)}$。2016 年，为了解决逼近最近邻搜索（Approximate Nearest Neighbor Search）问题，Becker、Ducas 和 Gama 等人提出了一种使用局部敏感滤波器（Locality-Sensitive Filter）的新方法，将时间复杂度改进为 $2^{0.292n+O(n)}$。2017 年，Herold、Kirshanova 设计了一个时间复杂度为 $2^{0.3717n+O(n)}$ 和空间复杂度为 $2^{0.1887n+O(n)}$ 的算法。2019 年，Albrecht、Ducas、Herold 等人设计了一般性的筛法核（General Sieve Kernel），称为 G6K，适用于各种格筛法。他们还创造了 SVP 的纪录：解决了 155 维的 TU Darmstadt SVP 挑战。在文献[7]中，Léo Ducas 等人专注于研究怎么使用带张量核（Tensor Core）的 GPU 来加速筛法。他们将 G6K 应用到上面提到的 3 个筛法中去，并利用新设计的

对偶哈希技术，不但获得了新的 SVP 纪录，而且与纯 CPU 攻击相比，更省能量。例如，与以前 155 维 SVP 纪录相比，文献[7]利用更少的时间和能量创下了 176 维 SVP 的新纪录。

NTRU 算法最早由 Hoffstein、Pipher 和 Silverman 于 1996 年提出，既是一个轻量级的算法，又是一个后量子算法，并且进入了 NIST 后量子算法标准化竞赛的第 3 轮。自从提出以来，人们一直对 NTRU 算法非常感兴趣，做了各种可能的密码分析，同时提出了更加安全的变形。2011 年，通过对 NTRU 加密算法做适当的修改，Stehlé 和 Steinfeld 将变形后的 NTRU 加密算法的安全基础建立在了理想格（Ideal Lattice）的最难情况（Worst-Case）的困难问题之上，而这类理想格与分圆域有关。由于 NTRU 格具有较丰富的代数结构，因此 NTRU 问题的困难性比单纯地寻找格中最短非零向量要低一些。当模数 q 足够大时，如 q 是 n 的超多项式时，往往可以利用 NTRU 格的代数结构来改进单纯的格基约化方法。这种情况被称为是过度伸展的（Overstretched）。2016 年，Albrecht、Bai 和 Ducas 发现，对于过度伸展的 q，子域攻击（Subfield Attack）比其他攻击方法的渐进效果更好。同年，Cheon、Jeong 和 Lee 设计了一种类似的子域攻击方法。2017 年，Kirchner 和 Fouque 发现，上述攻击方法真正有效的原因是 NTRU 格包含一个不寻常的稠密大维数子格。通过进一步的理论分析，他们发现，对于随机的三元秘密多项式，$q = n^{2.783+O(1)}$ 就已经是过度伸展的。在文献[8]中，通过理论分析和实验观察 SKR_κ 和 DSD_κ 两个事件，Léo Ducas 和 van Woerden 发现，当 n>100 时，$q \approx 0.004 \cdot n^{2.484}$ 才是更好的估计值。例如，实验显示，当 n=127、q=700 时就会显示出过度伸展的特征，而按照 Kirchner 和 Fouque 的估计，q=700000 时才会显示出过度伸展的特征。

自从 2005 年 LWE 问题被 Regev 提出以来，在密码学中已经得到了很多应用，如全同态加密、后量子加密算法、后量子密钥封装（Key Encapsulation Mechanism，KEM）、后量子签名算法。特别是在 NIST 后量子密码算法标准竞赛活动中，有一部分算法正是基于 LWE 问题设计的，包括进入第 3 轮的 CRYSTALS-Kyber 和 CRYSTALS-Dilithium。另外，人们也一直在寻找更好的密码分析方法来解决 LWE 问题。第一类解决 LWE 问题的方法是代数或者组合的方法，这类方法需要大量的具体实例才行。BKW 型算法是这类算法的典型代表，目前还在不断改进中。第二类算法是基于对偶格的密码分析方法。这类方法首先需要构造一个区分器，然后猜测密钥的一部分，通过区分器来判断猜测是否正确。2015 年，Duc、Tramèr 和 Vaudenay 使用多维傅里叶变换，改进了 Albrecht、Cid、Faugère 等人 2013 年提出的算法。2017 年，Albrecht 首次考虑了在对偶格攻击中生成很多短向量的问题。在文献[9]中，Guo 等人提出了新的对偶格攻击方法，得到了关于 LWE 问题的最好密码分析结果。此外，他们还将这个攻击方法应用到 CRYSTALS-Kyber 和 CRYSTALS-Dilithium 两个算法，得到了目前最好的密码分析结果。Guo 等人的主要工作如下：首先巧妙利用模 q 的特点，设计了一个区分器可以用于猜测密钥的最低位比特，并且用 FFT 变换，进一步提升了这个区分器的效率；其次利用 BKZ 格基约化算法和这个新的区分器，设计了一个更好的针对 LWE 问题的对偶格分析方法；

最后为了估计新算法在经典的 RAM（Random Access Machine）模型下的攻击效果，结合 BKZ 算法和筛法，Guo 和 Johansson 还设计了一个两步的格基约化策略。

对于任意多项式时间可以计算的程序，不可区分混淆（indistinguishability Obfuscation，iO）可以在保持该程序功能的前提下把它变得混乱，使得任何概率多项式时间的攻击者都无法将其与同样长度的真随机程序区分开来。iO 是一种新的密码学本原（Cryptographic Primitive），在密码学中有很多应用。目前，iO 的构造方法主要有两种：一种是利用双线性配对（Bilinear Pairing）或多线性配对（Multilinear Pairing）的；另一种是没有利用配对（Pairing）的。2020 年，使用配对技术，基于以下 4 个经过长期研究的密码学假设：F_p 上的 LWE 问题、F_p 上的 LPN 问题、双线性群上的 DLIN 假设和 NC^0 中的伪随机发生器，Jain、Lin 和 Sahai 给出了一种 iO 的构造方法。2020 年，基于分裂的全同态加密方案（Split FHE），Brakerski、Döttling、Garg 等人给出了一种新的 iO 构造方法。受到这个工作的启发，Gay 和 Pass 提出了一个新的循环安全（Circular Security）假设，并基于这个新的假设和 LWE 问题，给出了一种新的 iO 构造方法。基于健忘的 LWE 采样假设（Oblivious LWE Sampling），Wee 和 Wichs 也给出了一种新的构造方法。在文献[10]中，Hopkins 等人研究了上述两个构造中的密码学新假设。而后发现，这两个假设都是不成立的，并且给出了具体的反例。虽然反例目前还未能直接用于攻击这两个构造方法，但足以说明这两个构造的安全基础是不牢固的。

格基约化算法在密码学和数论中都有很重要的应用，从 1982 年 A K Lenstra、H W Lenstra 和 Lovâsz 提出 LLL 格基约化算法以来，人们就不断尝试改进格基约化算法的效率。1988 年，Schnorr 提出使用浮点运算代替有理数的精确运算，提高了 LLL 算法的效率。1994 年，Schnorr 和 Euchner 提出了 BKZ 算法。该算法的基本思想是，通过引入分块参数 k 来控制算法时间与算法输出质量之间的折中，k 越大，输出质量越好，但算法时间复杂度也越大。2009 年，Nguyen 和 Stehlé 设计了一个时间复杂度为 $O(d^5(d+B)B)$ 的 LLL 算法，其中 d 是格的维数，B 是向量元素的最大比特数。这个算法也被称为 L^2 算法。2011 年，Novocin、Stehlé 和 Villard 将 LLL 算法的时间复杂度改进到 $O(d^{5+\varepsilon}B+d^{\omega+1+\varepsilon}B^{1+\varepsilon})$，并将改进后的算法称为 $\tilde{L}^1$。2016 年，Neumaier 和 Stehlé 获得了更好的时间复杂度 $O(d^4B^{1+O(1)})$。在文献[11]中，基于 QR 分解、Block-Cholesky 分解和浮点运算，Paul Kirchner 等人得到了更好的时间复杂度 $\tilde{O}(d^\omega C)$。其中，对于密码学应用，如 Coppersmith 攻击 RSA 使用的格、背包类型的格，C 与 B 接近。文献[11]中得到的算法同样是 LLL 类型的。

在后量子密码学领域，基于编码和基于格的构造方法是相对让人更加信服的两种构造方

法。基于编码的构造方法，与 RSA 算法几乎有同样长的历史，最早由 McEliece 于 1978 年提出，经过人们长期的密码分析仍然是安全的。在数论领域，人类研究格的历史非常长。在密码学领域，使用格理论，除了后量子公钥加密、密钥封装和数字签名算法，人们还构造了全同态加密等算法。对于一个具体的密码算法，找到一个新的攻击方法，我们只是得到了新的时间复杂度上界。为了更好地评估上述两类后量子密码算法的安全性，我们更需要知道的是所有潜在攻击方法的时间复杂度下界。这个问题虽然非常难，但针对某一类具体的算法，我们有可能得到好的结果。使用最近邻技术（Nearest Neighbor Technique），最近几年，人们不断得到新的密码分析结果。在文献[12]中，Elena Kirshanova 等人考虑了最近邻技术的时间复杂度下界，从理论上得到了很好的结果。他们证明了，在使用最近邻技术的格筛法中，2016 年 Becker、Ducas、Gama 和 Laarhoven 的结果是最优的，他们的时间复杂度 $2^{0.292d+O(d)}$ 是最好的可能结果。此外，Kirshanova 和 Laarhoven 证明了，2016 年 Laarhoven 的量子筛法和 2018 年 Herold、Kirshanova 和 Laarhoven 的向量元组筛法（Tuple Sieving）也都是最优的。对于编码的情况，Kirshanova 和 Laarhoven 得到了一个下界，与 2015 年 May 和 Ozerov 的结果几乎匹配。这个结果在一定程度上说明，我们要想得到比 May 和 Ozerov 明显好的译码攻击，需要寻找其他途径。

群签名可以很好地保护签名者的隐私信息，可以用于设计更加复杂的隐私保护方案，因此人们对群签名方案的设计一直很感兴趣。面对高速发展的量子计算技术，设计可以抵抗量子计算攻击的高效群签名方案，是一个紧迫的问题。使用格来构造密码算法和协议，是后量子密码学领域中让人们极其关注的一条途径。2010 年，使用 LWE 问题，Gordon、Katz 和 Vaikuntanathan 构造了第一个基于格的群签名方案。2016 年，Libert、Ling、Nguyen 等人设计了一种基于格的零知识证明技术，并使用该技术构造了群签名方案。这两个方案有一个共同的缺点：签名规模太大，达 500MB 左右。2018 年，del Pino、Lyubashevsky 和 Seiler 设计了一个零知识证明协议来证明承诺的值属于某个特定的小集合。他们又进一步地给出了一个群签名方案，签名长度为 581KB。这是文献[13]之前的签名规模最小的群签名方案。基于这个构造，在文献[13]中，利用 2018 年提出的 BDLOP 承诺方案，Vadim Lyubashevsky 等人设计了一个更加高效的基于格的零知识证明协议，并在此基础上给出了一个迄今为止签名长度最短的群签名方案。在该零知识证明协议中，最核心的方程是 $[A \mid B + mG]\boldsymbol{s} = \boldsymbol{u}$，其中用户的身份是 m，用户的私钥是短向量 $\boldsymbol{s}$。这个方程有个特别之处，等式左边的矩阵并不是完全公开的，需要使用新的密码学技巧来处理。通过巧妙应用 BDLOP 承诺方案，Vadim Lyubashevsky 等解决了这个问题。进一步，通过 Fiat-Shamir 变换，将待签名的文件作为哈希函数的部分输入，进而可以将这个零知识证明协议转化为群签名方案。此外，这个方案还有一个优点：在协议最初的承诺方案里，用户身份 m 已经被封装到了承诺值里（也可以看作是对 m 的加密），群管理者通过简单的解密算法就可以恢复出用户身份 m。

给定随机矩阵 $\boldsymbol{A} \in Z_q^{m \times n}$，以及随机 $\boldsymbol{s}, \boldsymbol{e} \in \{-1,0,1\}^n$，计算 $\boldsymbol{b} \in Z_q^m$ 满足 $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{e} \bmod q$。

三元 LWE 问题（Ternary LWE）的含义是：已知 $(\boldsymbol{A}, \boldsymbol{b})$，求出 $\boldsymbol{s}$。在标准的基于格的加密和签名方案中，m=n。在 NTRU 算法的原始论文中，Hoffstein、Pipher 和 Silverman 提到了一种由 Odlyzko 提出的中间相遇攻击。由于使用这个攻击方法，因此可以计算出 $\boldsymbol{s}$，时间复杂度为 $O(3^{n/2})$。2003 年，Howgrave-Graham、Silverman 和 Whyte 提到了另一种中间人攻击，在文献[14]中，Alexander May 将其称为 Howgrave-Graham 中间人攻击。虽然 Howgrave-Graham 攻击的时间复杂度比 Odlyzko 要差，但是在文献[14]中 Alexander May 的新攻击算法的基础。此外，Alexander May 在提出这个新算法之前，也受到了如下两个工作的启发：2011 年 Becker、Coron 和 Joux 关于子集和问题的工作，以及 2012 年 Becker、Joux、May 和 Meurer 关于随机线性码译码的工作。令 $T^n = Z_q^n \cap \{-1,0,1\}^n$。对任意 $0 < w < n$，假设 $\boldsymbol{s}$ 是一个 Hamming 重量为 $\boldsymbol{W}$ 的三元向量。在文献[14]中，May 将原始的三元 LWE 问题转化为如下问题：$\boldsymbol{A}\boldsymbol{s}_1 + \boldsymbol{e}_1 = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{s}_2 + \boldsymbol{e}_2$，其中 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$ 的 Hamming 重量为 $w/2$，$\boldsymbol{e}_1 \in T^{n/2} \times 0^{n/2}$，$\boldsymbol{e}_2 \in 0^{n/2} \times T^{n/2}$。在此基础上，Alexander May 设计了一个新的中间人攻击算法，并借鉴子集和问题中的表示技术（Representation Technique），设计了三种不同的表示技术，复杂程度依次递增。通过分析发现，这个新算法的时间复杂度可以低到 $O(3^{0.25n})$，这比最好的量子攻击算法的时间复杂度 $O(3^{n/3})$ 还要低一些。此外，对于 NIST 后量子算法标准竞赛第 3 轮的两个算法 NTRU 和 NTRU-Prime，May 攻击算法的时间复杂度为 $O(3^{0.3n})$。

与群签名类似，群加密也可以很好地保护加密者的隐私信息，并且可以用于设计更加复杂的隐私保护方案。面对量子计算的巨大危险，设计可以抵抗量子计算攻击的高效群加密方案，也是一个非常有意义的紧迫问题。2007 年，Kiayias、Tsiounis 和 Yung 首次提出了群加密的概念，并且给出了一个基于数论的构造方案。这个方案的局限是，只能支持部分动态群（Partially Dynamic Group），用户可以动态地加入，但不能退出。构造能够支持完全动态群（Fully Dynamic Group）的群加密方案，是一个无法回避的问题，也是一个极有意义的问题。2016 年，Libert、Ling、Mouhartem 等人构造了第一个基于格的群加密方案。借鉴 Bootle、Cerulli 和 Chaidos 等人于 2016 年构造的群签名方案，2021 年 Nguyen、Safavi-Naini、Susilo 等人给出了一种群加密方案的构造方法，可以支持完全动态群。但这个方案的缺点是，构造非常复杂，效率较低，而且安全证明是在 RO 模型下做出的。在文献[15]中，Jing Pan 等人给出了完全动态群加密（Fully Dynamic Group Encryption）的严格定义和安全要求。然后，基于

Yang、Au、Zhang 等人于 2019 年给出的零知识证明协议，他们设计了一个新的零知识证明协议来证明格上的复杂关系。最后，Jing Pan 等人构造了一个完全动态群加密方案，并在标准模型下给出了安全证明。

与 LWE 问题类似，Ring-LWE 问题也可以用于构造一系列密码学性质很好的算法和协议，而且效率比基于 LWE 问题的构造更高。另外，由于 Ring-LWE 问题具有更丰富的代数结构，它的计算困难性很有可能低于 LWE 问题。2010 年，Lyubashevsky、Peikert 和 Regev 提出了 Ring-LWE 问题，并给出了一个多项式时间的量子归约，将理想格中的最坏情况下的困难性假设（Worst-Case Assumptions on Ideal Lattices）归约到 Ring-LWE 问题。2017 年，Peikert、Regev 和 Stephens-Davidowitz 进一步将上述的归约推广到了任意模和任意环的情况。由于理想格与 Ring-LWE 问题存在着紧密的联系，因此深入理解理想格中的 SVP 问题，无疑对我们理解 Ring-LWE 问题非常有帮助。在这个领域，人们已经得到了一系列的结果，设计了很多算法，包括量子算法。在文献[16]中，Yanbin Pan 等人发现，当 $N=2^n$ 时，如果素理想 P 中包含一个素数 $p \equiv \pm 3 (\mathrm{mod}\, 8)$，那么 P 中的 SVP 问题存在多项式时间的解法。进一步，他们还将该结论推广到了素理想乘积的情况。基于文献[16]中的方法和结论，即使是 $N=2^n$ 的特殊情况，研究者也并不能在多项式时间内解决 Ring-LWE 问题，因为前面提到的归约是基于理想格中的最坏情况下的困难性假设，而且归约是单向的。

人们研究 NTRU 问题，已经有 26 年的历史。虽然 NTRU 和 NTRU-Prime 进入了 NIST 后量子算法标准竞赛的第 3 轮，但是相对其他格问题的困难性，人们在计算复杂度的层面上对 NTRU 问题的困难性还是缺乏深入的理解。在文献[17]之前，研究者仅知道如下两个归约结果：从判定 NTRU 问题到求解 NTRU 问题（Search NTRU）的归约，以及一个从判定 NTRU 问题到求解 Ring-LWE 问题（Search Ring-LWE）的归约。第二个归约仅给出了 NTRU 问题的困难性上界，而研究者更关心的是 NTRU 问题的困难性下界。2016 年，Peikert 提出了如下问题：对于 NTRU 类型问题的困难性，是否存在最坏情况下的约化（Worst-Case Reduction），或者从求解问题到判定问题的约化（Search-to-Decision Reduction）。在文献[17]中，基于非常扎实的代数数论功底，Alice Pellet-Mary 和 Damien Stehlé 回答了上述两个问题，得到了关于 NTRU 问题困难性的两个归约结果。Alice Pellet-Mary 和 Damien Stehlé 的工作如下：①设计一个从平均情况 id-HSVP 问题（Average-Case id-HSVP）到平均情况 NTRU 问题（Average-Case NTRU）的归约，再利用已有的从最坏情况 id-HSVP 问题（Worst-Case id-HSVP）到平均情况 id-HSVP 问题的归约（2020 年由 Boer、Ducas、Pellet-Mary 等人给出），我们就得到了从最坏情况 id-HSVP 问题到平均情况 NTRU 问题的归约；②从平均情况求解 NTRU 问题到判定 NTRU 问题的归约。

在 NIST 后量子算法标准竞赛中，有些 KEM 算法是基于格构造的，如进入第 3 轮的 5 个算法：CRYSTALS-KYBER、NTRU、SABER、FrodoKEM 和 NTRU-Prime。评估这些 KEM 算

法抵抗密钥不匹配攻击（Key Mismatch Attack）的能力是全面评估它们的安全性的重要方面。在密钥不匹配攻击中，一个参与者的公钥被重复使用，然后通过比较两个参与者之间的共享密钥是否相同，最终恢复出该参与者的私钥。2018 年，Ding、Fluhrer 和 Saraswathy 首次提出了一种密钥不匹配攻击方法，该方法可以用于攻击 2012 年 Ding、Xie 和 Lin 提出的基于格的密钥交换协议。2019 年，Bauer、Gilbert、Renault 等人提出了一种针对 NewHope 的密钥不匹配攻击方法。后来，Qin、Cheng 和 Ding 改进了这个攻击。接着，Okada、Wang 和 Takagi 又进一步降低了这个攻击的询问次数。2012 年，Zhang、Cheng 和 Ding 提出了一种针对 NTRU-HRSS 的密钥不匹配攻击，能以 93.6%的概率完全恢复出密钥。虽然人们已经提出了若干种密钥不匹配攻击方法，但对于 NIST 候选算法还是缺乏系统全面的评估。在文献[18]中，针对 NIST 竞赛第 2 轮和第 3 轮中所有基于格的 KEM 算法，Yue Qin 等人进行了系统全面的研究。他们从理论上给出了最优的询问次数，并发现对于 NewHope、FrodoKEM 和 SABER，实际攻击所需的询问次数与该理论值之间还有较大的差距。

1993 年，Blum、Furst、Kearns 等人首次研究了 LPN 问题在密码学中的应用。12 年后，Regev 提出了 LPN 问题的推广 LWE 问题，而且将 LWE 问题的困难性建立在了最坏情况的格困难问题（Worst-Case Hard Lattice Problem）之上。但是，对于 LPN 问题，人们的理解还很不深入，一直没有找到类似 LWE 问题的很好的计算困难性归约。2000 年，Blum、Kalai 和 Wasserman 提出了著名的 BKW 算法来解决 LPN 问题，时间和采样复杂度都为 $2^{O(n/\log n)}$。2005 年，通过引入采样放大技术（Sample Amplification），Lyubashevsky 改进了 BKW 算法，新算法的时间复杂度为 $2^{O(n/\log\log n)}$，采样复杂度为 $q=n^{1+\varepsilon}$。2018 年，Brakerski、Lyubashevsky、Vaikuntanathan 等人利用噪声率为 $\log_2 n/n$ 的 LPN 问题构造了抗碰撞的哈希函数。2019 年，Brakerski、Lyubashevsky、Vaikuntanathan 等人证明了如下结论：如果平衡码上的 NCP 问题（Nearest Codeword Problem）在噪声率为 $\log^2 n/n$ 时具有最坏情况下的困难性，那么噪声率为 $1/2-1/\mathrm{poly}(n)$ 的 LPN 问题具有拟多项式（Quasi-Polynomial）困难性。在文献[19]中，基于上述 2019 年的工作，Yu 等人得到了一些新的结果。其中，文献[19]的一个主要结果如下：假设噪声率为 $w/m=n^{-c}$ 的 NCP 问题是 $(T=2^{\Omega(n^{1-c})},m=2^{\Omega(n^{1-c})})$ 困难的（对于平衡码或对于独立码），我们有：①当 $0<c<1/2$ 时，常噪声率 LPN 问题是 $(T=2^{\Omega(n^{1-c})},\varepsilon=2^{-\Omega(n^c)},q=2^{\Omega(n^c)})$ 困难的；②当 $1/2\leqslant c<1$ 时，常噪声率 LPN 问题是 $(T=2^{\Omega(n^{1-c})},\varepsilon=2^{-\Omega(n^{1-c})},q=2^{\Omega(n^{1-c})})$ 困难的。

本节作者：胡红钢（中国科学技术大学）

参考文献

[1] AGGARWAL D, LI Z, STEPHENS-DAVIDOWITZ N. A 2n/2-Time Algorithm for δ-SVP and δ-Hermite SVP, and an Improved Time-Approximation Tradeoff for (H)SVP [C]. EUROCRYPT 2021 (1): 467-497.

[2] ALBRECHT M R, BAI S, LI J, et al. Lattice Reduction with Approximate Enumeration Oracles: Practical Algorithms and Concrete Performance [C]. CRYPTO 2021 (2): 732-759.

[3] ALBRECHT M R, NADIA H. On Bounded Distance Decoding with Predicate: Breaking the "Lattice Barrier" for the Hidden Number Problem [C]. EUROCRYPT 2021 (1): 528-558.

[4] CHAILLOUX A, LOYER J. Lattice Sieving via Quantum Random Walks [C]. ASIACRYPT 2021 (4): 63-91.

[5] CONG K, COZZO D, MARAM V, et al. Gladius: LWR Based Efficient Hybrid Public Key Encryption with Distributed Decryption[C]. ASIACRYPT 2021 (4): 125-155.

[6] DE FEO L, DE SAINT GUILHEM C Delpech, FOUOTSA Tako Boris, et al. Séta: Supersingular Encryption from Torsion Attacks [C]. ASIACRYPT 2021 (4): 249-278.

[7] DUCAS L, STEVENS M, VAN WOERDEN W P J. Advanced Lattice Sieving on GPUs, with Tensor Cores [C]. EUROCRYPT 2021 (2): 249-279.

[8] DUCAS L, VAN WOERDEN W P J. NTRU Fatigue: How Stretched is Overstretched[C]? ASIACRYPT 2021 (4): 3-32.

[9] GUO Q, JOHANSSON T. Faster Dual Lattice Attacks for Solving LWE with Applications to CRYSTALS [C]. ASIACRYPT 2021 (4): 33-62.

[10] HOPKINS S B, JAIN A, LIN H. Counterexamples to New Circular Security Assumptions Underlying iO [C]. CRYPTO 2021 (2): 673-700.

[11] KIRCHNER P, ESPITAU T, FOUQUE PA. Towards Faster Polynomial-Time Lattice Reduction [C]. CRYPTO 2021 (2): 760-790.

[12] KIRSHANOVA E, LAARHOVEN T. Lower Bounds on Lattice Sieving and Information Set Decoding [C]. CRYPTO 2021 (2): 791-820.

[13] LYUBASHEVSKY V, NGUYEN N K, PLANÇON M, et al. Shorter Lattice-Based Group Signatures via "Almost Free" Encryption and Other Optimizations [C]. ASIACRYPT 2021 (4): 218-248.

[14] MAY A. How to Meet Ternary LWE Keys [C]. CRYPTO (2) 2021: 701-731.

[15] PAN J, CHEN X, ZHANG F, et al. Lattice-Based Group Encryption with Full

Dynamicity and Message Filtering Policy [C]. ASIACRYPT 2021 (4): 156-186.

[16] PAN Y, XU J, WADLEIGH N, et al. On the Ideal Shortest Vector Problem over Random Rational Primes [C]. EUROCRYPT 2021 (1): 559-583.

[17] PELLET-MARY A, STEHLÉ D. On the Hardness of the NTRU Problem [C]. ASIACRYPT 2021(1): 3-35.

[18] QIN Y, CHENG C, ZHANG X, et al. A Systematic Approach and Analysis of Key Mismatch Attacks on Lattice-Based NIST Candidate KEMs [C]. ASIACRYPT 2021 (4): 92-121.

[19] YU Y, ZHANG J. Smoothing Out Binary Linear Codes and Worst-Case Sub-exponential Hardness for LPN [C]. CRYPTO 2021 (3): 473-501.

3　同源密码

Boneh、Kogan 和 Woo 在 Asiacrypt 2020 提出了一个用于构建不经意伪随机函数（OPRF）的框架，并通过基于 SIDH 的可验证方案和基于 CSIDH 的方案进行实例化。Andrea Basso 等人在文献[1]中对基于 SIDH 的 OPRF 进行了分析，给出了对 Boneh 等人提出的同源假设（One-more Assumption）的攻击。Basso 等人首先提出一种多项式时间攻击，该攻击打破了 OPRF 的伪随机性。他们的攻击允许敌手进行一些初始评估和离线计算后评估 OPRF，而无须与服务器进一步交互。对 OPRF 协议的简单修改可以防止此类攻击。因此研究人员又提出了第二种攻击方法。该亚指数的攻击在上述对策存在的情况下仍能成功。这两种攻击都破坏了 Boneh 等人提出的安全参数。此外，该研究验证了概念的实现并给出一些攻击的时机。最后，检查了 Boneh 等人的一个 OPRF 参数生成，认为需要一个可信的第三方来保证其可证明安全性。

2016 年，Galbraith 等人提出了一种针对 SIDH 密钥交换协议的自适应攻击。在 SIKE 中，通过应用 Fujisaki-Okamoto 变换的变体可以避免上述攻击。这种变换迫使 Bob 向 Alice 透露他的加密密钥，随后 Alice 使用该密钥对 Bob 的密文重新加密并验证其有效性。因此，Bob 不能重用他的加密密钥。支持静态密钥交换的对策主要有 k-SIDH 及其 Jao-Urbanik 变体。然而这些对策需要运行多个并行的 SIDH 实例，花费较高。

文献[2]首先提出了一种针对 SIDH 的 GPST 自适应攻击的新对策。该对策不需要像 SIKE 那样公开密钥，也不需要像 k-SIDH 那样多个实例并行。该方案将对策转化为 SIDH 型方案的一种密钥验证方法。其次利用该方法设计了一个高效的 SIDH 型交互式静态密钥交换协议 HealSIDH（Healed SIDH），与 k-SIDH 相比，HealSIDH 的效率高出一个数量级。再次利用 HealSIDH 设计了一个 PKE 方案 SHealS。SHealS 使用比 SIKE 更大的素数，具有更大的密钥和密文，但在完全执行方案时仅计算 4 个同源，而 SIKE 中有 5 个同源。基于该方案引入的一个新假设，证明了 SHealS 是 IND-CPA 安全的，并猜测了它的 IND-CCA 安全性。最后提

出 HealS，即一种使用较小素数的 SHealS 变体，提供相同的安全级别、较小的密钥和密文。与 SHealS 相比，HealS 的缺点是私钥不能用作加密密钥。

文献[3]中提出了延迟加密（Delay Encryption）这一新的密码原语，然后根据超奇异椭圆曲线同源和双线性对构造出高效的延迟加密的例子。延迟加密与时间锁问题（Time-lock Puzzles）和可验证延迟函数（Verifiable Delay Function，VDF）有关，并且可以被简单地描述为基于加密的时间锁标识（Time-lock Identity Based Encryption）。延迟加密在分布式协议中有着许多的应用，如密封投标 Vickrey 拍卖和电子投票。在此基础上该研究还通过修改 Boneh 和 Frankiln 的基于身份的加密方案，构造了一个新的延迟加密示例。主要方法是利用一根长的同源链来替换原方案中的主私钥，与 De Feo、Masson、Petit 和 Sanso 在同源可验证延迟函数中所用的技巧类似。和基于同源构造的可验证延迟函数一样，文献中的延迟加密在安全的使用参数之前需要可靠的设置。文中受信任的设置和可验证延迟函数一致。同时文中讨论了基于同源的延迟协议、分布式可信设置、水印和实现问题。

文献[4]对超奇异椭圆曲线同源的密钥交换协议（SIDH）因其自同态环非交换而没有亚指数级别的量子攻击这一普遍被人接受的观点进行了反驳。文中强调了阿贝尔群作用在 SIDH 密钥空间上的存在性，并给出了在 SIDH 参数不合理的情况下，这种群作用可以通过 SIDH 中的挠点信息有效计算。这将减小同源密码的困难性到隐藏移位问题，此问题可以在量子亚指数时间内被解决。文中将新攻击描述为更一般环境中的一个特殊实例，这使得其能够将新的密码分析与其他量子攻击结合。在此基础上，提出了一个关于群作用的函数 Malleability Oracle 来定义密钥所需的属性。在一些额外的假设下，通过解决隐藏子群的移位问题，访问这个谕言就可以计算函数的原像。

文献[5]构造了一个存在恶意敌手情况下的基于同源密码 UC 安全的不经意传输协议。该方案基于 CSIDH 的框架，考查了同源的计算效率。该协议底层的困难假设称为计算互反 CSIDH 问题（Computational Reciprocal CSIDH Problem）。此困难问题被证明等价于计算 CSIDH 问题。该方案首先通过对 Diffie-Hellman 协议的改动设计了一种 1-out-of-2 的 OT 协议以实现具有可信公共曲线的紧致 OT 原型。接下来将 3 轮协议通过曲线的二次扭转化为 2 轮的方案。该方案在半诚实模型中是高效的基于同源的 OT 协议。基于这种修改可以建立一种安全机制，接收方将向发送方展示“解密能力”，以便进行单边模拟。此外，研究者还建立了一种新的陷门算法，在设置中使用二次扭曲线来达到完全可模拟的构造。

文献[6]利用了 Prouhet-Tarry-Escott（PTE）problem 的解给出了寻找连续的 Smooth Number 的具体算法。对于一个固定的光滑界 B，将搜索限制为整数对增加了寻找到界为 B 光滑整数的概率。其算法将一个简单的筛选与 PTE 问题的一系列解决方案给出的参数相结合。寻找这些孪生光滑整数的动机是这些数在基于同源的后量子协议中有着重要的应用。在 B-SIDH 和 SQISign 中需要两个连续的光滑整数，并且满足两个连续的光滑整数的和是一个素数。在寻

找密码参数的过程中，应用文中的方法可以找到素数 p，使得 p+1 和 p–1 是2^{15}-光滑的。在更高的安全性下，文中的筛法找到了一个 376 位的素数是2^{21}-光滑的、384 位的素数是2^{22}-光滑的、512 位的素数是2^{28}-光滑的。作者发现以前的文献中还未有计算满足光滑性要求素数的方法。

本节作者：王滨、袁思蒙、于伟（中国科学院信息工程研究所）

参考文献

[1] BASSO A, KUTAS P, MERZ S-P, et al. Cryptanalysis of an oblivious PRF from supersingular isogenies [C]. ASIACRYPT 2021 (1): 160-184.

[2] FOUOTSA T B, PETIT C. SHealS and HealS: isogeny-based PKEs from a key validation method for SIDH [C]. ASIACRYPT 2021 (4): 279-307.

[3] BURDGES J, DE FEO L . Delay Encryption [C]. EUROCRYPT 2021 (1): 302-326.

[4] KUTAS P, MERZ S-P, PETIT C, et al. One-Way Functions and Malleability Oracles: Hidden Shift Attacks on Isogeny-Based Protocols. EUROCRYPT 2021 (1): 242-271.

[5] LAI Y, GALBRAITH S D, Cyprien Delpech de Saint Guilhem. Compact, Efficient and UC-Secure Isogeny-Based Oblivious Transfer [C]. EUROCRYPT 2021 (1): 213-241.

[6] COSTELLO C, MEYER M, NAEHRIG M. Sieving for Twin Smooth Integers with Solutions to the Prouhet-Tarry-Escott Problem [C]. EUROCRYPT 2021 (1): 272-301.

4　多变量公钥密码

多变量公钥密码（Multivariate Public Key Cryptosystem，MPKC）被认为是抵抗量子计算攻击的一类候选密码算法，其安全性基于求解有限域上随机产生的一组多变量二次多项式（Multivariate Quadratic，MQ）方程组问题的困难性。不平衡油醋（Unbalanced Oil and Vinegar，UOV）体制和隐藏域方程（Hidden Field Equation，HFE）体制一直是 MPKC 研究的两大热点。进入 NIST 后量子密码标准征集第 2 轮的 3 个算法中 LUOV（Lifted UOV）和 Rainbow 体制是 UOV 类的数字签名算法，GeMSS 是 HFE 类数字签名体制。最终，Rainbow 进入了 NIST 标准征集的第 3 轮，GeMSS 入选了第 3 轮的数字签名替补算法。

LUOV 的核心思想是将体制中心映射多项式定义在$\mathbb{F}_2$的扩域$\mathbb{F}_{2^r}$上，而将其系数限制为二元域$\mathbb{F}_2$中的元素，从而达到降低公钥量的目的。当 r 是合数时，丁津泰等人提出了一种子域差分攻击（Subfield Differential Attack，SDA），将 LUOV 的公钥转换到$\mathbb{F}_{2^r}$与$\mathbb{F}_2$之间的一个中间子域上，成功地伪造了合法签名。随后，LUOV 的作者改进了方案，将 r 指定为素数，

以避免中间子域的出现。然而，丁津泰等人修改了 SDA 攻击方法，提出了嵌套式子域差分攻击（Nested Subfield Differential Attack，NSDA）。攻击表明，改进后的 6 个 LUOV 具体方案有一半没有达到 NIST 设定的安全目标等级。具体来说，文献[1]利用了扩域中元素可表示为子域上的多项式这一特性，逐步进行差分测试，最终获得了任意消息的合法签名。文献[1]给出了满足 NIST 第一类安全性的 LUOV 的签名伪造攻击实验，在 210 分钟内成功地伪造了合法签名。文献[1]还指出，SDA 和 NSDA 攻击方法并没有用到 UOV 体制的特殊结构，因此可看作是对任意“提升”类体制的攻击方法，但这些方法对于非“提升”类体制并不适用。

油醋类签名体制的结构比较特殊。体制中包含两类变量：油变量和醋变量。对于给定的消息，随机选取醋变量的值，就可得到关于油变量的线性方程组，求解该方程组即可得到给定消息的签名值。因此，油醋类签名体制的公钥构造是将油变量空间隐藏起来。一旦敌手找到了这个油变量空间，就可以伪造任意消息的合法签名。最早，Kipnis 和 Shamir 提出了一种针对平衡油醋签名方案的攻击（Kipnis-Shamir 攻击），该攻击能够找到公钥中的隐藏油空间，同样可以作用于醋变量是油变量个数两倍的情形。为了抵抗 Kipnis-Shamir 攻击，Rainbow 采用的是多层油醋结构。文献[2]在 Kipnis-Shamir 攻击基础之上，提出了交集攻击（Intersection Attack）。在该攻击下，UOV 和 Rainbow 的安全性受到影响，如 Rainbow 后量子标准征集第 3 轮提交的 Ia 级方案，抗现有攻击的安全强度从 2^{147} 降到了 2^{140}。具体来说，文献[2]发现当醋变量的个数小于油变量的个数 2.5 倍和 3 倍时，可以通过寻找油空间的两个仿射空间的交集来寻找油空间。文献[2]还提出一种长方形最小秩攻击（Rectangular MinRank Attack）用来分析 Rainbow 体制。长方形最小秩攻击将 Rainbow 的私钥恢复归约为一个最小秩问题。攻击方案中用的矩阵不是方阵而是长方形矩阵。在该攻击下，Rainbow 第 3 轮参数的安全强度分别降低了 $1/2^{20}$、$1/2^{40}$ 和 $1/2^{55}$。

在 UOV 类体制的改进方面，目前研究主要集中在如何降低其公钥规模。文献[3]在破解了分块反循环 UOV 的基础之上，推广了分块反循环的思想，提出了商环 UOV 方案（Quotient Ring UOV，QR-UOV）。合适的参数设置可使得该方案的安全强度达到 NIST 后量子密码标准征集的安全性要求 I、III、V，而且在同等安全性强度下，其公钥量比 Rainbow 要少 50%~70%，不过其签名长度要比 Rainbow 长一些。具体来说，QR-UOV 仍然采用分块矩阵的构造方式，其矩阵中的各个分块源于商环上多项式对应的多项式矩阵，存储该类矩阵仅需要存储多项式的系数即可，从而达到降低公钥量的目的。

GeMSS 作为 NIST 后量子密码征集第 3 轮的替补算法，同样受到了关注。GeMSS 是一类 HFE 结合了减方法和醋变量方法的变体（HFEv-）。现有的对于 HFE 类体制安全性分析的最小秩攻击都是在中心映射所在的扩域上进行分析。文献[4]给出了一种在基域上的最小秩攻击，能够找到 HFEv-类体制的等价密钥，从而可以成功伪造任何信息的合法签名。具体来说，文献[4]发现减方法并没有增强 HFE 类体制的安全性，而醋变量方法对其攻击的复杂度影响

也仅仅是增加了一个多项式因子。文献[4]指出在其攻击下，GeMSS 方案现有的参数集达不到 NIST 后量子密码标准征集的安全级别。要提高 HFE 类体制的安全性必须提高其中心映射多项式的次数，而这又将导致体制的效率降低。现有的技术无法同时确保 HFE 类体制的安全性和高效性。

对于多变量公钥密码的直接攻击就是求解有限域上非线性多项式方程组，受到了广泛的关注。该方法还可以应用于对称密钥密码体制的代数攻击。文献[5]设计了一种基于多项式方法的高效算法来求解$\mathbb{F}_2$上多变量多项式方程组。具体来说，文献[5]中提出的算法结合了现有算法的优化和简化方法，并利用$\mathbb{F}_2$上内存约减的默比乌斯变换的一个变体来优化内存，使得算法在保持时间复杂度同时降低了空间复杂度。文献[5]将其算法应用于分析 NIST 后量子密码征集第 3 轮中的数字签名备选算法 Picnic（该算法的构造是基于分组密码算法 LowMC），指出 Picnic 3 个新实例中有两个未达到相应的安全级别。文献[5]还将其算法应用于 Keccak 缩减轮数的原像攻击和碰撞攻击中，降低了原有攻击的复杂度。

本节作者：聂旭云、何晨宁（电子科技大学）

参考文献

[1] DING J,DEATON J, VISHAKHA Y B. The Nested Subset Differential Attack [C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology - EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12696, pp: 329-347. Springer, Cham.

[2] BEULLENS W. Improved Cryptanalysis of UOV and Rainbow [C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology - EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12696, pp: 348-373. Springer, Cham.

[3] FURUE H, IKEMATSU Y, KIYOMURA Y, et al. A New Variant of Unbalanced Oil and Vinegar Using Quotient Ring: QR-UOV [C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093, pp: 187-217. Springer, Cham.

[4] TAO C, PETZOLDT A, DING J. Efficient Key Recovery for All HFE Signature Variants [C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12825, pp: 70-93. Springer, Cham.

[5] DINUR I. Cryptanalytic Applications of the Polynomial Method for Solving Multivariate Equation Systems over GF(2) [C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology - EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12696, pp: 374-403. Springer, Cham.

安全协议

1 混淆/不可区分性

程序混淆源于软件工程领域反编译的需求，同时在计算复杂性和密码学领域有着很高的理论价值与应用前景。2001 年，Barak、Goldreich、Impagliazzo 等人首次提出了混淆的两种正式定义，它们分别是虚拟黑盒混淆和不可区分混淆（indistinguishability Obfuscation，iO）。同时，他们证明了虚拟黑盒混淆不可能将任意电路混淆为另一个电路。但是，这一结论并没有将量子的相关机制纳入考量范围，而经典条件下成立的结论在量子条件下有可能会被打破。在文献[1]中，作者进一步证明了，虚拟黑盒混淆不可能将任意电路混淆为某些量子状态。

由于虚拟黑盒混淆不可能存在，研究者们致力于构造可证明安全的 iO。自 2015 年以来，一系列漂亮的工作将构造 iO 这一大目标不断拆分成很多小目标——构造一些更为简单的密码学原语。到 2019 年时，构造一个特殊的伪随机数发生器成了构造 iO 的最后一块拼图，而这一目标在当时看来难以由标准假设实现。同时，由于在构造 iO 时引入了大量的密码学原语，使得整个 iO 的构造十分复杂，内部组件之间的归约关系较为混乱。在文献[2]中，作者考查了截至 2020 年构造 iO 所需要使用的假设，并提出了足以构造特殊的伪随机数发生器的新假设。对于新假设，他们给出了一系列标准，如假设应当是容易描述的。同时，作者对之前的构造框架进行了一定程度的简化。

由于文献[2]中的构造需要使用双线性映射相关的假设，而这些假设会被量子计算机打破。因此有一些研究试图基于格上困难问题来构造后量子安全的 iO。在 2020 年，Brakerski、Döttling、Garg 等人通过基于 LWE 的全同态加密构造了 iO，但是他们没能给出相应的安全证明，同时使用了 Random Oracle（RO）模型。在 2021 年，Gay 和 Pass 用公共引用串代替了 RO，给出了标准模型下的 iO 构造。同时，他们将这一构造的安全性归约到 LWE 假设和一个更强的循环安全假设上。在文献[3]中，作者总结了前面的工作，指出构造 iO 并不需要使用同态加密，即密文不需要对应于唯一的明文，他们将这一密码学原语命名为 Functional Encoding。同时，作者通过基于 LWE 的同态承诺和不经意 LWE 采样器构造出了 iO。不经意 LWE 采样器并没有已知的基于标准假设的构造，他们在文中给出了一个启发式的构造，这一构造同样基于一个特殊的循环安全假设。

本节作者：郁昱、姚立（上海交通大学）

参考文献

[1] ALAGIC G, BRAKERSKI Z, DULEK Y, et al. Impossibility of Quantum Virtual Black-Box Obfuscation of Classical Circuits [C]. CRYPTO 2021(1): 497-525.

[2] GAY R, JAIN A, LIN H, et al. Indistinguishability Obfuscation from Simple-to-State Hard Problems:New Assumptions, New Techniques,and Simplification [C]. EUROCR-YPT 2021(3): 97-126.

[3] WEE H, WICHS D. Candidate Obfuscation via Oblivious LWE Sampling [C]. EUROCRYPT 2021(3): 127-156.

2　零知识证明

零知识证明是现代密码学中一个强有力的工具，它允许证明者在不泄露私密信息的前提下向验证者证明一个断言的正确性。一方面，零知识证明可以被用来构造其他的关键密码学组件；另一方面，零知识证明是在不可信网络环境下建立信任和设计隐私保护方案的核心工具。本节简要介绍 2021 年零知识证明研究领域的主要进展。

2.1　零知识证明的理论进展

ZAP 是两轮公开掷币的证据不可区分协议，是由 Dwork 等人提出的一个零知识证明的重要变体，被广泛应用于构造零知识证明以及其他密码学组件。时至今日，基于离散对数衍生假设（如 DDH 假设）的 NIZK 和 ZAPs 的构造仍然是公开问题。Jain 等人在文献[1]中提出了第一个基于 DDH 假设的 NIZK 和 ZAP 方案，解决了这一公开问题。该工作基于 DDH 假设构造一个相关性难解哈希函数（CIH），利用这样的 CIH 与 FS 启发式方法对交互式零知识证明方案进行转化，进而得到公共随机串模型下基于 DDH 假设的统计零知识的 NIZK。如果将零知识性质弱化为计算零知识，那么可以得到自适应合理性。类似，作者利用 CIH 函数将一个带陷门的 Σ 协议转化为一个基于 DDH 假设的 ZAP 协议，这也是第一个基于群的统计不可区分 ZAP 协议。

由于近年来量子计算机的快速发展以及量子计算所展现出的强大能力，研究学者开始愈加重视量子环境下安全的密码方案的设计及相关理论。Ananth 等人在文献[2]中考查了量子零知识协议在量子环境下的并发合成，结合 Watrous 的量子重绕引理与区块重绕策略，实现了有界量子并发下安全的对 NP 断言与 QMA 断言的量子零知识证明方案的构造，并进一步地给出了有界量子并发下安全的对 NP 断言的量子零知识的知识证明协议的构造。Chia 等人[3]

提出了一种全新的量子重绕技术，其使得模拟器在某种意义上能够抽取到验证者承诺的消息，模拟验证者的内部状态。作者基于此技术以及 collapsing hash functions 构造了一个常数轮的抗量子黑盒ϵ −零知识证明协议，并基于抗量子的单向函数构造了常数轮的抗量子黑盒ϵ −零知识论证协议。Zhang 等人在文献[4]中给出了一类特殊的知识证明协议的定义，该定义考查经典意义下的验证者与量子能力的证明者，其中证明者拥有的证据是量子态的。该工作给出了这种定义的几种有趣的性质及应用，并提供了若干实例。

在经典环境下，（基于格的）Σ 协议可以通过 FS 启发式方法转化为 RO 模型下安全的非交互零知识协议。然而这种转换需要对 Σ 协议有特殊的限制，才能保证转化后得到的协议在量子 RO 模型下仍然安全。Unruh 等人给出第一个量子计算环境下安全的转化方法，但其有着较高的额外开销。Katsumata 等人在文献[5]中构造了一种新的原语——可抽取线性同态承诺协议，并利用该协议提出了一种新的转化方法。这种方法可以将多种新近提出的无法使用 FS 转换的 Σ 协议转化成量子 RO 模型下安全的 NIZK，且相较于另一种通用的 Unruh 转换，其证明尺寸要小得多。Shmueli 等人在文献[6]中给出一个将任意 Σ 协议转化为恶意指定验证者的非交互零知识证明协议的通用转化方式。尽管这项技术是经典的，但仍适用于量子协议，并可以对 QMA 语言构建可复用的恶意指定验证者非交互零知识证明协议。

与经典环境类似，量子承诺方案也是量子计算环境下零知识协议构造的关键组件。国内学者颜俊在文献[7]中考查了由量子比特承诺方案在并行合成下得到的量子比特串承诺方案的绑定性，并证明了这类方案满足一种更强的绑定性——谓词绑定性。利用这一观察，结合量子重绕引理，作者给出了第一个对任意 NP 断言的基于量子安全单向置换（函数）的量子完美（统计）零知识论证协议（合理性错误为 1/2）。

抗参数篡改零知识（sub-ZK）的 NIZK 在公共参考串恶意生成的情况下仍然保持零知识性。Fauzi 等人[8]在可抽取单向函数与广义（Generalized）可抽取单向函数的基础上提出了验证可抽取的（广义）可抽取单向函数（VE(G)OWFs）这一概念，并探究了这些概念间的内在联系以及若干实例化。该文献进一步地展示了 VE(G)OWFs 与 sub-ZK NIZK 的关系：利用 VE(G)OWFs，我们可以将满足特定性质的 NIZK 转化成 sub-ZK NIZK；给定 sub-ZK NIZK，我们也可以直接构造 VE(G)OWFs。

简洁非交互论证（SNARG）协议是一类通信效率表现较好且有实用意义的论证协议。但 Gentry 与 Wichs 曾给出一个重要的负面黑盒分离结果，即在黑盒归约框架下，对 NP 断言的自适应合理性的 SNARG 无法由可证伪假设构造得到。Lipmaa 等人在文献[9]证明了上述不可能性结论是紧的，该工作更具体地构造了第一个基于可证伪假设的非自适应合理性的 SNARG 方案。这个方案同时满足非自适应的知识的合理性与 Sub-zk 的性质（后两个性质需要非可证伪假设）。

RO 模型下的 SNARG 方案可以由轻量化的密码学组件（如密码学安全的哈希函数）实现

且无须可信初始化，因此引起了研究学者的大量关注。Micali 等人在 1994 年第一次给出了将 PCP（Probabilistically Checkable Proof，概率可检验证明）转化为 RO 模型下 SNARG 的方案。一个 RO 模型下基于 PCP 的 SNARG 方案，如果恶意证明者能够进行 t 次查询，且欺骗成功的概率最多为 ϵ，我们说该方案是(t,ϵ) −安全的。在已有的安全性归约证明中，要求$\epsilon_{\mathrm{PCP}} \leqslant \frac{1}{2} \cdot \frac{\epsilon}{t}$以及$\lambda \geqslant \log_2(8 \cdot \frac{t^2}{\epsilon})$，其中$\epsilon_{\mathrm{PCP}}$是所使用 PCP 方案的合理性错误。长期以来 RO 模型下(t,ϵ) −安全的 SNARG 方案的通信复杂度都是$O\left(\left(\frac{t}{\epsilon}\right)^2\right)$，Chiesa 等人在文献[10]中利用“chopped tree”技术优化了打开承诺的通信开销，第一次得到通信复杂度为$O(\frac{t}{\epsilon} \cdot \log_2 \frac{t}{\epsilon})$的 RO 模型下基于 PCP 的 SNARG 方案。

知识假设是一类特殊的不可证伪假设，通常是指存在一个多项式时间的提取器，以某些群元素（或某个哈希值）为输入，能够输出这些群元素对应的指数（或哈希值的前象）。这些被提取出来的值通常是一个断言的证据。一方面，知识假设常被用于 zkSNARK 方案的安全性证明；另一方面，由于密码方案的规模和复杂性不断增长，证明协议的合成也变得越来越重要。在对基于知识假设的证明协议进行合成时，由于这种特殊抽取器的存在，模拟器往往能够得到这个系统的所有信息。但这种能力在常用的合成技术，如通用合成和构造性密码技术中是不被允许的。知识假设的这种性质给基于知识假设的大量高效零知识证明方案的合成带来了极大困难，已有的解决方法通常会破坏证明方案原本的简洁性质。Kerber 等人在文献[11]中提出 Knowledge-Respecting 区分环境的概念，区分器需要解释其所输出的隐含知识的元素（如群元素或者哈希值）是如何计算出来的，且允许模拟器访问这些解释，然后探究了这种区分器存在的条件。基于此，作者给出结论：对基于知识假设方案完整的一般性合成是不存在的，但对基于不同知识假设的方案进行合成则是可行的。文献最终给出了第一个能够保证简洁性质且支持合成的 zkSANRK 方案。

对于通用目的零知识证明方案，即使要证明的关系或者相应的零知识证明协议都只需要以黑盒的方式调用某个功能函数，证明安全性时模拟器仍然需要该功能函数的完整代码，因为需将证明的断言归约为零知识证明协议适用的 NP 完全语言。这就使得黑盒构造（以黑盒的方式调用某个密码学功能函数）的通用目的零知识证明方案的设计面临着巨大挑战。2012 年 Rosulek 等人给出一个负面结果——即使是对一个简单的断言如对单向函数的范围成员关系证明，仍然需要以非黑盒的方式去访问所依赖的功能函数。Pandey 等人在文献[12]中引出 proof-based 的概念，即对于一个功能函数如某个单向函数 f，一个新的函数 F 仅仅以黑盒的方式访问 f 的输入-输出，则 F 是 f 的 proof-based 版本，通过对 F 构造相应的零知识证明方案

来实现黑盒构造。作者探究了各种密码学功能函数如单向函数、伪随机生成器的 proof-based 版本的构造。该文献给出了一个负面结果，如果所有的输入都由证明者选取，就无法由单向函数黑盒构造 proof-based 伪随机生成器（伪随机生成器能够由单向函数黑盒构造，这是伪随机生成器和 proof-based 伪随机生成器的一个分离结果）；还给出了一个正面结果，若对输入加以限制和形式上的调整，则能够给出单向函数、伪随机生成器、抗碰撞哈希函数的 proof-based 版本构造，并以此给出相应的范围成员关系证明的完全黑盒构造。这给未来黑盒构造的通用目的零知识证明方案的设计带来了希望。

一些高效零知识证明方案，尤其是交互式协议经由 FS 转换得到的 NIZK，如 Bulletproof，通常缺少具体的安全性分析。这使得在应用这些方案时的参数选择以及方案间同安全性下的性能对比变得困难。Jaeger 等人在 2020 年给出了（交互式的）Bulletproof 在一般群模型下的具体安全性分析。然而 Bulletproof 通常以非交互的形式被应用，且一般群模型是一种较为理想的模型。Ghoshal 等人在文献[13]中给出了代数群模型下状态恢复合理性的概念，这是一种更强的合理性要求，允许证明者重置验证者。证明者的这种行为与交互式协议经由 FS 方法转化为非交互协议的安全性分析中提取器的行为类似，因此建立了一个对 FS 转换得到的非交互协议进行安全性分析的桥梁。作者给出了一个将状态恢复合理性与证据扩展模仿真性质结合起来的框架，并将此框架应用于 Bulletproof 和 Sonic 中，第一次给出了对两种方案的代数群模型下的具体安全性分析。

2.2 实用零知识证明的进展

实际应用场景有着对效率要求高、影响因素和具体要求复杂等特点。研究学者从应用角度出发，结合不同场景的不同要求，给出了具有不同特征的实用零知识证明方案的构造。

Bottle 等人在 Bulletproof 中给出了一个将适用于内积断言的论证协议通信复杂度从线性优化到对数级别的迭代技巧。Cramer 等人在 2020 年的工作将这一技巧从内积断言扩展到线性关系断言，提出了压缩Σ协议理论。在此基础上，Attema 等人在文献[14]中将其应用于格上的Σ协议，得到一个效率更高的对格上问题的交互式零知识证明协议。Attema 等人在文献[15]中对压缩Σ协议理论进行了延伸扩展，并基于此构造了具有对数级别通信复杂度的关于 k-out-of-n 的部分知识的零知证明协议。作者还更进一步地介绍了如何在证据满足某些附加约束时进行证明以及如何将这些结果推广到非阈值访问结构。

2019 年，Lai 等人将 Bulletproof 中对算数电路的零知识证明协议扩展，首次得到对基于双线性群的电路的直接零知识证明方案（无须先将电路转化为算术电路）。Attema 等人在文献[16]中通过对压缩Σ协议的技巧进行扩展，提出了一种对基于双线性群电路直接设计零知识证明协议的方法，并在通信复杂度上有了 3 倍的提升。以此方法为工具，该文献给出了首个签名尺寸为$O(\log_2 n)$且无须可信初始化的门限签名方案；还给出了一个用于证明多个签名满

足某个公共约束关系的证明方案。

通用目的（对所有 NP 语言）的高效 zkSNARK 方案能够被广泛应用于各种实际场景，一直是实用高效零知识证明的研究热点。Couteau 等人在 2020 年给出一种将某些交互式协议转化为非交互协议的框架——CH 转换，其核心思路是将挑战嵌入到某个非对称双线性群的指数上。Couteau 等人在文献[17]中对 CH 框架进行了显著改进：给出设计用于证明某个加密向量属于特定代数集合的 NIZK 的通用方法，进而得到一个更为一般化、适用于更多语言的框架，且其依赖的假设更弱。以此框架为工具，作者构造了对所有 NP 语言的高效 NIZK 方案。Bünz 等人在 2019 年基于未知阶群给出了一个多项式承诺——DARK 方案，并以此构造了通信复杂度最优的、无须可信初始化的 zkSNARK 方案——supersonic。Rothblum 等人在文献[18]中指出 DARK 方案的安全性证明存在问题，并利用格上的相关技术，给出了 DARK 方案的一种改进，绕过了其安全性证明中存在的问题。同时改进后的多项式承诺方案，构造了时间与运行空间均高效的零知识论证协议。

在实际应用中，经常会遇到同时对多个断言进行证明的场景。最平凡的方法是对这些断言逐一证明，效率与断言的数量呈线性关系。研究者通过探究这些断言以及相应零知识证明协议之间的内在关联，给出了效率远优于平凡策略的批量证明方法。Choudhuri 等人在文献[19]中给出了第一个基于标准假设的公共参考串模型下对 NP 语言的批量论证协议，可以同时对多个断言进行批量证明，其通信开销远小于分别独立证明这些断言开销的和，从而实现了高效的批量证明。该文献在 Reingold 等人的工作基础之上，将批量承诺方案应用于 Spartan 协议，得到一个“双模”交互式的批量论证协议，然后利用 CIH 函数和 FS 启发式方法将其转化为标准假设下的非交互批量论证协议。Bünz 等人在文献[20]中将内积断言扩展到双线性映射，称为 pairing 内积断言，并对该断言给出了一个高效的、证明尺寸为$6\log_2 n$个群元素的论证协议。以此论证协议为工具，该文献给出了第一个验证者计算复杂度为对数级别、证明者计算复杂度为次线性、参考串长度为次线性的多项式承诺方案；并以此给出一个对 Groth 的 zkSNARK 方案的批量证明方法，与已有批量证明方法相比，效率有极大的提升。

Kaslasi 等人在 TCC 2020 上构造了对任意有非交互统计零知识证明的语言（NISZK）的批量验证协议。Kaslasi 等人在文献[21]中指出了该工作的两个主要不足之处：①不是公开掷币的；②要求验证者是诚实的。对于 NISZK 中的语言，该文献首先构造了一个公开掷币的诚实验证者零知识的批处理协议，然后在保持效率的前提下，将其转化为能应对恶意验证者的批处理协议，从而在批量处理效率与已有工作相似的前提下，解决了这两个不足。

基于线性 PCP 的 zkSNARK 方案（证明尺寸甚至只有常数个群元素）以及基于 pairing 的密码学方案，需要带结构的 CRS。为了保证协议的安全性，CRS 只能由可信第三方来生成，但现实中很难找到能够扮演可信第三方的角色。因此，如何应对恶意的 CRS 生成方，成了近年来研究者关注的热点问题。Groth 等人在 2018 年给出了可更新参考串的概念，即所有的参

与方都可以对参考串进行修改，只要有一个参与方是诚实的，即可保证方案的安全性，极大地降低了对可信第三方的依赖。Campanelli 等人在文献[22]中提出多项式全息 IOP 的概念并给出相关实现，基于 pairing 构造了对多项式的 CP-zkSNARK 方案。然后作者将两者结合起来，得到对多项式的 zkSNARK 方案，进而得到对算数电路可满足问题的可更新参考串的 zkSNARK 方案。在所有可更新参考串方案中，该方案有着最小的证明尺寸，在某些情况下还有着最优的证明者计算效率。

除此之外，Ananth 等人在文献[23]中提出了 CRS 生成问责的概念，即若 CRS 生成方将协议参与方的私密信息提供给他人，则我们有能力提供一个公共可验证的证据来证明 CRS 生成方有不法行为。通过让 CRS 生成方衡量“做坏事”的收益和“做坏事”被抓后的损失来限制其不法行为。作者探究了问责性质在不同场景如 NIZK 和两轮的两方安全计算中的具体内涵，并分别给出了具体实现。主要技术手段是将嵌入了某个秘密值的协议执行副本作为 CRS 的一部分（如 NIZK 证明），使得存在一个提取器能够通过恶意 CRS 生成方的不法行为将其秘密值提取出来作为不法行为的证据，从而进行问责。其他的解决方法是使用分布式的 ceremony 协议取代可信第三方来生成公共参考串。文献[24]考查了这类协议的安全性，并给出了一个关于带 ceremony 协议的 zkSNARk 的安全性框架。基于该框架，该文献重新考查了已有方案中为经典的 Groth16 所设计的 ceremony 协议，简化了其构造，并避免了该构造中对 random beacon 模型的依赖。

现如今多数高效 NIZK 的构造都致力于在证明尺寸上进行优化，没有考虑证明生成过程中对设备内存的要求。降低证明方案对内存的要求有助于扩展零知识证明的应用范围，使其可以部署在小内存设备。内存效率优化通常采用 gate-by-gate 的模式，即分别对每个门进行证明。Carsten Baum 等人[25]给出了一种内存优化的 zkSNARK——Mac'n'Cheese。该方案采用 commit-and-prove 范式，将信息论安全的 MAC（Message Authentication Code，消息认证码）与 sVOLE（subfield Vector Oblivious Linear Evaluation，向量不经意线性估值）结合起来得到一个承诺方案并对电路计算的证据进行承诺，然后还给出一般性的 OR 断言合成证明方法，使得对 m 个断言的 OR 合成的证明长度仅与最长的断言长度相关（而不是所有断言长度的和），极大地优化了 OR 断言合成证明的效率。以此为工具，证明者只需要证明某个分支上的计算是正确的，极大地优化了效率。Mac'n'Cheese 有着优于已有工作的实际效率表现，且对批量处理的支持更好。

Feng 等人[26]在标准假设下构造了一个新的原语——可认证证据的非交互零知识证明系统。该证明系统使得拥有某个断言x的证据w的人能够确认该断言的一个有效证明是否是利用证据w产生的。此外，该文献还给出了这样的证明系统在不可延展（完全单向）哈希函数、验证者本地撤销的群签名以及明文可验证公钥加密上的应用。

2.3　零知识的应用

零知识证明被广泛应用于具体场景下隐私保护方案的设计。

ECDSA 近年来被用于比特币以及某些其他密码学货币，适合分布式部署的门限 ECDSA 签名方案成了近年来的研究热点。Castagnos 等人在 2020 年基于类群上的 CL 加密方案给出了一个通信复杂度最优的门限 ECDSA 方案。Yuen 等人在对 CL 密文相关的零知识证明中，通过引入额外一轮挑战用于消除证明者首轮消息中的低阶元素，避免了由保证合理性而导致的过多重复执行，从而得到了一个对 CL 密文相关断言的效率更高的零知识证明协议，并以此构造了效率更优的门限 ECDSA 方案。但这两种方案都需要假设通过某种初始化过程可以得到不存在不安全的小阶元素的类群。这种假设是一种全新的未经检验的假设，且没有说服力强的安全性阐释。

国内学者邓燚等人在文献[27]中，通过对 CL 密文相关断言的证明以及门限 ECDSA 方案实际所需安全性进行深入观察，提出弱化了合理性要求的 Promise Σ协议的概念，并证明这种弱化了的合理性仍然能够满足一些实际场景的需求。作者以此为工具，在不引入小阶元素假设的前提下，给出了两方/多方门限 ECDSA 方案，且相比已有方案在签名密钥生成阶段通信/计算效率都更高。

范围证明通常被用来证明某个被承诺/密文隐藏起来的整数属于某个声称的范围，在密码货币以及区块链交易系统等隐私保护方案的构造中发挥着重要作用。Couteau 等人在文献[28]中给出了一个将有限域上承诺方案转化为有界整数上的承诺方案的通用转化方法，得以将平方分解方法用于有限域上的承诺方案，进而得到构造范围证明的模块化方法。在离散对数假设下，得到最有效和最有影响的范围证明。与同样基于离散对数假设的 Bulletproof 相比，范围大小和安全参数缩短 12%~20%，而且无论对证明者还是验证者效率都要高出不止一个数量级；在 LWE 假设下得到的范围证明与现有方案相比，支持更多数量的批量处理；与类群上的假设结合，得到了第一个类群上的无须可信初始化的高效范围证明方案。

本节作者：邓燚、汪海龙、张心轩、谭陶莉、朱旭东（中国科学院信息工程研究所）

参考文献

[1] JAIN A, JIN Z. Non-Interactive Zero Knowledge from Sub-exponential DDH[C]. EUROCRYPT 2021 (1): 3-32.

[2] ANANTH P, CHUNG K, ROLANDO L, et al. On the Concurrent Composition of Quantum Zero-Knowledge[C]. CRYPTO 2021 (1): 346-374.

[3] CHIA N, CHUNG K, YAMAKAWA T. A Black-Box Approach to Post-Quantum Zero-

Knowledge in Constant Rounds[C]. CRYPTO 2021 (1): 315-345.

[4] ZHANG T, VIDICK T. Classical proofs of quantum knowledge[C]. EUROCRYPT 2021 (2): 630-660.

[5] KATSUMATA S. A New Simple Technique to Bootstrap Various Lattice Zero-Knowledge Proofs to QROM Secure NIZKs[C]. CRYPTO 2021 (2): 580-610.

[6] SHMUELI O. Multi-theorem Designated-Verifier NIZK for QMA[C]. CRYPTO 2021 (1): 375-405.

[7] YAN J. Quantum Computationally Predicate-Binding Commitments with Application in Quantum Zero-Knowledge Arguments for NP[C]. ASIACRYPT 2021 (1): 575-605.

[8] FAUZI P, LIPMAA H, SIIM J, et al. Verifiably-Extractable OWFs and Their Applications to Subversion Zero-Knowledge[C]. ASIACRYPT 2021 (4): 618-649.

[9] LIPMAA H, PAVLYK K. Gentry-Wichs Is Tight: A Falsifiable Non-Adaptively Sound SNARG[C]. ASIACRYPT 2021(3): 34-64.

[10] CHIESA A, YOGEV E. Subquadratic SNARGs in the Random Oracle Model[C]. CRYPTO 2021 (1): 711-741.

[11] KERBER T, KIAYIAS A, KOHLWEISS M. Composition with Knowledge Assumptions[C]. CRYPTO 2021 (4): 364-393.

[12] LIANG X, PANDEY O. Towards a Unified Approach to Black-Box Constructions of Zero-Knowledge Proof[C]. CRYPTO 2021 (4): 34-64.

[13] GHOSHAL A, TESSARO S. Tight State-Restoration Soundness in the Algebraic Group Model[C]. CRYPTO 2021 (3): 64-93.

[14] ATTEMA T, CRAMER R, KOHL L. A Compressed Sigma-Protocol Theory for Lattices[C]. CRYPTO 2021 (2): 549-579.

[15] ATTEMA T, CRAMER R, FEHR S. Compressing Proofs of k-Out-Of-n Partial Knowledge[C]. CRYPTO 2021 (4): 65-91.

[16] ATTEMA T, CRAMER R, RAMBAUD M. Compressed Sigma-Protocols for Bilinear Group Arithmetic Circuits and Application to Logarithmic Transparent Threshold Signatures[C]. ASIACRYPT 2021(4): 526-556.

[17] COUTEAU G, LIPMAA H, PARISELLA R, et al. Efficient NIZKs for Algebraic Sets[C]. ASIACRYPT 2021 (3): 128-158.

[18]BLOCK A R, ROSEN A, ROTHBLUM R D, et al. Time- and Space-Efficient Arguments from Groups of Unknown Order[C]. CRYPTO 2021 (4): 123-152.

[19] CHOUDHURI A R, JAIN A, JIN Z. Non-Interactive Batch Arguments for NP from

Standard Assumptions[C]. CRYPTO 2021 (4): 394-423.

[20] BÜNZ B, MALLER M, Mishra P, et al. Proofs for Inner Pairing Products and Applications[C]. ASIACRYPT 2021 (3): 65-97.

[21] KASLASI I, ROTHBLUM R D, VASUDEVAN P N. Public-Coin Statistical Zero-Knowledge Batch Verification against Malicious Verifiers[C]. EUROCRYPT 2021 (3): 219-246.

[22] CAMPANELLI M, FAONIO A, FIORE D, et al. Lunar: a Toolbox for More Efficient Universal and Updatable zkSNARKs and Commit-and-Prove Extensions [C]. ASIACRYPT 2021 (3): 3-33.

[23] ANANTH P, ASHAROV G, GOYAL V, et al. Towards Accountability in CRS Generation[C]. EUROCRYPT 2021 (3): 278-308.

[24] KOHLWEISS M, MALLER M, SIIM J, et al. Snarky Ceremonies [C]. ASIACRYPT 2021 (3): 98-127.

[25] BAUM C, MALOZEMOFF A J, ROSEN M B, et al. Mac'n'Cheese: Zero-Knowledge Proofs for Boolean and Arithmetic Circuits with Nested Disjunction [C]. CRYPTO 2021 (4): 92-122.

[26] FENG H, TANG Q. Witness Authenticating NIZKs and Applications [C]. CRYPTO 2021 (4): 3-33.

[27] DENG Y, MA S, ZHANG X, et al. Promise Sigma protocol: How to Construct Efficient Threshold ECDSA from Encryptions Based on Class Groups [C]. IACR Cryptology ePrint Archive 2022.

[28] COUTEAU G, LIN H, KLOOß M, et al. Efficient Range Proofs with Transparent Setup from Bounded Integer Commitments [C]. EUROCRYPT 2021 (3): 247-277.

3　承诺协议

承诺协议（Commitment）是一个包含两个参与方的两阶段交互协议，其中一个参与方称为承诺者（或发送者），另一个参与方称为接收者。承诺协议分为承诺阶段和打开阶段（或揭示阶段）。在承诺阶段，承诺者对一个秘密信息m做出承诺，计算得到密文形式的承诺值 c 并发送给接收者，即接收者在此阶段无法获取该密文中的任何消息，可有效保证发送者秘密信息的隐私性，这个性质也称为承诺方案的隐藏性质（Hiding）；在打开阶段，承诺者将秘密信息m以及密钥发送给接收者或公开，接收者可利用密钥验证该消息与承诺者在承诺阶段所发送消息是否一致，该阶段任何恶意的承诺者都不能将原始承诺消息m打开为另一个与实际承诺不同的值m'，并且可通过验证，这个性质也称为承诺方案的绑定性质（Binding）。承诺协议作为一个基本的密码学原语被广泛用于构造零知识证明系统以及现代密码学其他安全计算

协议领域，如抛币协议、不经意传输协议、两方计算与安全多方计算、（否认）认证协议、密钥协商协议和电子投票等各种高层协议的设计中。

零知识证明系统，特别是非交互式证明在密码学的理论研究和实践应用中都发挥了重要作用。经过国内外众多学者长期研究，已有高效的基于配对的零知识简洁非交互式知识论证（zero-knowledge Succinct Non-interactive ARguments of Knowledge，(zk)SNARKs）等人的研究成果。然而，这些 SNARKs 结构的某些方面仍然不能令人满意，其中存在最大问题且未完全解决问题是对长期受信任、结构化和电路相关参数[与电路相关的结构化参考串（Structured Reference String，SRS）]的依赖性。尽管通过构建透明参数，即在统一随机串（Uniform Random String，URS）模型中寻找替代方案以绕过受信任第三方的需求方面做出了重要的研究工作。但是，基于配对的(zk)SNARKs 的原始结构主要依赖于可信设置的可靠性，可信初始设置中似乎仍然是最实用的替代方案，因为它们的验证速度非常快，这一特性区块链应用程序中是必须的；另外，该问题的多方解决方案不能完全扩展。作为可信 SRS 模型的替代方案，Groth 等人（2018 年美密会）定义了可更新模型（Updatable Model）。在这个模型中，SRS 可以由任意一个参与方以非交互方式和可验证的方式更新，产生一个正确生成的结构化参考串，并且如果在所有的参与方中至少有一方是诚实的，那么所有各方都不知道模拟陷门。密码学有一个重要研究趋势，即提倡以模块化方式构建协议，通过将复杂的协议分解为更简单的步骤，它们变得更容易分析；此外，更加重要的是实现工作间的可比性（Comparability），特别是在零知识领域，由于大量的应用场景、实现细节、效率需求、加密假设和信任模型，导致难以实现各工作成果间的可比性。文献[1]提出一种基于配对的通用可更新（Universal and Updatable Pairing-based）(zk)SNARKs 模块化结构的代数框架。首先将先前工作的技术核心确定为可检查子空间采样（Checkable Subspace Sampling，CSS）论证的实例，这是一种新的信息论交互式证明系统。在该系统中，证明者根据验证者的硬币表明向量已在子空间中被采样。然后通过这个代数公式（Algebraic Formulation），可以立即看到 CSS 论证用作线性空间中成员资格参数的构建块。将 R1CS 约束系统简化为 3 个代数关系：内积、Hadamard 积和 CSS 参数。这种代数公式非常简洁明了，使得通用可更新 SNARKs 的进展与使用类似语言的其他工作联系起来变得更加容易。

目前，在构建非交互式不可延展承诺方面取得了令人瞩目的进展，所有方法均分为两个步骤：①基于各种亚指数困难假设，为非常小的标签/身份空间获得简单的“基本”（Base）承诺方案；②假设亚指数非交互式证据不可区分证明（Non-Interactive Witness Indistinguishable Proofs，NIWIs）和无密钥抗冲突哈希函数的变体，构建非交互式编译器，将针对较小标签空间的基于标签的不可延展性承诺转换为针对较大标签空间的基于标签的不可延展性承诺。在先前工作中，唯一无须初始设置（Setting）和使用黑盒基本方案（Goyal 等人 FOCS 2012 年）的标签扩增方法增加了多轮交互。然而构造高效的具有非延展性的承诺方案作为一个公开问

题，一直是研究者努力的方向。文献[2]首次提出了非交互式具有不可延展性的承诺黑盒构造，其中关键技术贡献是一种实现标签扩增过程所需的非交互式一致性证明的新方法。文献[2]中所构造的方案满足已知最强的不可延展性定义，即选择承诺攻击（Chosen Commitment Attack，CCA）安全性。此外，该方案不是只使用了黑盒构造且消除了先前所有工作中对亚指数 NIWIs 的依赖性，而是依赖可以基于广泛的假设（如亚指数 CDH 或 LWE）获得具有亚指数的隐式 PRGs。

文献[3]主要针对如下应用场景：两个参与方，其中 Alice 想让 Bob 相信 NP 语言中 k 个语句（$x_1, x_2, \cdots, x_k$）的真实性。最简单直接（Naive）的方法是 Alice 发送 k 语句中每个语句的证据（Witness）w_i给 Bob，Bob 来验证每个（x_i, w_i）对，该方法实现了非交互式（Non-Interactive）、可公开验证（Publicly Verifiable）的论证。然而，该方法存在明显的局限性，即实现代价非常昂贵，因为通信成本随着证据的总长度线性增长。文献[3]针对是否能以通信成本远小于$k \cdot m$（其中$m = m(|x|)$是证据的长度）非交互式地实现的 k 个 NP 语句的证明展开研究，即研究重点是如何在公共参考串（Common Reference String，CRS）模型下实现批量论证（Batch Arguments，BARG），首次实现了可公开验证的非交互式批量论证系统的构造。具体地，文献[3]所构造方案首先在基于标准加密假设的公共参考串模型中为 NP 语言提供了一个论证系统的构造，定义和构造了面向 NP 语言的双模交互式批量论证（Dual-Mode Interactive Batch Arguments），所构造的论证系统允许证明 NP 语言的多个实例，同时仅需消耗大约证明单个实例的通信和验证成本；该论证系统在 CRS 模型下实现了两种模式，即正常模式（Normal Mode）或陷门模式（Trapdoor Mode）；然后展示了如何将 Fiat-Shamir 协议的相关性-难解性（Correlation-Intractability）框架应用于此类交互论证；最后证明 Fiat-Shamir 变换应用于双模交互式批量论证仍然可靠（Sound），构造了面向 NP 语言的非交互式批量论证（Non-Interactive Batch Arguments for NP）。

理解零知识的轮回复杂性一直是一个重要的问题，尤其对于有界概率多项式（Bounded Probabilistic Polynomial，BPP）以外语言的零知识论证，在没有任何可信初始设置时，需要至少 3 条交互消息。同时，引出了一个自然的问题：什么样的零知识的有意义的松弛（Meaningful Relaxations）能以非交互式和无须初始可信设置实现？现在已经有几类 Relaxations of Zero-Knowledge，具体包括弱零知识（Weak Zero-Knowledge）、证据隐藏（Witness Hiding）、强证据不可区分（Strong Witness Indistinguishability，Strong WI）、证据不可区分（Witness Indistinguishability，WI）。然而，有意义的非交互式保密论证的已知构造中存在如下问题：①不够可靠（Sound）且依赖于非标准的困难假设（Non-Standard Hardness Assumptions）；②不提供有意义的保密性，尤其是在考虑具有唯一证据的断言（Statement）时。文献[4]引入了非交互式分布不可区分的论证（Non-Interactive Distributionally Indistinguishable Arguments，NIDI）来解决（Non-Interactive Witness Indistinguishable Proof，NIWI）证明存在的一个重大

缺陷，即在用唯一的证据（Witness）证明 NP 语言的断言时缺乏有意义的保密性。有助于解决现有非交互式论证的一些缺点，并使非交互式承诺和证明（Commit-and-Prove）等应用无须可信初始设置。

Proof-Carrying Data（PCD）用于一组不信任的参与方执行分布式无限计算，并保证每个计算的中间状态都可以被简洁的验证。PCD 概括了增量可验证的计算（Incrementally Verifiable Computation），可用于构建 SNARKs。然而，直到目前唯一已知的构建 PCD 的方法需要昂贵的 SNARKs 递归操作。于是，一个直接的公开问题是：PCD 是否可以从比 SNARKs 更弱的原始系统中构建，并使用简洁的积累方案？如果有，那么能否基于此来构造出 PCD 并提高其具体效率。文献[5]针对上述问题展开研究，而多项式承诺方案（Polynomial Commitment Schemes，PCS）因其在构建 SNARKs 中的发挥的关键作用而受到关注。Halo 利用了 Bulletproofs 内部参数的聚合特性，首次展示了一种不需 SNARKs 构建 PCD 的新方法。该构造是启发式的（Heuristic），因为它非黑盒地使用了 Fiat-Shamir 变换的具体实例。文献[5]在这种构造方法的基础上进行了扩展，其研究工作表明即使 PCS 评估证明既不简洁也不高效，也可以启发式地从任何同态多项式承诺方案（PCS）构建 PCD。事实上，Halo 方法可扩展到任何具有更一般属性的 PCS，即能够将承诺的线性组合聚合成一个新的简洁承诺，该承诺稍后可以对该线性组合开放。因此，文献[5]的研究结果给出了 SNARKs 和 PCD 的新结构，这些结构以前在文献中没有描述过，并且也可以作为未来结构的蓝图。

在文献[6]中，作者展示了一种如何在不依赖 SNARKs 的情况下获得 PCD 的方案。引入的一种弱累加实现构建一个 PCD 方案，给定任意非交互式的知识论证（如具有线性大小的论证），该论证系统具有关系拆分累加方案（Split Accumulation Schemes for Relations）。此外，还为 R1CS 构建了一个透明的非交互式知识论证，其拆分累加可以被验证，在基于离散对数难题的随机谕言模型中，所设计的构造被证明是安全的，并且通过随机谕言机启发和上述的结果，实现了 PCD 的效率提升。最后，文献[6]还为 Pedersen 承诺下的 Hadamard 积以及基于 Pedersen 承诺的简单多项式承诺方案构建了一个拆分累加方案。

Micali（FOCS 1994 年）开创性地给出了随机谕言模型（Random Oracle Model，ROM）中第一个简洁非交互论证（Succinct Non-Interactive ARGument，SNARG）。该构造结合了 PCP 和密码承诺，且有以下特征：它似乎是后量子的；它可以通过轻量级密码学启发式地实例化；它有一个透明的（Public-Coin）参数设置。但是，该构造同时存在一个明显的缺点，即参数过大（A Large Argument Size）。文献[7]中提供自 25 年前 Micali 构造引入以来首次进展，给出了一个新的构造且可以实现更小的参数。准确地说，如果每个t-Query 恶意证明者都能以最多ε的概率说服验证者，那么 ROM 模型中的 SNARG 就是(t,ε)-安全的。对于(t,ε)-安全性，即使依靠的是远远超出现有技术的推测概率证明，ROM 中所有已知的 SNARGs 构造的参数大小（包括 Micali 构造）均是$\tilde{O}((\log_2(t/\varepsilon))^2)$比特规模。在实践中，这些参数开销导致 SNARGs

比基于其他（前量子和昂贵的）工具的构造要大得多。因此，这使得许多研究者认为 ROM 模型中的 SNARGs 本质是二次的。然后，实际情况确实如此吗？文献[7]中证明了事实并非如此，并在 ROM 模型中构建了一个具有亚二次参数大小：$\tilde{O}(\log_2(t/\varepsilon)\cdot\log_2 t)$的 SNARG，该构造依赖于对 PCPs（Probabilistically Checkable Proofs）的强可靠性概念和对承诺的弱绑定性概念。

文献[8]研究了形式如$V_T^{(T)}\cdot z = s\cdot w$的（对偶）范德蒙（Vandermonde）系统何时在环$\mathcal{R}$上有一个解z，其中V_T定义为集合T的范德蒙矩阵，其中“松弛”s是解质量的度量。为此，文献[8]提出环$\mathcal{R}$上（s,t）-减法集（(s,t)-subtractive Sets）的概念及其性质，即当S是(s,t)- subtractive的，那么由任意t子集$T\subseteq S$定义的上述（对偶）范德蒙系统在$\mathcal{R}$上是可解的。一个有趣的挑战：当给定环$\mathcal{R}$时，如何在最小化（范数）s的同时找到大集合S。文献[8]通过在素数p上的环$\mathcal{R}=\mathbb{Z}[\xi_{p^\ell}]$上构造大小为$n=\mathrm{poly}(\lambda)$的$(s,t)$-减法集$S$族，基于 SIS（Short Integer Solution）关系$A\cdot x = s\cdot y \bmod q$具有$O(1/n)$知识误差的知识证明，构造了 Schnorr-like 格（若$s=1$，则$p=\mathrm{poly}(\lambda)$。该技术很自然地融入在美密会 2020 上的基于格 Bulletproof 框架中，为 NP 语言生成了基于格的简洁论证，并具有更好的参数。除此之外，文献[8]中还将(s,t)-减法集的概念将基于群的门限密码学与格设置关联起来，并通过将其与分布式伪随机函数相关联进行证明。

Sumcheck 协议是 Lund 等人在 1992 年引入的交互式证明，在复杂性理论的概率证明理论以及最近的密码学发展中发挥了基本作用；Sumcheck 协议已广泛用于一系列关于简洁论证（Succinct Arguments）的工作中，这避免了诸如在时间和内存上都是昂贵的快速傅里叶变换，该协议在其他简洁论证中是很常见的操作。另外，一系列工作成果基于离散对数设置中 Pedersen 承诺的折叠技术（Folding Techniques）（Bootle 等人于 2016 年欧密会）构建了简洁论证。非正式地，为了证明一条长消息知识打开了给定的 Pedersen 承诺，证明者与验证者进行了一次缩减，通过将消息“围绕”（Around）验证者的挑战折叠起来，从而将消息的长度减半。尽管已有众多研究工作以及应用，但 Pedersen 承诺（相关）的折叠技术并没有完全被理解。例如，在将 Fiat-Shamir 变换应用于（公共硬币）交互式论证后通常用作非交互式论证。然而，该非交互论证的安全性在随机谕言模型中只能通过超多项式时间提取器（Superpolynomial-time Extractor）或代数群模型（Algebraic Group Model）来证明。此外，几乎所有简洁的论证都是通过某种类型的概率证明获得的（并且有些设置是固有的），然而在折叠技术中没有明显的概率证明。Sumcheck 协议和折叠技术间存在差异，但同时有几个共同特点：两种协议都有一个可通过线性数量的操作进行实现的证明器，或者作为流算法；此外，两种协议都满足有助于证明安全特性的强可靠性概念。因此，引发一个思考：两种协议的相似之处仅仅是巧合吗？文献[9]引入了一类交互式协议，称为 Sumcheck 论证（Sumcheck Arguments），该论证系统可用于在 Sumcheck 协议和 Pedersen 承诺的折叠技术之间建立了一种新颖的连接。并基于模（Modules）定义了一类 Sumcheck 友好的承诺方案，这些方案覆盖

了许多有趣的例子，并且证明了 Sumcheck 协议应用于与承诺方案相关的多项式，构建了一个简洁的关于打开承诺的知识论证。文献[9]在此基础上还获得了 NP 完全语言 R1CS 在某些环上的简洁论证。Sumcheck 论证能够实现将不同密码初始设置（离散对数 Discrete Logarithms、配对 Pairings、未知阶群 Groups of Unknown Order、格 Lattices）中的许多先前研究成果作为特例进行恢复，从而提供了一般框架来理解上述先前工作成果。此外，文献[9]回答了先前工作中提出的公开问题，如从 SIS 假设中获得基于格的简洁论证，以解决环上的满足性问题。

本节作者：林昌露、黄可可（福建师范大学）

参考文献

[1] RÀFOLS C, ZAPICO A. An Algebraic Framework for Universal and Updatable SNARKs[C]. CRYPTO 2021: 774-804.

[2] GARG R, KHURANA D, LU G, et al. Black-box Non-Interactive Non-Malleable Commitments[C]. EUROCRYPT 2021: 159-185.

[3] CHOUDHURI A R, JAIN A, JIN Z. Non-Interactive Batch Arguments for NP from Standard Assumptions[C]. CRYPTO 2021: 394-423.

[4] KHURANA D. Non-Interactive Distributional Indistinguishability (NIDI) and Non-Malleable Commitments[C]. EUROCRYPT 2021: 186-215.

[5] BONEH D, DRAKE J, FISCH B, et al. Halo Infinite: Proof-Carrying Data from Additive Polynomial Commitments[C]. CRYPTO 2021: 649-680.

[6] BÜNZ B, CHIESA A, LIN W, et al. Proof-Carrying Data without Succinct Arguments[C]. CRYPTO 2021: 681-710.

[7] CHIESA A, YOGEV E. Subquadratic SNARGs in the Random Oracle Model[C]. CRYPTO 2021: 711-741.

[8] ALBRECHT M R, LAI R W F. Subtractive Sets over Cyclotomic Rings: Limits of Schnorr-Like Arguments over Lattices[C]. CRYPTO 2021: 519-548.

[9] BOOTLE J, CHIESA A, SOTIRAKI K. Sumcheck Arguments and Their Applications[C]. CRYPTO 2021: 742-773.

4 门限签名/Σ协议

文献[1]展示了首个不依赖于 PCP 的格基承诺-证明（Commit-and-Prove）透明电路零知识

（ZK）证明协议。该工作是美密会 2020 上提出的压缩 Σ 协议的扩展。压缩 Σ 协议首先使用 Bulletproof 中的递归“折叠技术”进行压缩，其通信代价为对数轮次。在 ZK 中证明一个秘密向量满足给定的约束，可通过算术秘密共享技术将其降低到线性复杂度。当在任何电路 ZK 证明之前创建对秘密向量的承诺时，常用到“承诺-证明”技术，该技术仅需要对数级别的通信代价。然而，当考虑建立低通信量 SIS 平台系统时，Σ 协议无法确定能否拥有多项式小的挑战空间。当考虑压缩步骤（非常量轮次）及并行化减少误差的必要性时，目前并没有相关的研究结果可证明：复合协议可具有有效的知识提取器。该文通过两个独立结果来回答上述问题：第一个结果展示了非常数轮与多项式小的挑战空间相结合下，有效知识提取的严格分析，而第二个结果表明并行重复确实可以快速减少知识误差。

现有的基于离散对数（DL）的多重签名方案若在实践中以 256 位组的形式实现，其方案证明的安全性仅能提供弱保证。这是因为标准模型下基于离散对数假设的安全归约，大多是松归约。文献[2]表明，若放松模型或假设要求，则均可获得相应的紧归约证明。该文给出了代数群模型中离散对数困难假设的紧归约证明，以及标准模型下除离散对数之外的假设的紧归约证明。该文首先对经典的 3 轮方案（BN 和 MuSig）提出相应的紧归约证明。然后作者给出了一个新的两轮多重签名方案 HBMS，效率上该方案与之前的方案相仿。该方案的安全证明通过链式归约框架实现，其中一个归约将被分解为一条涉及若干中间问题的子归约链。总体而言，该文的结果提高了基于离散对数困难问题的多重签名方案在实践中的安全保证。

Lai 等人（CCS 2019）展示了 Bulletproof 的算术电路零知识协议（Bootle 等人欧密会 2016 和 Büunz 等人 S&P 2018）如何直接扩展到双线性群算术电路，即无须将这些电路进行算术电路转换。简而言之，双线性群算术电路是一种标准算术电路，其增加了群幂运算或配对的特殊门。然而，使用标准算术电路表示这些特殊门会导致电路大小开销的显著增加，因此通过标准算术电路实现零知识的方法可能会产生大量的额外成本。Lai 等人通过将额外的零知识技术集成到 Bulletproof 框架中来避免这种情况，从而非常有效地处理特殊门。而文献[3]则使用了不同的方法实现对特殊门的处理：通过扩展压缩 Σ 协议理论（美密会 2020），将算术电路关系扩展到双线性群算术电路关系。除了概念上更为简洁，该方法还具有将 Lai 等人的协议的通信成本降低原来的 1/3 的优势。最后，该文展示了该研究结果的一个应用：构建了第一个 k-out-of-n 门限签名（Threshold signature，TSS），其允许透明初始化以及门限签名大小为 n 的对数级别。该门限签名方案可隐藏最多 k 个签名者的身份，且阈值 k 可在聚合阶段动态选择。

多重签名允许一组签名者能够在联合消息上进行联合签名。最近，Drijvers 等人（S&P 2019）表明，迄今为止在纯离散对数环境（无双线性对）中提出的所有两轮多重签名方案在并发签名会话下都是不安全的。同时，Drijvers 等人提出了一种安全的两轮方案，其签名效率仅为 Schnorr 签名的 2 倍多。由于该方案的实用性，其在密码系统中变得流行（如可用于比

特币）。如果需要一种多重签名方案来替代 Schnorr 签名，就不得不借助于 3 轮方案或一系列的签名会话，这两种方案在实践中都是不可取的选择。文献[4]提出了一种简单且高度实用的双轮多签名方案 MuSig2。这是第一个同时满足以下条件的多重签名方案：①在并发签名会话下是安全的；②支持密钥聚合；③输出 Schnorr 签名；④只需要两轮通信；⑤具有与普通 Schnorr 签名相似的签名者复杂性。此外，它是纯 DL 环境中的第一个多重签名方案，支持除一轮之外的所有预处理，高效地实现了非交互式签名，且不会弱化并发会话下的安全性。该文在随机谕言模型中证明 MuSig2 方案的安全性，以及在随机谕言机和代数群模型相结合的模型下，证明其变体方案的安全性。

门限签名允许 n 方共享发布数字签名的能力，从而使得群内任意 $t+1$ 位以上成员合作均可以签名，而 t 或更少参与者的团体则不能。目前已知的基于类组的门限 ECDSA 构造要么效率低下（需要将底层挑战空间很小的零知识证明并行重复），要么需要非标准的低阶假设。文献[5]提出了使用基于类组的加密方案构造的高效门限 ECDSA 协议，该协议既不使用低阶假设，也不并行重复底层零知识证明。与以前的构造相比，显著提高了密钥生成的效率。在此过程中，该文引入了 Promise Σ 协议概念，它只需满足一种称为 Promise 可提取性的弱化可靠性（Soundness）。与基于类组的加密相关的 Promise Σ 协议事实上并不能保证声明的真实性，但在该文提出的应用场景中，该协议可提供足够的安全保证（承诺可提取性）。该文进一步展示了如何在基于类组的加密上模拟同态操作。这些技术具有实际意义，且适用于其他需要对与类组相关的陈述进行有效零知识证明的场景。

在集合成员证明中，公开信息由一组元素和一个承诺组成。然后，证明者生成一个零知识证明，表明该承诺确实是集合中的某个元素。这个原语与匿名性和隐私协议的概念密切相关，如环签名和“one-out-of-many”证明等。文献[6]提出来一个新的格基集合成员证明，其大小与集合大小成对数关系。另外，该文将该集合成员证明转换为环签名方案，其环签名大小与公钥集大小也成对数关系。对于2^5个元素的集合，群签名大小为 16 KB，而对于大小为2^{25}的集合，群签名大小为 22KB。在大约 128 位的安全级别上，这些输出分别比 Beullens 等人（亚密会 2020）和 Esgin 等人（CCS 2019）的方案小 2/3 和 1/7。该文还展示了一个新的环签名方案，结合了其他一些技术和优化，可以转化为一个相当高效的基于 Esgin 等人（CCS 2019）的 MatRiCT 框架的 Monerolike 安全交易系统。借助该新技术，能够将交易证明的大小比上述工作量减少 10%~25%。

One more discrete logarithm（OMDL）困难假设是身份认证协议、盲签名和多签名方案的安全分析的基础，如 blind Schnorr 签名和最近的 MuSig2 多重签名。由于这些方案输出标准的 Schnorr 签名，因此可与现有的系统兼容（如区块链系统等）。此外，OMDL 假设可用于某些安全性归约的不可能性证明中。尽管使用广泛，但是 OMDL 缺乏严格的分析，甚至在通用群模型（Generic Group Model，GGM）中也无法给出相应证明。文献[7]在 GGM 模型中给出

了 OMDL 的形式证明，并证明了一个相关的假设，即 one-more CDH 假设。

Schnorr 身份认证和签名方案一直是过去 30 年中最有影响力的密码协议之一。尽管对这两种方案最著名的攻击是通过离散对数计算，但是已知的离散对数问题的安全性遇到了“平方根障碍”。特别是，在任何 p 阶群的离散对数问题中，Shoup 困难性结果被认为是成立的（因此常被用于设置具体的安全参数），对 Schnorr 身份认证和签名方案的最著名的 t 时间攻击的成功概率为$\frac{t^2}{p}$，而现有的安全证明抵抗成功概率仅分别为$\left(\frac{t^2}{p}\right)^{\frac{1}{2}}$和$\left(q_H\frac{t^2}{p}\right)^{\frac{1}{2}}$的攻击，其中$q_H$表示随机谕言模型中攻击者发起的哈希查询次数。文献[8]为基于 Σ 协议的身份认证和签名方案（特别是基于离散对数困难假设的 Schnorr 身份认证和签名方案）建立了更为严格的安全保证。该文通过引入经典分叉引理的高矩泛化来规避平方根障碍，其依赖于 d-moment 困难关系：任何算法在为一个随机实例生成证据的成功概率，受算法运行时间的第 d 时刻动态控制。在离散对数问题的具体背景下，Shoup 的原始证明已表明，由于离散对数问题在 GGM 模型中是 2-moment 困难的，因此该假设可以被视为在没有比通用算法更高效的群中，对离散对数假设的高度合理增强。在这种情况下，高矩分叉引理表明，假设离散对数问题是 2-moment 困难的，则在任何 t 时刻，攻击者最多以$\left(\frac{t^2}{p}\right)^{\frac{2}{3}}$的概率和$\left(q_H\frac{t^2}{p}\right)^{\frac{2}{3}}$的概率分别攻破 Schnorr 身份认证和签名方案的安全性。

文献[9]通过对每个签名者提供的两个预先承诺进行去线性化，构建了一个基于 Schnorr 的两轮签名方案（DWMS）。在代数群模型（Algebraic Group Model，AGM）和随机谕言模型（Random Oracle Model，ROM）下，DWMS 可基于 OMDL 困难假设与二纠缠和问题的困难假设来证明其安全性。该文提出的新的 m 纠缠和问题的本质是在标量域中修改 k-sum 问题。假设相关群中离散对数问题是困难的，该文证明了在 AGM 模型中这个新问题的困难性。该新困难性假设可简化使用承诺去线性化的多重签名方案的安全性证明。

本节作者：黄琼（华南农业大学）、马莎（华南农业大学）、黄建业（澳大利亚伍伦贡大学）

参考文献

[1] ATTEMA T, CRAMER R, KOHL L. A Compressed Σ-Protocol Theory for Lattices[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12826. Springer, Cham.

[2] BELLARE M, DAI W. Chain Reductions for Multi-signatures and the HBMS Scheme[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in

Computer Science, vol 13093. Springer, Cham.

[3] ATTEMA T, CRAMER R, RAMBAUD M. Compressed ς-Protocols for Bilinear Group Arithmetic Circuits and Application to Logarithmic Transparent Threshold Signatures[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093. Springer, Cham.

[4] NICK J, RUFFING T, SEURIN Y. MuSig2: Simple Two-Round Schnorr Multi-signatures[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

[5] DENG Y, MA S, ZHANG X, et al. Promise ς-Protocol: How to Construct Efficient Threshold ECDSA from Encryptions Based on Class Groups[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology - ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093. Springer, Cham.

[6] LYUBASHEVSKY V, NGUYEN N K, SEILER G. SMILE: Set Membership from Ideal Lattices with Applications to Ring Signatures and Confidential Transactions[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology - CRYPTO 2021. Lecture Notes in Computer Science, vol 12826.

[7] BAUER B, FUCHSBAUER G, PLOUVIEZ A. The One-More Discrete Logarithm Assumption in the Generic Group Model[C]. In: Tibouchi, M., Wang, H. (eds) Advances in Cryptology-ASIACRYPT 2021. Lecture Notes in Computer Science, vol 13093. Springer, Cham.

[8] ROTEM L, SEGEV G. Tighter Security for Schnorr Identification and Signatures: A High-Moment Forking Lemma forΣ-Protocols[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology-CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

[9] ALPER H K, BURDGES J. Two-Round Trip Schnorr Multi-signatures via Delinearized Witnesses[C]. In: Malkin, T., Peikert, C. (eds) Advances in Cryptology-CRYPTO 2021. Lecture Notes in Computer Science, vol 12825. Springer, Cham.

5 安全多方计算

在 2021 年的国际三大密码年会中，安全多方计算（Secure Multi-Party Computation，SMPC）仍然是研究的热点方向，其中美密会包含 2 个相关专题，收录 21 篇论文，欧密会包含 3 个相关专题，收录 12 篇论文，亚密会包含 1 个相关专题，收录 5 篇论文。从研究内容看，主要涉及 7 个方面：①电路优化；②轮复杂度及通信复杂度优化；③非交互安全多方计算协议；

④编译器构造；⑤安全性及安全模型；⑥隐私集合运算；⑦量子安全多方计算协议。

5.1　电路优化研究

混淆电路（Garbled Circuits，GC）是解决安全多方计算问题的一个通用方法，可以安全地计算任意的函数，但其通信复杂度过高以至于在实际环境中很难使用。在最初提出的混淆电路方法中，对于每个 AND 门和 XOR 门都需要 $8k$ 比特的通信，其中 k 为安全参数。如何降低混淆电路门的密钥规模是一个非常重要的研究领域。自混淆电路提出的 30 年以来，人们提出了很多降低混淆电路密钥规模的方法，Free-XOR 和 half-gates 是其中两个非常关键的技术。Free-XOR 使得在求值 XOR 门时完全不需要通信，half-gates 使得 AND 门的密钥规模降低到 $2k$，且 half-gates 技术与 Free-XOR 技术是兼容的。也就是说，结合 Free-XOR 和 half-gates，求值一个 XOR 门没有开销，求值一个 AND 门只需要 $2k$ 比特的通信。同时，在 half-gates 技术中结合了当时所有的技术提出了 linear garbling 模型，并证明了在 linear garbling 模型下，求值 AND 门的通信下界就是 $2k$ 比特。

在 2021 年美密会上 Rosulek 和 Roy[47]提出了一种全新的技术称为 slicing and dicing，并对 AND 门给出了新的优化。slicing and dicing 技术绕过了 linear garbling 模型，首次打破了 half-gates 中给出的 $2k$ 比特的下界，使得求值一个 AND 门只需要（$1.5k$+5）比特的通信，同时求值 XOR 门不需要开销。对于安全参数 k=128 的情况，相比于 half-gates 技术降低了 23% 的通信开销，这大大提升了混淆电路的实现效率。

在基于混淆电路的两方安全计算中，电路构造方通常需要将完整的混淆电路发送给电路计算方，这包括了那些没有被用到的条件分支。2020 年堆叠混淆电路（Stacked Garbling Circuit，SGC）[34-35]是一种改进的混淆电路（Garbled Circuit，GC）技术，能够减少协议的通信量，使得通信复杂度与程序最长执行路径成正比，而非整个电路规模。然而，与经典的混淆电路技术相比，堆叠混淆电路协议中参与方的计算量却有所增加。为了进一步降低计算量，文献[37]提出堆叠交错（Stack-and-Stagger）技术，在保持通信优势的同时，将混淆电路的分支总数从 $O(nk)$降到 $2n$-k，其中 n 表示电路的分支总数量，k 表示计算使用的分支数量，这显著减少了双方的计算量。在 1Gbps 的局域网环境，在总计 128 个分支中计算 16 个分支的运行时间比标准 SGC 协议大约快 7.68 倍。该技术适用于参与方知道所使用的条件分支索引的两方计算场景，可用于构造半诚实模型下的安全协议。

与文献[37]类似，另一种算法将标准 SGC 协议的计算量从 $O(b^2)$减少到 $O(b\ \log_2 b)$，其中 b 表示分支数量，而且没有增加通信量。标准 SGC 协议的计算量大都是因为茫然收集了程序中非活动部分对应的无用标签。这些无用标签是由一个多路复用器收集的，其生成成本很高。文献[36]重新设计了堆叠和无用标签收集过程，避免了平方扩展。同时，标准 SGC 算法需要 $O(b)$空间，而文献[36]的算法只需要 $O(\log_2 b)$空间。这使得即使是小规模的运行环境也能处理

大量的分支。对于少于 16 个分支的程序，运行时间与标准 SGC 相当；对于较大分支数量的程序，其性能明显优于标准 SGC，对于 1024 个分支的程序，能够比标准 SGC 快 31 倍。

自 1976 年 Valiant 给出了 2 路和 4 路通用电路（Universal Circuit，UC）的理论构造，其渐进大小分别为$5n\log_2 n$和$4.75n\log_2 n$，它们与渐近下界$\Omega(n\log_2 n)$某个常数因子相匹配。在后续不断发展优化中，Zhao 等人（亚密会 2019）提出了渐近大小为$4.5n\log_2 n$的 4 路通用电路，并证明了 Valiant 框架下 UC 的下界为$3.64n\log_2 n$。随着计算规模超过 1000 万门级（$n = 10^7$）甚至 10 亿门级（$n = 10^9$），UC 规模中的常数因子在应用程序性能中扮演着越来越重要的角色。文献[42]提出的通用电路渐进大小为$3n\log_2 n$，比已知的$4.5n\log_2 n$提高了 33%，并超过了 Zhao 等人（亚密会 2019）构造的 UC 下界$3.64n\log_2 n$。该工作表明，其新框架下的 UC 的下界为$2.95n\log_2 n$。文献还实现了 2 路通用电路，其大小为$3n\log_2 n$。

目前，MPC 的大多数应用都集中在少数参与方，这很大程度上是因为大多数已知协议在通信和计算复杂度在参与者数量上产生线性乘法开销，即复杂度为$O(n|C|)$（其中 n 是参与者的数量，$|C|$是电路的大小）。文献[14]提出了一个复杂度与参与方数量无关的大规模 MPC 协议。该工作发现在 MPC 的许多重要应用中使用的电路，如用于创建机器学习模型的训练算法，都具有高度重复的结构，因此提出了一个有意义的电路类，称为（A,B）-重复电路，其中参数 A、B 为变量。文中表明对于$\big(\Omega(n),\Omega(n)\big)$-重复电路，有效的乘法和 PSS（Packed Secret Sharing）技术可以结合，实现 n 个参与方的$O(|C|)$安全多方计算。这是第一个比 SIMD 电路更大的电路类的结构。实验结果表明，对于涉及大量参与方的计算，该协议优于最先进的已实现的 MPC 协议。

在“非诚实大多数”情况下，高效的 MPC 协议几乎都是在预处理模型中设计的，如 SPDZ 等。但是，当参与方的数量非常大时（大规模 MPC），这类协议将不再具有实用性。文献[11]提出了一种基于混淆电路的大规模多方计算协议。该协议不仅实现了主动安全，还支持 free-XOR 技术，每一方的通信复杂度为 $O(n)$，计算复杂度为 $O(1)$。在该协议构造中，使用了一个基于 LPN 的加密方案新变体，然而这使得混淆阶段的计算代价过高。为了解决这一问题，文献[11]提出了第二个协议，该协议假设至少有 n/c 个参与方是诚实的（对于一个任意的固定值 c），通过小幅度牺牲了在线阶段的效率，使得预处理阶段更加轻巧。实验表明，该协议计算阶段具有实用性，AES-128 的计算时间仅为 1.72s。此外，当参与方数目大于 100 时，该协议的性能优于现有最好协议的性能。

5.2 轮复杂度、通信复杂度优化研究

轮复杂度、通信复杂度的优化是近几年安全多方计算领域的热门研究内容，其主要关注安全多方计算协议的轮数及通信量的优化，并针对不同场景下、实现不同安全性的安全多方计算协议研究其轮复杂度、通信复杂度的相关问题。

针对基于黑盒调用的安全多方计算通用协议，文献[38]研究黑盒安全多方计算协议的最小轮复杂度问题，即协议黑盒调用简单密码学原语，并实现针对任意数量恶意敌手的安全性。在一般（朴素）模型中，之前的黑盒协议的轮数超过 15 轮，这与使用黑盒模拟器的协议的 4 轮下限值相差甚远。该工作改进并提出了以下类型的黑盒协议。①在朴素模型中实现 4 轮“pairwise MPC”，该轮最优协议允许每对有序的参与方都能使用自己的输入计算某个函数，其输出被发送给第二个参与方。该协议可以黑盒调用任何使用伪随机公钥的公钥加密方案。作为特例，可以在每对有序参与方之间实现黑盒轮最优的 OT 协议。②基于 OT 相关性的 2 轮 MPC 协议，该轮最优协议黑盒调用具有增强半诚实安全性的任意一般化的 2 轮 MPC 协议。③在朴素模型中的 5 轮 MPC 协议，该协议黑盒调用带有伪随机公钥的公钥加密方案，以及具备“半恶意”安全性的 2 轮 OT 协议。与之相类似，文献[45]还构造了针对任意函数的 3 轮安全多方计算协议，其黑盒调用一个 2 轮 OT 协议。其结论表明，针对半诚实敌手，他们的协议调用一个朴素模型中具备半诚实安全的 2 轮 OT 协议。该结论解决了基于最小假设的黑盒（半诚实）MPC 协议的轮复杂度问题，并回答了 Applebaum 等人[1]提出的一个公开问题。而针对恶意敌手，他们的协议需要调用一个在公共参考串（CRS）模型中具有恶意安全性的 2 轮 OT 协议，且该 OT 协议针对接收方需要满足自适应安全性的一种变体。

除上述基于黑盒的 MPC 通用协议之外，文献[32]研究诚实方大多数场景下高效可扩展的 MPC 协议。针对算术电路，他们在诚实方大多数场景下研究无条件安全 MPC 协议的通信、计算和轮复杂度问题，从算法和实现方面均有所改进。具体而言，针对 Damgård 和 Nielsen[26] 在半诚实模型中提出的最知名的结果，他们将其通信复杂度改进了 33%。并将其推广至恶意情况，从而得到具备恶意安全性的无条件诚实方大多数 MPC 协议。此外，对于 Damgård 和 Nielsen 协议的轮复杂度，他们的工作将其改进了 2 倍。实验表明，他们的构造与之前工作相比，运行时间有30%~50%的提高。与之类似，文献[46]研究二元域（Binary Fields）上无条件安全 MPC 协议的通信复杂度。到目前为止，还没有一个 n 方协议能实现每个门上 $O(n)$比特的通信复杂度开销。他们首次在诚实方大多数场景中针对恶意敌手给出一个无条件安全的 MPC 协议，其计算一个布尔电路仅需要每个门上 $O(n)$比特的通信复杂度摊销。与上述工作研究无条件安全 MPC 协议类似，文献[4]研究如何使用再生码（Regenerating Codes）减少 MPC 协议的轮数。他们在并行计算的摊销设置场景下，构造了一个协议能够以 $d+O(1)$轮对深度为 d 的算术电路进行茫然计算，且每次乘法需要交互 $O(n^2)$个环元素。该协议主要依赖于函数的预处理，能够抵抗腐化 $t<n/2$ 个参与方的敌手。而在此场景中所有已知的方法均需要 $\Omega(n^2)$的复杂度。

针对自适应安全的 MPC 协议，文献[8]提出一个新的框架可以实现轮数最优（2 轮）。具体地，他们提出了一个相对较弱的 OT 概念，称为接收者不经意采样的不可区分 OT（indistinguishability OT with receiver oblivious sampleability，r-iOT），并说明了该类 OT 在

CRS 模型中构建针对恶意敌手的 2 轮自适应安全 MPC 协议是足够的。他们展示了如何从 CDH、LPN 以及基于同构（Isogeny-Based）的假设构造 r-iOT 协议，并基于此给出了第一个针对恶意敌手的 2 轮自适应安全 MPC 协议。之后，他们将非同构（Non-Isogeny）结论推广至朴素模型，并在朴素模型中基于 LPN 问题构造了第一个半诚实安全的 2 轮自适应安全 MPC 协议。除此之外，他们基于同构和 LPN 的协议构造为基于 LWE 的轮最优自适应安全 MPC 提供了第一个后量子备选方案。在此过程中，他们还证明了 r-iOT 也可以推出非承诺加密（NCE），并基于同构和 LPN 给出 NCE 的第一个构造。

此外，文献[7]研究无约束（Unbounded）MPC 协议的轮数最优问题。具体地，在第一轮中，各参与方分别公布一条仅与自己输入相关的消息。在第二轮中，由任意参与方构成的子集都可以在单轮广播中安全地联合计算任意函数 f。这里的无约束指的是协议不对参与方的数量或可计算的函数的大小施以任何先验约束。他们的主要工作就是在朴素模型下基于标准 LWE 困难问题，构造了一个半诚实安全的 2 轮无约束 MPC 协议。而之前的工作依赖于双线性映射这一困难假设，因此，他们给出的协议是第一个满足后量子安全的无约束 MPC 协议。

目前，已知的具有最优腐化阈值的 n 方无条件多方计算协议在每个门上通信复杂度至少为$O(n)$。即使在次优（Sub-Optimal）的腐化情况中，是否存在每个门通信复杂度为$O(1)$的协议一直是一个公开的问题，文献[49]的工作给出了答案。该工作构造了一个基于打包秘密分享（Packed Secret Sharing）的无条件多方计算协议，计算单个算数电路每个门具有$O(1)$的摊销通信复杂度（Amortized Communication Complexity），且该协议实现了半诚实安全性和恶意安全性。在单个电路中使用打包秘密分享的技术难点是确保一批门所需的所有秘密都出现在单个打包秘密分享中，此外需要确保这些秘密保持正确的顺序。文章的主要技术贡献就是通过使用二分图中的完美匹配技术（特别是 Hall 婚配定理）进行秘密的安全置换来实现秘密保序。

针对算术电路的基于秘密分享的安全多方计算协议，首先将输入分享于各参与方，然后参与方利用自己持有的秘密份额，依次计算算术电路中的加法门、乘法门，而电路的输出也以秘密份额的方式分享于各参与方。为了提高乘法门的计算效率，Beaver 提出线上线下的计算模式，将大量的计算和交互在线下阶段执行，产生大量与输入无关的随机 Beaver 三元组，而线上阶段只需少量的计算及交互。为了实现恶意敌手模型下的安全，协议需要利用一类同态消息认证码（Message Authentication Code，MAC）方案来实现对秘密份额的认证，这类方案一般是在$\mathbb{Z}_p$上设计的。而现有的 CPU 以及 GPU 在处理算术运算时，是在$\mathbb{Z}_{2^k}$上进行的，因此$\mathbb{Z}_p$上的方案，在实现上首先需要 CPU 或 GPU 模拟$\mathbb{Z}_p$上的计算。SPD$\mathbb{Z}_{2^k}$协议设计了一个高效的$\mathbb{Z}_{2^k}$上的 MAC 方案，实现了恶意参与方大多数的情况下安全的 SMPC 协议，提高了线上协议的效率，而线下预计算 Beaver 三元组阶段，特别是$\mathbb{Z}_{2^k}$上基于格上同态加密协议生成 Beaver 三元组的过程，仍是一个待解决的问题。在这个背景下，Cheon 等人[22]提出针对格上

同态加密方案的高效$\mathbb{Z}_{2^k}$消息封装方法、一种更简单的电路层封装消息重分享协议，以及一个关于环$\mathbb{Z}[X]/\Phi_M(X)$上明文知识的更有效的零知识证明协议。将上述工具综合使用，得到一个更高效的$\mathbb{Z}_{2^k}$上安全多方计算协议，与之前的最佳结果相比，该协议的平均通信效率调高了 2.2~4.8 倍。该文的技术不仅大大提高了$\mathbb{Z}_{2^k}$上 SMPC 的效率，还提供了一个工具包，可用于在$\mathbb{Z}_{2^k}$上设计其他密码学原语。

长期以来，所有标准的安全计算方法都需要与电路大小成比例的通信量。相关度为 $\omega(1)$ 或亚线性通用安全计算协议的假设集只限于 LWE、DDH 和 DCR 的循环安全变体。文献[23]在 LPN 超多项式困难假设下，基于伪随机数相关生成器（Pseudorandom Correlation Generator，PCG）提出了一种新的同态秘密分享方案（HSS），可以应用于两方设置中的所有分层电路（布尔电路或算术电路）及多项式计算。若电路大小为 s，计算该电路时协议的总通信量为 $O\left(\frac{s}{\log_2\log_2 s}\right)$。在 LPN 的$s^{2^{k(s)}}$-困难性假设下（任何$k(s)\leqslant \log_2\log_2\frac{s}{4}$），文献[23]实现了$O\left(\frac{s}{k(s)}\right)$的通信量。其结果可以使用 GMW 编译器直接推广到恶意设置。

5.3　非交互安全多方计算研究

针对非交互安全多方计算（NIMPC）问题，文献[15]研究多方可重用（Reusable）非交互安全计算（mrNISC）协议，其本质上是一个 2 轮 MPC 协议，其中第一轮信息是对参与方的隐私输入进行可重用承诺。他们给出了一个 mrNISC 的协议构造，它能像经典的多轮 MPC 协议那样实现标准的模拟安全性。协议主要基于多项式模的 LWE 假设，以及 NC^1 中伪随机函数（PRF）的存在。考虑安全性，协议在朴素模型中可以实现半恶意安全，之后通过进一步的可信设置（这对 mrNISC 来说是不可避免的）可以实现了恶意安全性。相比之下，唯一已知的 mrNISC 构造要么使用双线性映射，要么使用强原语（如程序混淆）。使用上述 mrNISC 构造可以实现新的具有门限解密的多密钥全同态（MKFHE）方案。在 CRS 模型中针对 NC^1 构造的门限 MKFHE 方案基于多项式模的 LWE 假设及 NC^1 中的伪随机函数，而之前的构造均依赖于具有超多项式模噪比的 LWE 假设。在朴素模型中，针对 P 构造的门限层级 MKFHE 方案主要基于多项式模的 LsWE，NC^1 中的 PRF、NTRU，以及另一种基于次指数模噪比 LWE 的常数数量参与方案。而朴素模型中唯一已知的门限 MKFHE 方案在开始时限制了可以共同计算的参与方集合。

与文献[15]研究通用协议不同，文献[27]则研究针对对称函数的非交互安全 MPC 协议。具体如下。①他们构造了针对阿贝尔程序（其可以实现任何对称函数）的 NIMPC 协议，其改进了当前最优的通信复杂度。如果输入取一个阿贝尔群 G 的任意值，那么其协议实现了通信复杂度$O(|G|(\log_2|G|)^2)$，改进了 Beimel 等人[12]的$O(|G|^2n^2)$。如果参与方被限制于大小不超过 d 的子集的输入，那么其协议实现了$|G|(\log_2|G|)^2(\max\{n,d\})^{(1+o(1))t}$，其中 t 是一个腐化

阈值。该结果改进了 Beimel 等人[12]的$|G|^3(nd)^{(1+o(1))t}$，甚至改进了 Benhamouda 等人[16]的$|G|^{\log_2 n+o(1)}n$，其中$t=O(\log_2 n)$和$|G|=n^{o(1)}$。②他们首次给出针对线性分类器的 NIMPC 协议，比那些基于通用构造的协议更为高效。③他们指出在 Benhamouda 等人[16]提出的从 PSM（Private Simultaneous Messages）到 NIMPC 的转换方式中，所使用到的子协议不满足特定的安全性。他们修改了上述协议，仅在通信复杂性上增加了常数开销。此外，他们还针对指示函数（Indicator Function）设计了一个 NIMPC 协议，其具备与输入长度相关的渐进最优通信复杂度。

此外，文献[20]结合 Ad Hoc 安全计算和非交互安全多方计算（Non-Interactive Multi-Party Computation，NIMPC）提出了一种新的门限 NIMPC 协议。在门限 NIMPC 协议中，共有 n 个持有输入值的参与方和 1 个计算方，但只需要 k 个参与方参与计算即可得到正确的输出值。利用门限 NIMPC 的方法，可以解决基于混淆电路的 NIMPC 协议中计算方可能在某条线路上拥有两个密钥或不拥有密钥的问题。

5.4 编译器构造研究

在安全多方计算协议的形式化安全模型中，按敌手的攻击行为，可将敌手划分为半诚实敌手、恶意敌手及隐蔽敌手。半诚实敌手是“诚实但好奇的”，他会严格按照协议约定执行，但想要从协议交互信息中获取额外的信息。恶意敌手为了实现自己的攻击目标，会任意偏离协议的约定来参与协议。隐蔽敌手则是介于半诚实敌手和恶意敌手之间，具备按照恶意敌手行为来执行协议的意愿，但是如果他这样做了，那么他将以一定的概率被诚实的参与方发现。安全多方计算协议的设计与分析，往往先设计一个半诚实敌手模型下安全的协议，再通过某种“编译器”，将其编译到恶意敌手或隐蔽敌手模型下安全的协议。

GMW 编译器是将半诚实敌手协议转化为恶意敌手模型下的通用框架，大多数恶意敌手模型下的安全多方计算协议都是基于此编译器而构造的。与基础的半诚实协议相比，经 GMW 编译器编译后的协议所增加的计算、存储与通信开销和电路门数量呈线性关系。Boyle 等人[13]在基于预计算 Beaver 三元组类型的安全多方计算协议的场景下，提出了一种 GMW 编译器的变体，使得增加的存储与通信开销和电路门数量的对数相关，因此具备了亚线性的关系。具备这一性质的根本原因在于，系统采用了一种“星型”的秘密分享结构，而位于星型中心的参与方 D，持有 n 个随机数，而每个随机数与其他任何一个参与方的份额形成了对数据的（2,2）分享，在这种结构下，D 可以看作是一个特殊的可信方，在 D 的参与下，可以提高协议的存储与通信效率。

实现半诚实敌手到隐蔽敌手协议的编译，通常做法是提供公开可验证性，使得诚实参与方对恶意行为产生一个可公开验证证书，从而为外部方（如法官）对恶意参与方进行惩罚提供证据。以前关于公开可验证秘密（Publicly Verifiable Covert，PVC）安全的工作主要针对两

方安全计算的情况，Faust 等人[29]介绍了一种新型的无隐私输入的多方 PVC 安全编译器，支持的协议类型包括在“离线–在线”模型中设计的预处理多方计算协议。该工作基于时间锁加密方案来实现公开可验证性。时间锁加密方案可以对消息进行加密，使其只能在某个截止日期后解密。传统的恶意敌手模型下的方案，参与方需要提前对自己交互的数据进行某种承诺，之后再通过打开承诺来检验自己是否实施了恶意行为，但是这里存在一个时间差敌手可以在执行打开承诺之前终止协议，从而诚实方无法拿到敌手作恶的证据。而在 Faust 等人的工作中，是通过一种时间锁加密方案来提供对输入交互信息的承诺，这样即使敌手作恶，且不主动打开承诺，诚实方也可以通过付出一定的计算代价，在某时间后，自己打开承诺，得到敌手作恶的证据。

5.5 安全性及安全模型研究

安全计算理论研究领域中一个非常重要的课题是研究在某种特定条件下安全计算的可行性或不可行性结果。例如，在主动腐化模型下，如果参与方是诚实方占大多数的，那么安全计算可以在两轮内完成（一轮是不能完成的）。但是，现有诚实方占大多数场景下的 2 轮 MPC 协议大多或是仅能达到最弱的安全性保证，即选择性中止，或者是在两轮交互中均需要广播信道。文献[25]深入研究了 2 轮 MPC 协议在任意一轮有广播信道或无广播信道时，实现不同安全性保证的可行性问题。在该研究工作之前，学术界有以下共识：在诚实方大多数场景下，若两轮均为广播信道，输出可达性（Guaranteed Output Delivery）这一最强的安全性保证则能够达到。作者研究发现，仅在第一轮提供广播信道的情况下，输出可达性依然可以实现；如果仅在第 2 轮提供广播信道，作者通过一个新的构造实现了“可识别中止”安全性，并给出结论：仅在第 2 轮提供广播信道无法实现公平性，也就无法保证输出可达性；如果只有点对点信道，在腐化阈值$t > 1$和$t = 1$、$n = 3$的情况下，只能实现最弱的安全性保证，即选择性中止，在其余情况下，即$t = 1$和$n \geqslant 4$时，可实现输出可达性以及所有更弱的安全性保证。文献[44]考虑了抛硬币协议的公平性问题，证明了以黑盒方式使用公钥密码系统的抛硬币协议是$1/\sqrt{r}$不公平的（r为协议消息数）。此外，作者还给出如下结论：若安全函数求值功能函数f是完备的，即在f-混合模型（f-hybrid）下可以安全实现茫然传输，则最优的公平抛掷硬币在f-混合模型下是可以实现的；若f不是完备的，则以黑盒方式在f-混合模型下使用公钥加密的抛掷硬币协议至少是$1/\sqrt{r}$不公平的。

信息论安全的 MPC 协议是安全多方计算领域中的一个重要研究分支。此类协议不依赖于任何计算问题的困难性假设，一般来说协议效率会更高一些。现有的信息论安全 MPC 协议一般工作在有限域上，因为有限域中具备较好的运算性质，可以保证协议设计简单、安全性强。也有一些研究工作关注有限环$\mathbb{Z}_{2^k}$上的协议构造，在有限环$\mathbb{Z}_{2^k}$上进行协议构造的优势是有限环$\mathbb{Z}_{2^k}$上的运算和硬件具有很好的兼容性，且更加适用于二进制分解、安全比较、定点

算术的安全截断等基于二进制的协议。然而，目前对于非交换环上进行安全计算的研究工作较少。相比于普通有限环，非交换环可以帮助探索代数结构上所需要的最小假设，且更加适用于一些实际应用场景，如机器学习模型训练和推理、统计分析等计算任务。文献[28]构造了第一个高效的、无条件安全的 MPC 协议，只需对非交换环R进行黑盒访问即可。在此之前的研究成果在相同的设置下仅在被腐化方数量为常数或在进行分支程序和公式的计算时才是高效的。作者将基于对非交换环的 Shamir 秘密共享方案进行一般化处理。当环的中心包含一个集合$A=\{\alpha_0, \cdots, \alpha_n\}$，使得$\forall i \neq j$，$\alpha_i-\alpha_j \in R^*$成立时，所得到的秘密共享方案是强乘性的，可以将现有的构造在有限域上进行推广。作者所构造的 MPC 协议主要针对集合A中元素不与所有R中元素交换而是集合内部彼此交换的情况。作者构造了具有高效在线阶段和对环R黑盒访问的 MPC 协议，所构造的协议与基于电路摊销友好编码的最先进协议相比，减少约$\lceil \log_2(n+1) \rceil/2$的通信量和$2\lceil \log_2(n+1) \rceil$的计算量。

现有的 MPC 协议都需要参与方承诺在协议执行过程中不得离线。但是随着人们对 MPC 的兴趣越来越浓厚，计算任务也不可避免地变得越来越复杂，导致一个计算任务需要几个小时甚至几天才能完成。在此类场景中，需要为 MPC 构建一个动态参与模型，模型中的参与方可以根据自身情况灵活退出或（重新）加入计算。这种动态模型不仅可以促进“MPC-as-a-Service”范式，还可以将 MPC 部署到志愿者网络（如区块链）中，大大降低了隐私保护计算的门槛。文献[19]提出流动安全多方计算（Fluid MPC）的概念，在流动 MPC 中，参与方可以动态加入或退出计算。每个参与方需要承诺的最低在线时间称为流动性，以通信轮数来衡量。作者探索了多种建模方案并提出了流动安全多方计算的正式解决方案。在诚实方大多数的场景中构建了信息论安全的流动安全多方计算协议，实现了最大流动性，即参与方在进行一轮接收消息和发送消息的动作后便可退出计算。

Yao 协议[50]中所提出的混淆电路方案是最基本的密码构造之一。Lindell 和 Pinkas（Journal of Cryptography 2009）[41]在选择性敌手（敌手可以在获得混淆电路之前选择挑战输入）下对 Yao 协议进行了正式的安全性证明。针对自适应敌手下该方案的安全性，Applebaum 等人（美密会 2013）[5]根据不可压缩论证，给出该方案不能满足自适应安全性的结论；Jafargholi 和 Wichs（TCC 2017）[39]指出，在在线阶段发送输出映射关系和混淆输入可以避开 Applebaum 等人得出的负面结果，并证明了对于深度较浅的电路，该方案是自适应安全的。特别地，对于深度为δ的电路，安全性损失最多是关于δ指数级的。文献[40]指出，文献[39]中给出的安全性损失上限在某种意义上是最优的，并表明存在着深度为$\delta \in N$的布尔电路族，使得任何由任意对称加密方案构建的用于证明 Yao 混淆方案的自适应不可区分性的黑盒归约方法，必定会失去一个在$\sqrt{\delta}$上呈指数级的因子。由于不可区分性是一个比模拟性弱的安全概念，因此他们的安全损失界限也适用于自适应模拟。

近年来，一种新的安全模型得到关注，即协议中的敌手除控制被腐化方之外，还可以篡

改诚实参与方的设备。Mironov 和 Stephens-Davidowitz[43]在 2015 年的欧密会上提出了反向防火墙（Reverse Firewall，RF）的思想，以应对诚实方设备受到腐化的情况。直观上，参与方 $\mathcal{P}$ 的 RF 是一个位于 $\mathcal{P}$ 和外部世界之间的外部实体，其作用是在 $\mathcal{P}$ 的计算机遭到破坏时对其接收或传出的消息进行净化。此外，作者还构建了一个半诚实安全的两方计算协议。在 2020 年的美密会上，Chakraborty、Dziembowski 和 Nielsen[18]利用防火墙构建了一个安全计算协议，将文献[43]中的工作扩展到多方计算场景，并考虑了静态敌手下的恶意安全性。文献[21]对自适应敌手下的 MPC 反向防火墙进行了初步研究，并提出了反向防火墙环境下自适应安全 MPC 的定义，探讨了各种安全概念之间的关系，然后在更强的自适应安全模型下为 MPC 协议构建反向防火墙。该工作还构造了用于自适应安全的增强硬币投掷协议和自适应安全零知识协议的反向防火墙，并在没有可信 setup 的情况下构造了一个 RF 场景下的常数轮自适应安全的 MPC 协议。

在安全多方计算协议安全性的研究中，公平性是最难实现的一个安全目标，在基于理想现实模拟范例的安全多方计算安全模型中，由于通信信道异步传输的特性，即使在理想世界中的协议，也无法达到绝对的公平。已有结论表明，在恶意大多数的 n 方掷币协议中，不能达到强公平性，也就是要满足恶意参与方不能使掷币结果产生任何偏差。在强公平性无法实现的情况下，研究者提出了弱公平性的概念，当协议出现偏差时，诚实方可以持有足够的数据，通过自己额外的计算来获得相对公平，或者持有证据并向一个权威机构申诉，通过外力来实现相对公平。在一个弱公平的协议中，往往需要多轮执行以逐步实现公平，如何减少交互轮数，就成为一个研究问题。现有的民俗锦标赛树协议，可以在与参与方数量的对数相关的交互轮复杂度下，实现基于博弈论的公平性。Chung 等人[17]在 n 方领导人选举协议的背景下，构造了一个亚对数交互轮次的协议，该协议是在公平性程度与协议交互轮次之间的一种平衡。Chung 等人首先给出了一个成为顺序近似公平性（Sequential Approximate Fairness）的公平性近似定义，通过一个与交互轮次相关的参数，定义了公平性与博弈论公平性的近似程度，然后证明了在一个 n 方领导人选举协议中，可以在r轮中实现$(1-\frac{1}{2^{\Theta(r)}})$-公平，其中 $\Theta(\log_2\log_2 n)\leqslant r\leqslant\Theta(\log_2 n)$。同时，Chung 等人证明了一个下界，表明如果使用与锦标赛树结构相似的协议，要达到“完美”的博弈论公平性，那么交互轮次的下界就是参与方的对数。

由于现实中对有状态环境进行长期维护是非常困难的，因而催生了无服务器计算范式。无服务器计算范式中大部分无状态的组件按需部署以处理计算任务，并在任务完成后被拆除。由这些组件构建的协议，能够通过隐藏参与协议的物理机器，来提高抵抗拒绝服务攻击的能力。为实现这样的保护，协议需要使用无状态的参与方，每个参与方仅发送一条消息，且不再发送其他消息。文献[30]对上述模式的协议进行了研究，并称这种模式为“You-Only-Speak-Once”。作者给出了形式化定义和研究此类协议的形式化模型，称为 YOSO 模型。作者将无

状态的参与方称为 roles，roles 在被销毁之前只能发送一条消息，是执行协议操作并与其他 roles 通信的抽象形式的实体。作者描述了两种不同的安全 MPC 协议，两个协议在假定诚实方占大多数时都能够实现输出可达性。

2015 年 Garg 等人[31]关注了如下问题：发送方能否将一对消息(m_0, m_1)进行编码，并且通过二进制擦除通道（Binary Erasure Channel）进行发送，使得接收方仅能够准确解码出两条消息中的一条，同时发送方却不知道是哪条消息？Garg 等人[31]的研究成果表明，这在信息论上是不可能实现的。然而在文献[6]中，作者通过假设接收方计算能力受限，允许存在逆多项式统计安全误差（Inverse-Polynomial Statistical Security Error），并依靠 Ideal Obfuscation，可以规避这种不可能性。在文中给出的构造中，密码原语 Ideal Obfuscation 既可以通过使用（无状态的）防篡改硬件直接实现，也可以使用现有的不可区分混淆方案（Indistinguishability Obfuscation）在平凡模型中进行实例化。

5.6 隐私集合运算研究

隐私集合求交（Private Set Intersection，PSI）是隐私集合运算的主流工作，指的是每个参与方拥有自己的私有集合，目的是联合计算这些私有集合的交集，但不会泄露交集之外的任何私有信息。PSI 在各种隐私保护大数据应用中可作为核心组件，如社交网络联系人发现、人类基因组医学研究、政府机构嫌疑人检测等。

文献[48]构造了一种新的可批处理的不经意伪随机函数（Oblivious Pseudo-Random Function，OPRF），该构造基于向量不经意线性函数计算（Vector Oblivious Linear-Function Evaluation，VOLE）和字符串探测与异或（Probe and XOR of Strings，PaXoS）结构，并且从该 OPRF 得到了一个新的 PSI。其中，VOLE 是一个两方协议，允许参与方随机采样向量$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{F}^m$及$\Delta \in \mathbb{F}$，其中$\mathbb{F}$表示一个有限域，$m$ 是一个正整数，使得$\boldsymbol{C} = \Delta \boldsymbol{A} + \boldsymbol{B}$，并且最终接收方持有$\boldsymbol{A}$和$\boldsymbol{C}$，而发送方持有$\boldsymbol{B}$和$\Delta$；PaXoS 是一个高效的线性方程组求解器，并将其看作一个数据结构。进一步，文献[48]对 PaXoS 进行修改，给出一个在某些约束下均匀分布的扩展，称为 XoPaXoS，构造了一个不经意可编程伪随机函数（Oblivious Programmable Pseudo-Random Function），以此为基础，允许 PSI 的输出由双方秘密共享，并执行其他的安全多方计算（Secure Multi-Party Computation）协议，构造了一个新的 Circuit PSI，其通信复杂度和计算复杂度均为$O(n)$，可以实现恶意模型下的安全性，并且当输入集合大小为$n = 2^{20}$时，该协议只需要 6.2 秒的运行时间和少于 52MB 的通信开销。

文献[33]将类似于不经意多项式估值（Oblivious Polynomial Evaluation，OPE）的性质抽象化，定义了不经意密钥值存储（Oblivious Key-Value Store，OKVS），将其表示为一种期望映射关系：$k_i \to v_i$。若v_i是随机值，则 OKVS 数据结构可以隐藏k_i，然后使用该结构替代以往 PSI 方案中可抽象化为 OKVS 的组件，得到更加高效的 PSI 协议。文献[33]给出了两种

OKVS 的具体构造，分别是基于 3-哈希混淆布谷鸟哈希表的方案和基于简单哈希及稠密矩阵的方案，这两种 OKVS 构造均具有线性复杂度。为了更容易得到 OKVS 参数以保证 OKVS 有一个较低的错误概率上界，文献[33]给出了将一个失败概率为p的较弱的随机 OKVS 结构，扩展到一个具有相似开销，失败概率降低为p^c的更强的 OKVS 技术，其中 c 为大于 1 的常数。使用文献[33]中的 OKVS 构造替换目前最好的 PSI 及其变体中的可以抽象为 OKVS 的组件，如 OPE 或特定数据结构，就可以得到更好的 PSI。具体而言，目前最好的 PSI 是 2020 年 Pinkas 等人在欧密会上提出的 PaXoS-PSI，若将其中的 PaXoS 组件使用文献[33]中的 OKVS 构造替代，在百万个元素的数据集下，减少了 40%的通信开销，并且运行效率提升了 20%~40%。

5.7 量子安全多方计算协议研究

在现有的量子安全多方计算领域（指基于量子计算的安全多方计算协议，而非基于经典计算，但抵抗量子计算敌手攻击的安全多方计算协议），已有的结论表明，如果量子茫然传输协议存在，意味着量子安全两方计算协议及多方计算协议的存在。Bartusek 等人[9]证明了量子茫然传输协议可以通过黑盒调用任意一个统计绑定、量子计算隐藏的承诺协议来构造，而这样的承诺协议可以通过黑盒调用一个量子计算困难的单向函数来构造。因此，在 Bartusek 等人的工作理论上证明了量子计算困难的单向函数的存在性，意味着量子安全多方计算协议的存在性。Bartusek 等人首先给出了一个量子比特承诺协议，该协议具备可抽取性及多义性，然后基于此承诺协议，给出了量子茫然传输协议。

文献[3,10]主要研究安全量子计算（Secure Quantum Computation，SQC）的轮复杂度问题。文献[3]首次在 CRS 模型中，针对两方量子计算（2PQC）和多方量子计算（MPQC）场景，构造了抵抗恶意敌手的常数轮安全量子计算协议。具体结论如下。①假设两消息 OT 存在，可以得到两消息的 2PQC，以及 5 轮的 MPQC 协议，其中只有 3 轮需要在线通信。上述 OT 可以基于量子困难的 LWE 问题（QLWE）构造得到。②假设 QLWE 问题的亚指数困难性，可以得到需要 2 轮在线的 3 轮 2PQC 协议，以及需要 2 轮在线的 4 轮 MPQC 协议。③当只有一方（两方中的一方）接收输出时，可以使用两消息 OT 实现最少的交互（两条消息）；一般来说，此类协议被称为非交互式安全计算（NISC），我们的结果实现了第一个恶意安全的量子 NISC。文献[3]提出了第一个允许可识别中止安全性的量子协议，其允许诚实方在中止的情况下就被腐化方的身份达成一致。而且上述协议首次实现所需量子通信轮数与电路复杂度独立。此外，他们还证明如果存在一个后量子安全经典协议，那么其轮复杂度与电路复杂度无关，那么他们构造的协议也具有此属性。其协议在具有对数深度解密电路的经典抗量子全同态加密方案的存在假设下是安全的，他们的构造还可以实现从量子公平安全计算归约到经典公平安全计算。

在恶意量子敌手模型下，文献[2]对安全多方计算中一些经典的、平凡模型中的功能函数

进行了研究。尽管使用现有的技术可以很容易地构造出一个多项式轮的协议，但是作者在该文中给出了一个常数轮后量子多方计算的构造。作者采用以下两种假设：LWE 的温和超多项式（Mildly Super-Polynomial）量子困难性、基于 LWE 循环（Circular）安全假设的多项式量子困难性，并提出了以下密码学原语。①基于 LWE 变体的量子困难性，提出通过量子电路实现关系可计算的幽灵加密方案（Spooky Encryption），将使用经典密钥的全同态加密首次实现了具有量子多密钥全同态加密性质的方案；②基于幽灵加密方案构造出具有常数轮零知识协议，可以抵御多个并行的量子验证者；③提出直线非黑盒模拟技术来抵抗并行的验证者，这里的验证者不会克隆敌手状态；④基于 LWE 的温和超多项式量子困难性，提出具有常数轮的后量子非延展性承诺方案。

2020 年，Dulek 等人[24]提出了可计算任何量子电路的安全协议，该协议在不具备诚实大多数的条件下仍然是安全的。然而，该协议容易受到拒绝服务攻击，甚至允许单个的被腐化方强制中止协议。文献[3]提出了第一个允许可识别中止安全的（Security-with-Identifiable-Abort）量子协议，当协议终止时诚实方能就被腐化方的身份达成一致。此外，当量子通信为必需时该协议执行的轮数与电路复杂度无关，是第一个具备该性质的协议。如果存在一个后量子安全经典协议，那么其轮复杂度与电路复杂度无关，作者设计的协议具有同样的性质。在具有对数深度解密电路的经典抗量子全同态加密方案的假设下，协议证明是安全的。作者所提出的构造接受从量子公平安全计算到经典公平安全计算的归约。

本节作者：王皓（山东师范大学）、蒋瀚（山东大学）、赵川（济南大学）、魏晓超（山东师范大学）、王丽萍、孙思维、范良（中国科学院信息工程研究所）

参考文献

[1] APPLEBAUM B, BRAKERSKI Z, GARG S, et al. Separating Two-Round Secure Computation From Oblivious Transfer[C]. ITCS 2020: 71:1-71:18.

[2] AGARWAL A, BARTUSEK J, GOYAL V, et al. Post-Quantum Multi-Party Computation[C]. EUROCRYPT 2021 (1): 435-464.

[3] ALON B, CHUNG H, CHUNG K-M, et al. Round Efficient Secure Multiparty Quantum Computation with Identifiable Abort[C]. CRYPTO 2021 (1): 436-466.

[4] ABSPOEL M, CRAMER R, ESCUDERO D, et al. Improved single-round secure multiplication using regenerating codes[C]. ASIACRYPT 2021 (2): 222-244.

[5] APPLEBAUM B, ISHAI Y, KUSHILEVITZ E, et al. Encoding Functions with Constant Online Rate or How to Compress Garbled Circuits Keys[C]. CRYPTO 2013 (2): 166-184.

[6] AGRAWAL S, ISHAI Y, KUSHILEVITZ E, et al. Secure Computation from One-Way

Noisy Communication, or: Anti-correlation via Anti-concentration[C]. CRYPTO 2021 (2): 124-154.

[7] ANANTH P, JAIN A, JIN Z, et al.Unbounded Multi-party Computation from Learning with Errors[C]. EUROCRYPT2021 (2): 754-781.

[8] ALAMATI N, MONTGOMERY H, PATRANABIS S, et al. Two-Round Adaptively Secure MPC from Isogenies, LPN, or CDH[C]. ASIACRYPT 2021 (2): 305-334.

[9] BARTUSEK J, COLADANGELO A, KHURANA D, et al. One-Way Functions Imply Secure Computation in a Quantum World[C]. CRYPTO 2021 (1): 467-496.

[10] BARTUSEK J, COLADANGELO A, KHURANA D, et al. On the Round Complexity of Secure Quantum Computation[C]. CRYPTO 2021 (1): 406-435.

[11] BEN-EFRAIM A, CONG K, OMRI E, et al. Large Scale, Actively Secure Computation from LPN and Free-XOR Garbled Circuits[C]. EUROCRYPT 2021 (3): 33-63.

[12]BEIMEL A, GABIZON A, ISHAI Y, et al Non-Interactive Secure Multiparty Computation[C]. CRYPTO 2014 (2): 387-404.

[13] BOYLE E, GILBOA N, ISHAI Y, et al. Sublinear GMW-Style Compiler for MPC with Preprocessing[C]. CRYPTO 2021 (2): 457-485.

[14] BECK G, GOEL A, JAIN A, et al. Order-C Secure Multiparty Computation for Highly Repetitive Circuits[C]. EUROCRYPT 2021: 663-693.

[15] BENHAMOUDA F, JAIN A, KOMARGODSKI I, et al. Multiparty Reusable Non-interactive Secure Computation from LWE[C]. EUROCRYPT 2021 (2): 724-753.

[16] BENHAMOUDA F, KRAWCZYK H, RABIN T. Robust Non-interactive Multiparty Computation Against Constant-Size Collusion[C]. CRYPTO 2017 (1): 391-419.

[17] CHUNG K-M, HUBERT CHAN T-H, WEN T, et al. Game-Theoretic Fairness Meets Multi-party Protocols: The Case of Leader Election[C]. CRYPTO 2021 (2): 3-32.

[18] CHAKRABORTY S, DZIEMBOWSKI S, NIELSEN J B. Reverse Firewalls for Actively Secure MPCs[C]. CRYPTO 2020 (2): 732-762.

[19] CHOUDHURI A R, GOEL A, GREEN M, et al. Fluid MPC: Secure Multiparty Computation with Dynamic Participants[C]. CRYPTO 2021 (2): 94-123.

[20] CIAMPI M, GOYAL V, OSTROVSKY R. Threshold Garbled Circuits and Ad Hoc Secure Computation[C]. EUROCRYPT 2021 (3): 64-93.

[21] CHAKRABORTY S, GANESH C, PANCHOLI M, et al. Reverse Firewalls for Adaptively Secure MPC Without Setup[C]. ASIACRYPT 2021 (2): 335-364.

[22] CHEON J H, KIM D, LEE K. MHz2k: MPC from HE over with New Packing, Simpler Reshare, and Better ZKP[C]. CRYPTO 2021 (2): 426-456.

[23] COUTEAU G, MEYER P. Breaking the Circuit Size Barrier for Secure Computation Under Quasi-Polynomial LPN[C]. EUROCRYPT 2021: 842-870.

[24] DULEK Y, GRILO A B, JEFFERY S, et al. Secure Multi-party Quantum Computation with a Dishonest Majority[C]. EUROCRYPT 2020 (3): 729-758.

[25] DAMGÅRD I, MAGRI B, RAVI D, et al. Broadcast-Optimal Two Round MPC with an Honest Majority[C]. CRYPTO 2021 (2): 155-184.

[26] DAMGÅRD I, NIELSEN J B. Scalable and Unconditionally Secure Multiparty Computation[C]. CRYPTO 2007: 572-590.

[27] ERIGUCHI R, OHARA K, YAMADA S, et al. Non-interactive Secure Multiparty Computation for Symmetric Functions, Revisited: More Efficient Constructions and Extensions[C]. CRYPTO 2021 (2): 305-334.

[28] ESCUDERO D, SORIA-VAZQUEZ E. Efficient Information-Theoretic Multi-party Computation over Non-commutative Rings[C]. CRYPTO 2021 (2): 335-364.

[29] FAUST S, HAZAY C, KRETZLER D, et al. Generic Compiler for Publicly Verifiable Covert Multi-Party Computation[C]. EUROCRYPT 2021 (2): 782-811.

[30] GENTRY C, HALEVI S, KRAWCZYK H, et al. YOSO: You Only Speak Once-Secure MPC with Stateless Ephemeral Roles[C]. CRYPTO 2021 (2): 64-93.

[31]GARG S, ISHAI Y, KUSHILEVITZ E, et al. Cryptography with One-Way Communication[C]. CRYPTO 2015(2): 191-208.

[32] GOYAL V, LI H, OSTROVSKY R, et al. ATLAS: Efficient and Scalable MPC in the Honest Majority Setting[C]. CRYPTO 2021 (2): 244-274.

[33] GARIMELLA G, PINKAS B, Rosulek M, et al. Oblivious Key-Value Stores and Amplification for Private Set Intersection[C]. CRYPTO 2021 (2): 395-425.

[34] HEATH D, KOLESNIKOV V. Stacked Garbling for Disjunctive Zero-Knowledge Proofs[C]. EUROCRYPT 2020 (3): 569-598.

[35] HEATH D, KOLESNIKOV V. Stacked Garbling - Garbled Circuit Proportional to Longest Execution Path[C]. CRYPTO 2020 (2): 763-792.

[36]HEATH D, KOLESNIKOV V. LogStack: Stacked Garbling with O(b log b) Computation[C]. EUROCRYPT 2021(3): 3-32.

[37] HEATH D, KOLESNIKOV V, PECENY S. Garbling, Stacked and Staggered-Faster k-out-of-n Garbled Function Evaluation[C]. ASIACRYPT 2021 (2): 245-274.

[38] ISHAI Y, KHURANA D, SAHAI A, et al. On the Round Complexity of Black-Box Secure MPC[C]. CRYPTO 2021 (2): 214-243.

[39] JAFARGHOLI Z, SCAFURO A, WICHS D. Adaptively Indistinguishable Garbled Circuits[C]. TCC 2017 (2): 40-71.

[40] KAMATH C, KLEIN K, PIETRZAK K, et al. Limits on the Adaptive Security of Yao's Garbling[C]. CRYPTO 2021 (2): 486-515.

[41] LINDELL Y, PINKAS B. A Proof of Security of Yao's Protocol for Two-Party Computation[J]. Cryptol. 22(2): 161-188 (2009).

[42] LIU H, YU Y, ZHAO S, et al. Pushing the Limits of Valiant's Universal Circuits: Simpler, Tighter and More Compact[C]. CRYPTO 2021: 94-124.

[43] MIRONOV I, STEPHENS-DAVIDOWITZ N. Cryptographic Reverse Firewalls[C]. EUROCRYPT 2015 (2): 657-686.

[44] MAJI H K, WANG M. Computational Hardness of Optimal Fair Computation: Beyond Minicrypt[C]. CRYPTO 2021 (2): 33-63.

[45] PATRA A, SRINIVASAN A. Three-Round Secure Multiparty Computation from Black-Box Two-Round Oblivious Transfer[C]. CRYPTO 2021 (2): 185-213.

[46] POLYCHRONIADOU A, SONG Y. Constant-Overhead Unconditionally Secure Multiparty Computation Over Binary Fields[C]. EUROCRYPT 2021 (2): 812-841.

[47] ROSULEK M, ROY L. Three Halves Make a Whole? Beating the Half-Gates Lower Bound for Garbled Circuits[C]. CRYPTO 2021: 94-124.

[48] RINDAL P, SCHOPPMANN P. VOLE-PSI: Fast OPRF and Circuit-PSI from Vector-OLE[C]. EUROCRYPT 2021 (2): 901-930.

[49] GOYAL V, POLYCHRONIADOU A, SONG Y. Unconditional Communication-Efficient MPC via Hall's Marriage Theorem[C]. CRYPTO 2021 (2): 275-304.

[50] YAO A C. How to Generate and Exchange Secrets (Extended Abstract) [C]. FOCS 1986: 162-167.

6 秘密共享/不经意传输

基于环$Z/P^{\lambda}Z$上的 Shamir 秘密共享，已有的文献构造了环$Z/P^{\lambda}Z$上信息论意义下安全的多方计算（Multiparty Computation，MPC）协议。然而，①离线乘法门在参与方数量上具有超线性的通信复杂度；②在环$Z/2^{\lambda}Z$上时，份额规模加倍，这是因为自正交码从域到环的提升是不可行的；③由于缺乏强乘法，BGW 模型无法使用此前给出的秘密共享。Cramer、Rambaud 和邢朝平在文献[1]中利用环上曲线提升的存在性与代数几何码，克服了以上所有的缺点。一方面，构造了具有强乘法的算术秘密共享方案，这是 BGW 模型中最重要的原语；

另一方面，将 Reverse Multiplication Friendly Embeddings（RMFE）从域提升到环上，且保持了域上原有的（线性）复杂度。特别是，RMFE 已成为 MPC 中针对同一电路通信复杂度的标准摊销（Amortization）技术，且文献[1]在$Z/2^{\lambda}Z$上实现了与域上 MPC 相同的复杂度。

Boyle 等人（TCC 2019）在基于函数秘密共享（Function Secret Sharing，FSS）的预处理模型中提出一种新的安全计算方法，其中门 g 是 FSS 方案的相关偏移族$g_r(x) = g(x + r)$。他们进一步提出了基于任何伪随机生成器（Pseudorandom Generator，PRG）的有效 FSS 方案，用于在“混合模式”安全计算中需求计算门的偏移系列，包括用于零测试、整数比较、ReLU 和样条函数的门。与基于混淆电路或秘密共享的技术相比，基于 FSS 的方法显著减小了在线通信和轮复杂度。文献[2]改进和扩展了 Boyle 等人的结果，做出以下 3 个贡献：①降低了分布式比较函数（Distributed Comparison Function，DCF）的密钥规模和实现门所需 DCF 的数量，从而降低了 FSS 方案的密钥规模；②提出了第一个基于 PRG 的 FSS 方案，可用于算术和逻辑移位门（Arithmetic and Logic Shit Gate），以及在Z_{2^n}上共享输入和输出的位分解，这些门对于与定点算术和机器学习相关的许多应用至关重要；③通过有序调用 FSS 方案进行乘法和移位，实现了对“multiply-then-truncate”的 2 轮基于 PRG 的安全赋值，这也表明单个 FSS 方案实现一轮基于 PRG 的安全计算将需要解决 FSS 的一个公开问题，即适用于比特位联合函数（Bit-Conjunction Functions）类的基于 PRG 的 FSS。

秘密共享方案的授权结构可以用单调函数$f:\{0,1\}^n \mapsto \{0,1\}$来表示。文献[3]聚焦于所有最小项规模为$a$的单调函数，以及它们的对偶函数（最大项规模为$b$的单调函数），分别称为$(a,n)$-upslices 和$(b,n)$-downslices。这些函数族分别对应于单调$a$-正则 DNF 和单调$(n-b)$-正则 CNF。文献[3]的主要贡献如下：①对于任意的 Downslice，其总份额规模为$1.5^{n+O(n)} < 2^{0.585n}$。由于每个单调函数都可以分解为 n 个 Downslices，因此一般授权结构也有类似的结果，从而改进了 Applebaum 等人（STOC 2020）研究结果的$2^{0.637n+O(n)}$复杂度；②考虑单调 DNF 上的一般分布F：对于每个宽度值$a \in [n]$，均匀采样k_a个规模为a的单调项，其中$\boldsymbol{k} = (k_1, k_2, \cdots, k_n)$是非负整数的任意向量。除了指数级小概率，$F$可以被$2^{0.5n+O(n)}$的总份额规模来实现，并且可以被严格小于 2/3 的指数线性实现，这为“指数级难度”的授权结构提供了候选分布。

考虑如下应用场景：Alice 使用(t,n)秘密共享方案将秘密数据存储在n个服务器上。Trudy 想要获得 Alice 的秘密数据，并愿意为此付费。Trudy 在互联网上发布广告来阐述其精心设计的加密方案，并指出它可用于从n个服务器收集份额，并保证每个提交其份额的服务器都会获得巨额的金钱奖励，且不会因违反与 Alice 的服务协议而被法律制裁。Bob 是服务器之一，在仔细检查 Trudy 的加密方案后，确定 Alice 无法证明他提交了份额。此时，Bob 可能使用广告中的加密方案，将其份额提交给 Trudy。针对该应用场景，文献[4]提出了可追溯秘密共享（Traceable Secret Sharing），该新密码原语能够提供作弊服务器不诚实行为的有效证据，致使作弊服务器总是可以被追溯和制裁的。文中给出了相关定义，并展示了它们是如何构建起来

的；在存在安全两方计算协议的前提下，构造了有效的可追溯秘密共享；讨论了该原语在多服务器计算委托的可追溯协议中的应用。

文献[5]研究了两类多项式秘密共享方案：①每个授权集都是使用多项式来重构秘密的方案；②秘密的分发与重构都是利用多项式的方案。文献[5]指出，对于线性秘密共享方案来说，具有线性分发的方案和具有线性重构的方案是等价的，而对于多项式秘密共享方案来说，具有多项式分发的方案可能比具有多项式重构的方案更强。文献[5]还证明了具有多项式重构的方案份额规模的下界；构造了具有二次分发和重构的秘密共享方案与有条件秘密披露（Conditional Disclosure of Secrets，CDS）协议；推广了 Liu 等人（美密会 2017）的构造方法，构建了最优二次k-服务器 CDS 协议；研究了如何将构建的二次k-服务器 CDS 协议转换为鲁棒 CDS 协议，并使用鲁棒 CDS 协议来构造份额规模为$O(2^{0.705n})$的具有任意授权结构的二次秘密共享方案，这比线性秘密共享方案的最优份额规模$O(2^{0.7576n})$好，比已知的一般秘密共享方案的最小份额规模$O(2^{0.585n})$差。

文献[6]给出了解决 Paillier 群中分布式离散对数问题的简单方法，它允许两方在本地将秘密的乘法份额（在指数中）转换为加法份额。与其他具有逆多项式错误概率的方法不同，该算法是完善正确的。文献[6]的主要贡献如下：①为具有可忽略的正确性误差且支持指数级明文的分支程序构造了同态秘密共享，其安全性取决于决策合数剩余（Decisional Composite Residuosity，DCR）假设；②基于二次剩余（Quadratic Residuosity，QR）或 DCR 假设，为不经意传输（Oblivious Transfer，OT）和向量不经意线性计算（Vector Oblivious Linear Evaluation，VOLE）相关性构造了伪随机相关函数（Pseudorandom Correlation Function，PCF）。基于 DCR（或 QR）和带噪声假设的学习奇偶性的组合，为一般 2 阶相关性（包括 OLE）构造了伪随机相关生成器；③进一步升级所构造的 PCF 使其具有公钥设置，在独立发布公钥后，每方都可以在本地获得其 PCF 密钥。这允许基于 QR、DCR、CRS 和随机谕言机，在完全静默的场景下生成任意数量的 OT 或 VOLE，无须 PKI 之外的任何交互。这里的公钥设置是基于一种新的非交互式矢量 OLE 协议，可以看作是 Bellare-Micali 不经意传输协议的变型。

函数f的同态秘密共享（Homomorphic Secret Sharing，HSS）允许输入方为其私有输入分配份额，进而在本地计算输出份额，并从中恢复f的函数值。HSS 可直接用于非门限敌手结构的两轮多方计算协议，并且其通信复杂度与f的规模无关。文献[7]构造了两种 HSS 方案，支持对单个低次多项式实施并行赋值，能够容忍多方和一般敌手结构。在单个赋值的特定情况下，这里的多方方案与其他多方方案相比，能够容忍更广泛的敌手结构，并且与一般结构相比具有指数级更小的份额规模。虽然限制了可容忍敌手结构的范围，但方案执行ℓ个并行赋值，其通信复杂度大约比仅使用ℓ个独立赋值小$\log_2 \ell / \ell$。考虑到其他门限方案不适用的场景，文献[7]还形式化了两类敌手结构，然后执行$O(m)$个并行赋值，其通信成本几乎与单个赋值相同，其中m是参与方的数量。

文献[8]构造了第一个同时满足如下 3 个性质的同态秘密共享（HSS）：①具有可忽略的正确性误差；②支持指数级的整数；③依赖于一个未知的隐含 FHE 的假设。具体来说，决策合数剩余（Decisional Composite Residuosity）假设。这解决了 Boyle、Gilboa 和 Ishai（美密会 2016）提出的公开问题。除了密文在两个非共谋参与者之间共享，同态秘密共享与全同态加密类似。此前的 HSS 要么具有不可忽略的正确性误差和多项式规模的明文空间，要么基于更强的 LWE 假设。文献[8]还给出了所构造方案的两个应用：具有恒定带宽开销的两台服务器 ORAM，以及错误率可忽略的 rate-1 陷门哈希函数。

文献[9]提出新的不经意传输扩展协议和向量 OLE 协议，称为 Silver，用于静默向量不经意线性赋值（Silent VOLE）和不经意传输。Silver 具有极高的性能：在标准笔记本电脑的一个内核上生成 1000 万个随机 OT 只需要 300ms 和 122KB 的通信。与标准 IKNP 协议相比，计算量减少了 37%，通信量减少了 1/1300。与 Yang 等人（CCS 2020）的协议相比，计算量减少了 1/4，通信量减少了 1/14。Silver 是静默的：在一次廉价交互之后，两方可以存储小种子，然后在本地生成大量 OT，同时保持离线，而 IKNP 和 Yang 等人的协议都不具有这一特性。与 Boyle 等人（CCS 2019）最著名的静默 OT 扩展协议相比，Silver 的计算量减少了 1/19，并且通信相同。Silver 大幅度改进了众多 MPC 协议的效率，其方法的本质归结为构造具有大的极小距离和极低编码时间的新线性码族，与构造 MPC 协议的标准范式大相径庭。

考虑攻击者可以得到每个份额泄露的任意m比特的应用场景。对于常数m，文献[11]指出对应于维数为k的随机线性码（在足够大的素数域上）的 Massey 秘密共享方案是局部抗泄露的，其中$k/n > 1/2$，而 Benhamouda 等人（美密会 2018）的构造需要满足$k/n > 0.907$。由于所有可能m比特局部泄露函数的数量是随机线性码数量的指数倍数，因此文献[10]的技术创新地从确定一个适当的伪随机性启发的测试系列开始，通过它们就足以确保抗泄露性。研究表明，大多数线性码能够通过所有测试。这种具有局部抗泄露性的线性秘密共享方案的蒙特卡罗构造方法可应用于抗泄露安全计算。Benhamouda 等人引入了一种分析指标来研究秘密共享方案的抗泄露性；当指标很小时，该方案具有抗泄露性。然而，文献[10]给出了 1 比特位局部泄露函数，证明了其逆命题是错误的，这就需要新的函数来更准确地分析抗泄露性。

文献[11]研究了素数域F上 Shamir 秘密共享方案的抗泄露性，其中敌手可以从每个份额中独立获得m比特。对于任意的重建门限值$k \geqslant 2$、物理比特泄露参数$m \geqslant 1$和参与方数量$n \geqslant 1$，文献[11]指出当域F的阶数足够大时，拥有随机赋值性的 Shamir 秘密共享方案具有高概率的抗泄露性；当忽略多个对数因子时，需要确保$\log_2 |F| \geqslant n/k$。该结果（不包括多对数因子）表明，只要泄露总量$mn < k\lambda$，Shamir 秘密共享方案仍然是安全的。此外，文献[11]提出了一种物理比特泄露攻击，主要针对$n = k$个秘密份额和满足$|F| = 1 \pmod{k}$的任何素数域F的$m = 1$个物理比特泄露。在$|F| \to \infty$时，给出了该攻击优势的计算公式。

自适应提取器与传统的随机提取器不同，它要求即使敌手在观察提取器的输出后获得源

上的泄露值，也能保证安全性。在 FOCS 2020 上，Chattopadhyay 等人构造了一种自适应安全的抗泄露秘密共享（LRSS）方案，其信息率和泄露率均为$O(1/n)$，其中n是参与方的数量。文献[12]构造了一个自适应安全的 LRSS 方案，攻击者可以从信息率、泄露率中获取泄露的份额总数之间进行权衡。这些方案可应用于非延展秘密共享和安全信息传输。

使用不经意传输（Oblivious Transfer，OT）的协议很少只需要一个 OT 实例，通常需要批量的 OT，尤其是在为 OT 扩展生成基础 OT 时。在生成批量的 OT 时，可以通过在所有实例中重用某些协议消息来优化 2 轮 OT 协议。文献[13]指出这种批量优化容易出错，并对许多批次优化处理不正确的论文进行了分类，其中一些导致 OT 扩展协议中的泄露。进一步提供了正确优化 2 轮 OT 协议批量设置的应对方法，改进 McQuoid、Rosulek 和 Roy（ACM CCS 2020）的 OT 协议的多项性能。特别是文献[13]给出了一个极其简单的 OT。

本节作者：林昌露（福建师范大学）、丁健（巢湖学院）

参考文献

[1] CRAMER R, RAMBAUD M, XING C. Asymptotically-good arithmetic secret sharing over $Z/p^{\ell}Z$ with strong multiplication and its applications to efficient MPC[C]. CRYPTO 2021: 656-686.

[2] BOYLE E, CHANDRAN N, GILBOA N, et al. Function secret sharing for mixed-mode and fixed-point secure computation[C]. EUROCRYPT 2021: 871-900.

[3] APPLEBAUM B, NIR O. Upslices, downslices, and secret-sharing with complexity of 1.5^{n}[C]. CRYPTO 2021: 627-655.

[4] GOYAL V, SONG Y, SRINIVASAN A . Traceable secret sharing and applications [C]. CRYPTO 2021: 718-747.

[5] BEIMEL A, OTHMAN H, PETER N. Quadratic secret sharing and conditional disclosure of secrets[C]. CRYPTO 2021: 748-778.

[6] ORLANDI C, SCHOLL P, YAKOUBOV S. The rise of Paillier: Homomorphic secret sharing and public-key silent OT[C]. EUROCRYPT 2021: 678-708.

[7] ERIGUCHI R, NUIDA K. Homomorphic secret sharing for multipartite and general adversary structures supporting parallel evaluation of low-degree polynomials[C]. ASIACRYPT 2021: 191-221.

[8] ROY L, SINGH J. Large message homomorphic secret sharing from DCR and applications[C]. CRYPTO 2021: 687-717.

[9] COUTEAU G, RINDAL P, RAGHURAMAN S. Silver: silent VOLE and oblivious transfer from hardness of decoding structured LDPC codes[C]. CRYPTO 2021: 502-534.

[10]MAJI H K, PASKIN-CHERNIAVSKY A, SUAD T, et al. Constructing locally leakage-resilient linear secret-sharing schemes[C]. CRYPTO 2021: 779-808.

[11]MAJI H K, NGUYEN H H, PASKIN-CHERNIAVSKY A, et al. Leakage-resilience of the Shamir secret-sharing scheme against physical-bit leakages[C]. EUROCRYPT 2021: 344-374.

[12]CHANDRAN N, KANUKURTHI B, OBBATTU S L B,et al. Adaptive extractors and their application to leakage resilient secret sharing[C]. CRYPTO 2021: 595-624.

[13]MCQUOID I, ROSULEK M, ROY L. Batching base oblivious transfers[C]. ASIACRYPT 2021: 281-310

7 水印/版权保护

加密水印研究主要集中在伪随机函数的设计上，可通过将标识符嵌入到伪随机函数的密钥中同时实现加密和水印功能。针对现有可水印伪随机函数（watermarking PseudoRandom Functions，watermarking PRFs）安全概念的局限性，文献[1]基于公钥叛逆者追踪的定义引入可追踪 PRF 的新概念，既能满足可水印 PRF 的实例化应用，又能解决现有可水印 PRF 定义的局限性，并在现实场景中提供有意义的安全保障。可追踪 PRF 方案由 Setup、KeyGen、Eval 和 Trace 4 种算法组成。Setup 算法对 PRF 密钥和用于追踪的追踪密钥进行采样；KeyGen 算法以 PRF 密钥和一个标识符作为输入，并输出一个被标记的密钥；Eval 算法对 PRF 密钥或身份密钥进行 PRF 评测；对标识符有全局访问权限的 Trace 算法则输入追踪密钥，并最终输出被破坏的密钥。此外，该工作还展示了如何从私有约束伪随机函数构造非抗共谋可追踪伪随机函数，以及从不可区分混淆构造完全抗共谋可追踪伪随机函数。

如何利用量子状态的不可克隆性实现复制保护，一直是研究者关注的焦点。文献[2]通过利用制定的复制保护的新定义来捕获更广泛的攻击类型，给出了两个关于复制保护的更普遍的结论：与经典的谕言机相比，任何不可学习的功能都可以被复制保护；只需要假设公钥量子货币存在，在某种意义上任何可加水印的功能都可以被复制检测。在推导过程中，作者基于经典谕言机和子空间状态，将谕言机使用后量子混淆结构实例化，为所有不可学习的功能提出一个量子复制保护构造方案。此外，作者还定义了一个具有额外检查过程的复制检测，用户可以通过运行这个检查过程来公开验证程序的有效性，并针对任何可加水印的函数族构造了一种复制检测方案，证明了量子版权复制方案的正确性、有效性、反盗版安全性。

隐子空间思想已被广泛应用于量子货币、签名令牌、复制保护和不可克隆加密与解密等享有某种形式的不可克隆性的密码学原语中。文献[3]将隐子空间的思想推广到隐陪集，探讨了隐陪集状态的不可克隆性质和其在签名令牌、不可克隆解密和复制保护方面的应用。具体贡献如下：①在假设不可区分性混淆的情况下，证明了隐陪集状态具有一定的直积困难特性，

能够在无须预言的普通模型中实现令牌的签名；②结合令牌签名与可提取证据加密，构造了适用于普通模型的不可克隆解密方案；③基于隐陪集状态满足一定的自然纠缠单偶性的猜想，取消了构造的不可克隆解密方案中对可提取见证加密的要求，并依赖对不可预测分布的计算和比较混淆；④提出了一个普通模型下的伪随机函数版权保护方案。

本节作者：廖鑫（湖南大学）

参考文献

[1] GOYAL R, KIM S, WATERS B, et al. Beyond software watermarking: traitor-tracing for pseudorandom functions[C]. ASASIACRYPT 2021 (2): 250-280.

[2] AARONSON S, LIU J, LIU Q, et al. New Approaches for Quantum Copy-Protection[C]. CRYPTO 2021 (1): 526-555.

[3] COLADANGELO A, Liu J, Liu Qipeng, et al. Hidden cosets and applications to unclonable cryptography[C]. CRYPTO 2021 (1): 556-584.

应用密码

1 区块链/数字货币

区块链是一种分布式账本技术，其具有分布式、去信任、不可篡改等优点，在数字货币、跨境支付等领域得到了广泛的应用。在 2021 年的国际三大密码年会上，均有关于区块链的研究成果发表，主要关注的方向包括利用博弈论方法分析区块链安全、基于时间的密码原语、提高区块链交易吞吐量的通用通道协议等。

利用博弈论方法对数字货币等基于区块链的分布式账本应用进行分析，可以帮助我们更深地理解其系统鲁棒性和行为合理性。在文献[1]中，Badertscher 等人利用他们在 2018 年欧密会上提出的合理协议设计（Rational Protocol Design，RPD）框架[2]，分析了针对比特币的 51%双花攻击，并提出了一种针对数字货币底层协议参数的通用解决方法。该方法可以阻止控制大部分系统资源的敌手对区块链一致性进行攻击，包括 51%双花攻击。研究成果可以用于修补遭受此类攻击的系统，如以太坊经典（Ethereum Classic，ETC）和比特币现金（Bitcoin Cash，BCH）。该研究体现了博弈论分析对于研究数字货币经济行为和系统鲁棒性的重要性。

时钟同步器允许各方利用较弱的同步性假设来建立一个公共的全局时钟。尽管现有分布式容错计算技术已经推动了相关研究，但现有的同步器解决方案无法适用于参与方未知、系统配置随时间动态变化的应用场景。在文献[3]中，Badertscher 等人设计了一种支持动态复杂环境的同步器，可以容忍少数参与方的腐化。在此基础上，他们设计了一种基于权益证明（Proof of Stake，PoS）的区块链协议 Ouroboros Chronos。该协议集成了上述同步器，是第一个仅依赖本地时钟的 PoS 区块链协议，可以容忍最坏情况的参与方腐化，同时协议支持动态波动模型。在区块链实际应用中，Ouroboros Chronos 协议不仅适用于静态权益分配，还可以应用于权益转移和部分权益重分配。

基于时间的密码原语，如时间锁难题（Time-Lock Puzzles，TLP）等，已经被广泛应用在区块链中。然而，时间锁难题无法满足通用可组合（Universally Composable，UC）安全，这使得研究者难以在 UC 模型下安全地应用该原语。在文献[4]中，Baum 等人在 UC 框架下建立了 TLP 模型，并进行模块化证明。作者提供了在 UC 模型下使用时间锁原语证明协议安全性的基础工作，构建了基于随机谕言机和满足 UC 安全的 TLP，展示了随机谕言机在证明过程中的必要性。具体而言，作者的贡献主要有 3 点：一是提出了 UC 模型下的抽象时间概念，

即不再以时钟角度描述相关事件的排序；二是证明了可编程随机谕言机在构建 UC 安全 TLP 的必要性，即刻画了可编程谕言机与非可编程谕言机的一种新分离方式；三是对 TLP 进行可组合定义和构建。最后，作者构造并实现了一种特殊的 UC 安全两方计算示例应用。

无许可分布式账本要求每笔交易必须记录上链，无法满足交易的高吞吐需求。一种可行的做法是利用链下协议来解决该问题。在文献[5]中，Aumayr 等人提出了一种通用通道（Generalized Channels，GC）用来支持区块链的链下操作。通用通道通过扩展现有的支付通道功能并弱化现有状态通道的定义来实现。作者提出了一种与大部分区块链兼容的具体结构，该区块链需要支持交易授权、时间锁和常量布尔值与、或操作。具体而言，作者使用适配器签名来构造通用通道，并严格地证明了通用通道满足 UC 安全，实验结果表明通用通道的性能表现优于现有的支付通道实例闪电网络。该构造可以降低一半的通信复杂度，在存在争议的情况下可以将链上存储开销从线性级降低到常数级。最后，作者通过原型实现来评估通用通道的实用性，并讨论了包括两方公平计算的各类应用。

本节作者：包子健、安浩杨、刘洋、何德彪（武汉大学）

参考文献

[1] BADERTSCHER C, LU Y, ZIKAS V. A Rational Protocol Treatment of 51% Attacks[C]. CRYPTO 2021(3):3-32.

[2] BADERTSCHER C, GARAY J, MAURER U. But Why Does It Work? A Rational Protocol Design Treatment of Bitcoin[C]. EUROCRYPT 2018(2):34-65.

[3] BADERTSCHER C, GAŽI P, KIAYIAS A, et al. Dynamic Ad Hoc Clock Synchronization[C]. EUROCRYPT 2021(3):399-428.

[4] BAUM C, DAVID B, DOWSLEY R, et al. TARDIS: A Foundation of Time-Lock Puzzles in UC[C]. EUROCRYPT 2021(3): 429-459.

[5] AUMAYR L, ERSOY O, ERWIG A, et al. Generalized Channels from Limited Blockchain Scripts and Adaptor Signatures[C]. ASIACRYPT 2021(2): 635-664.

2　可搜索加密

文献[1]探讨了如何在基于固态硬盘（Solid State Drive，SSD）的存储介质上高效访问可搜索对称加密（Searchable Symmetric Encryption，SSE）密文数据的问题，并提出了针对 SSD 存储特性进行优化的 SSE 方案。在以往大多数 SSE 方案中，搜索操作的性能瓶颈通常并非源于对称密码运算，而是受制于存储介质的访问开销。这是因为 SSE 构造中应用了不经意的存

储访问技术，使得服务器在每次执行搜索时都必须执行一个超常数级别的不连续的存储访问请求，这在使用传统硬盘时读取效率会非常低。因此，以往的 SSE 方案主要针对传统硬盘进行搜索效率和算法复杂度的优化。文献[1]中指出，在经常使用 SSD 的今天，不连续的存储介质访问不再是限制 SSD 读取效率的主要问题。在 SSD 中，影响读取效率的因素主要在于所读取存储页面的数量。为此，文献[1]定义了页面效率（Page Efficiency）的概念，并提出了一种具有高效页面访问特性的 SSE 方案 Tethys。该方案的技术核心是利用布谷鸟哈希技术将可变大小项目打包到固定大小的桶中，从而实现了 SSD 存储页面的高利用率。实验结果表明，Tethys 具有良好的搜索效率和恒定的存储开销。

文献[2]针对近几年提出的公钥密文流模式匹配问题，提出了一种更高效且支持模式匹配的可搜索公钥加密算法。一般而言，流量分析和隐私之间的矛盾通常可以通过可搜索加密技术来解决。然而经典的可搜索加密方案通常只能支持固定的关键字检索，无法满足实际中所需的能够在任意长的数据流中搜索可变大小字符串的要求。现有的两种支持该特性的可搜索加密算法都使用了双线性映射来构造，它们在效率和安全性方面都有明显缺陷。文献[2]构造了两种新的算法来解决这些问题。第一个算法消除了过去的方案中依赖单个字符生成密文的局限，这大大减少了密文和公钥的大小。在一个具有较长搜索请求的典型应用程序上，新算法的公钥比以往的方案要小两个数量级。方案的第二个算法设法保留了第一个算法的大部分良好特性，同时提升了理论安全性。简单来说，第二个算法的安全性依赖于 DDH 问题的一个变体，而不像以往的方案一样依赖于一个交互式的困难问题。

2017 年，Kamara 和 Moataz 利用对称密码原语构造了第一个支持合取范式（Conjunctive Normal Form，CNF）查询，且具有最优通信开销、最坏情况下的亚线性搜索效率与非平凡泄露的加密多重映射（Encrypted Multi-Map，EMM）方案 BIEX。在文献[3]中，Sarvar Patel 等人对 BIEX 方案进行了进一步优化，并构造了方案 CNFFilter。在执行 CNF 查询时，CNFFilter 的信息泄露显著低于 BIEX。同时，CNFFilter 实现了更快的搜索时间和更少的通信开销，并保持了 BIEX 的所有良好特性，包括最坏情况下的亚线性搜索效率。另外，与 BIEX 相比，CNFFilter 对泄露滥用攻击（Leakage Abuse Attack）具有额外抵抗力。对于大多数 CNF 查询，CNFFilter 避免泄露任何 CNF 查询中出现的单子查询的结果集。例如，对于形式为$(l_1 \vee l_2) \wedge l_3$的 CNF 查询，CNFFilter 不会泄露子查询l_1、l_2或l_3的结果。CNFFilter 的核心是一种新的过滤算法，该算法通过牺牲一定的服务器端存储开销，显著降低了执行集合求交过程中的泄露量。

结构化加密（Structured Encryption，STE）是一类能够加密数据结构，并支持在所生成密文上执行安全查询请求的对称加密体制。STE 的一个特例是可搜索对称加密 SSE。2018 年，Kamara 等人针对 STE 构造了一个查询相等性泄露抑制框架，该框架能够隐藏两个查询请求是否相等。然而，Kamara 等人所构造的框架是静态的，即不支持对已有的密文数据进行添加与删除，这极大地限制了该框架的实用性。针对 Kamara 等人所提出框架的缺陷，在文献[4]

中，George 等人设计了一个具有动态性的泄露抑制框架。该框架将能够隐藏查询结果数量的半动态（支持添加操作）或可变的 STE 方案转换为支持数据动态添加和删除，并隐藏两个查询请求是否相等的 STE 方案。George 等人在 3 个基本 EMM 方案上应用了所提出的泄露抑制框架，并得到了对应的具有零或几乎零泄露的动态 EMM 方案。最后，George 等人在自然假设下证明了所获得的 3 个 EMM 方案具有比以往黑盒不经意随机访问机（Oblivious Random Access Machine，ORAM）模拟更高的渐进复杂度。

本节作者：徐鹏、王蔚、陈天阳、郑宇博（华中科技大学网络空间安全学院）

参考文献

[1] BOSSUAT A, BOST R, FOUQUE P A, et al. SSE and SSD: page-efficient searchable symmetric encryption[C]. Annual International Cryptology Conference. Springer, Cham, 2021: 157-184.

[2] BOUSCATIÉ E, CASTAGNOS G, SANDERS O. Public Key Encryption with Flexible Pattern Matching[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 342-370.

[3] PATEL S, PERSIANO G, SEO J Y, et al. Efficient boolean search over encrypted data with reduced leakage[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 577-607.

[4] GEORGE M, KAMARA S, MOATAZ T. Structured encryption and dynamic leakage suppression[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 370-396.

3　隐私信息检索

隐私信息检索（Private Information Retrieval，PIR）是一种允许一个客户端在多个（大于等于两个）服务器上实现秘密查询请求的密码学原语。传统的 PIR 方案能够实现对数多项式级别的查询通信复杂度，但是需要每个服务器消耗与数据集大小线性相关的计算开销。近期的部分 PIR 方案通过在初始化阶段令服务器对数据库进行预处理，将查询时的服务器计算开销降低为与数据集大小呈亚线性关系，然而这些方案要么增加了查询时的带宽开销，要么需要服务器存储大量的客户端状态，或者依赖于更强的安全性假设。针对上述问题，文献[1]构造了一个双服务器的预计算 PIR 方案，该方案通过推广私有可穿刺伪随机集（Privately Puncturable Pseudorandom Set，PRSet）的定义，设计出更通用的 PRSet 方案，并在此基础上

进行构造。该方案的安全性依赖于标准 LWE 假设，实现了预处理阶段服务器的$\tilde{O}(n)$计算复杂度，与客户端的$\tilde{O}(\sqrt{n})$计算与带宽开销，并对后续查询实现了$\tilde{O}(\sqrt{n})$的计算开销和$\tilde{O}(1)$的带宽开销与$\tilde{O}(\sqrt{n})$的客户端存储开销。

私有连接和计算（Private Join and Compute，PJC）是一种支持对分布式数据进行安全计算的基本功能。在文献[2]中，作者提出了默认隐私信息检索（PIR-with-default）的概念，并在此基础上构造了一个两方 PJC 协议。该协议能够隐藏两个参与方数据交集的大小，其通信开销与两方中较大数据集的大小呈亚线性关系。作者提供了两种在半可信服务器模型下的实例化方案。在方案 1 中，服务器在初始化阶段对自身（较大）数据集中的内容进行预计算，以降低后续客户端在线查询时的开销。该方案在初始化阶段消耗与服务器端数据集大小呈线性关系的计算开销，以达到后续客户端运行时的计算与通信复杂度均与客户端本地（较小）数据集线性相关的效果。方案 2 直接利用 PIR 进行实例化构造，其客户端查询请求的计算与通信复杂度和客户端数据集大小呈渐进线性关系，且与服务器数据集大小呈对数关系。上述两种构造都能够与差分隐私技术相结合，来隐藏不同查询输出之间的相关性。

本节作者：徐鹏、王蔚、陈天阳、郑宇博（华中科技大学网络空间安全学院）

参考文献

[1] SHI E, AQEEL W, CHANDRASEKARAN B, et al. Puncturable pseudorandom sets and private information retrieval with near-optimal online bandwidth and time[C]. Annual International Cryptology Conference. Springer, Cham, 2021: 641-669.

[2] LEPOINT T, PATEL S, RAYKOVA M, et al. Private join and compute from PIR with default[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 605-634.

4 不经意随机存储

不经意随机访问机（Oblivious Random Access Machine，ORAM）由 Goldreich 和 Ostrovsky 首次提出，其作用是一个隐藏其上所运行程序的内存访问模式，实现对于每个输入，外部所观察到程序所访问内存单元的分布是相似的。

目前，对于一组常见参数，ORAM 的性能下界是未知的，即使在受限的计算模型（如球箱模型）中也无法得到具有理论意义的下界值。具体来说，令 N 表示一个程序运行所需的存储单元的个数，其中每个存储单元的长度为 ω 比特，令 b 表示 ORAM 的存储单元大小，所有已有 ORAM 的下界均以 ω/b 作为其复杂度的一个因子。这样的下界是平凡的，并不具有理

论意义。在文献[1]中，Komargodski 等人基于上述参数证明了一个新的 ORAM 下界，并且即使针对基于其他参数的 ORAM 实现，该下界也并不高于已有方案所证明的下界。具体来说，Komargodski 等人证明了对于所有的 ORAM 方案，程序的一个逻辑内存操作，将导致$\Omega(\log(\frac{N\omega}{m})/\log(\frac{\omega}{b}))$的 ORAM 内存访问开销，其中$m$表示 ORAM 运行时的本地存储大小。

Komargodski 等人应用上述结果首次分别得到了在线 ORAM（ORAM 无法提前得知所有的内存访问序列）与离线 ORAM（ORAM 能够提前得知所有的内存访问序列）不同性能开销情况。具体来说，他们证明了当$\omega = \log N$，并且b与m均与N呈对数多项式关系时，存在一个对每个逻辑操作均进行$O(1)$次内存访问的离线 ORAM，而相对地，每个在线 ORAM 均需消耗$\Omega(\log N/\log\log N)$次内存访问。

ORAM 构造的一种方式是层次化 ORAM（Hierarchical ORAM）。在层次化 ORAM 的构造中，服务器端将 ORAM 存储结构划分为多层，第i层存储$2^i \cdot b$个 ORAM 数据块，其中b为正整数。当客户端需要读或写某个数据块时，会从所有层分别查找选取若干个数据块到本地，再对第一层的数据块进行重新写入。每当客户端执行了2^{i-1}次读或写操作，便将第$i-1$层中的内容写入到第i层中。这样的读写方式实现了$\log N$的带宽与客户端计算开销，其中N表示 ORAM 的最大容量。层次化 ORAM 利用哈希函数将数据的逻辑地址映射到数据块的物理存储地址。2012 年，Goodrich 等人采用带缓冲区的布谷鸟过滤器（Cuckoo Hashing with a Combined Stash）代替了传统哈希函数，从而实现了更高的地址映射效率。然而，在文献[2]中，Falk 等人发现 Goodrich 等人的方案存在安全问题，即在对某层数据块进行写入时，若该层的某个数据存在于缓冲区中，对该数据进行布谷鸟哈希的过程可能会泄露访问模式。该问题会导致层次化 ORAM 提供的安全性质失效，并影响了至少 5 个后续基于 Goodrich 等人的工作所构造的层次化 ORAM 方案。在文献[2]中 Falk 等人详细描述了基于所发现安全问题的攻击流程。为了解决这一安全问题，Falk 等人在对单层数据写入时引入了冗余操作，使得存在于缓冲区中的数据与不在缓冲区中的数据无法区分，从而隐藏了访问模式。

在文献[3]中，Health 等人构建了一种高效的配合零知识（Zero Knowledge，ZK）证明（Zero Knowledge Proof，ZKP）计算验证的 ORAM 系统（PrORAM）。相比其他的 ZKP 协议，这一方案和 ZKP 协议结合时，对 RAM 的内存访问效率被提升至每次访问仅需$2\log N$不经意传输（Oblivious Transfer）协议，其中N为 RAM 中数据块的总数。基本设计流程可以被简单概括为：证明者对某一秘密值进行分享后，由验证者选择一次性掩码，证明者利用验证者所选择的掩码加密数据并保存至本地 RAM。验证者在使用零知识技术验证证明者 RAM 中保存的数据时，这个 RAM 保证验证者无法根据数据的访问顺序来推断出存入顺序，并借此推导出秘密和保护掩码之间的对应关系。为了实现这一点，Health 等人利用了 OT 协议来实现验证者对 RAM 的访问，并设计了一个按序写且仅能一次读的 ORAM 结构，称为 swordORAM。在

每次读写之后，swordORAM 都需要刷新。在此基础上，通过简单的归约，即可将 swordORAM 转化成文献[3]设计的核心：PrORAM。

在以往的 ORAM 研究中，研究者主要关注 ORAM 经过摊销（Amortized）的访问效率（ORAM 在长时间运行后的平均访问效率），而忽视了最坏情况下的访问效率，这导致了在最坏情况下，这些 ORAM 会导致$\Theta(N)$的访问开销，其中N表示 ORAM 数据块的数量。这一问题源于 ORAM 在使用过程中需要周期性对密文数据进行重建，从而为未来的数据更新预留空间。针对这一问题，Asharov 等人在文献[4]中构造了最坏情况下最优的 ORAM，即使在最优情况下，Asharov 等人的构造也能保持$O(\log N)$的带宽与计算开销，以及$O(1)$的客户端存储与$O(N)$的服务器端存储开销。为了实现上述特性，Asharov 等人基于层次化 ORAM 进行了方案构造。对层次化 ORAM 的每层，Asharov 等人都维护两个不同的副本，这样当第一个副本中的某层进行重建操作时，第二个副本中的对应层可用于访问查询；当第二个副本中的某层被重建时，第一个副本中对应层的下一次可用于访问查询。在此基础上，Asharov 等人将对 ORAM 中每层的重建工作分散到使用对 ORAM 的过程中，而非像以往的工作一样仅当某些时刻才进行重建工作，从而避免了最坏情况下的$\Theta(N)$开销。

本节作者：徐鹏、王蔚、陈天阳、郑宇博（华中科技大学网络空间安全学院）

参考文献

[1] KOMARGODSKI I, LIN W K. A logarithmic lower bound for oblivious RAM (for all parameters)[C]. Annual International Cryptology Conference. Springer, Cham, 2021: 579-609.

[2] HEMENWAY FALK B, NOBLE D, OSTROVSKY R. Alibi: A flaw in cuckoo-hashing based hierarchical oram schemes and a solution[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 338-369.

[3] HEATH D, KOLESNIKOV V. PrORAM: Fast O(logn) Private Coin ZK ORAM [C]. Annual International Conference on the Theory and Applications of Cryptographic Techni-q ues. Springer, Cham, 2021: 495-525.

[4] ASHAROV G, KOMARGODSKI I, LIN W K, et al. Oblivious RAM with Worst-Case Logarithmic Overhead[C]. Annual International Cryptology Conference. Springer, Cham, 2021: 610-640.

5 隐私保护与机器学习

技术的不断迭代发展，社会效率得到不断攀升，人们生活得到极大的丰富，同时因为过

度依赖技术导致越来越多的隐私泄露问题。隐私保护一直是安全领域研究的重点内容。解决隐私保护问题，不能靠构造封闭的围墙来断绝技术的发展，要在技术进步时考虑用更多的技术手段来达到安全目标。

差分隐私一直是用来保护隐私、满足监管要求的重要技术，其具有在隐去用户真实信息的情况下精确计算一些函数的功能。但是差分隐私模型往往需要一个中心，这个中心能够获得用户的原始数据。这提供了用户数据泄露的可能，增加了系统的安全风险。为了解决这个问题，本地差分隐私被更多学者所研究，但减少信任假设的同时给计算带来很大的偏差甚至错误。扰乱模型（Shuffle Model）被提出，这是介于本地差分隐私模型和中心化差分隐私模型的解决方案。Badih Ghazi 等人在文献[1]中研究扰乱模型下差分隐私的一些特性。将频次估计（Frequency Estimation）问题在单消息和多消息扰乱模型下都进行了优化，达到了比较紧的界，同时对于选择（Selection）问题在单消息模型下进行了归约，获得了目前最紧的界。

匿名路由是在线隐私的基础研究，已经被学者研究了数十年。解决匿名路由问题，可以帮助计算机实现真正的匿名性。Elaine Shi 等人在文献[2]中提出非交互匿名路由（Non-Interactive Anonymous Router，NIAR）的解决方案，旨在通过使用密码学手段来实现匿名路由。非交互路由相对于之前的匿名路由方案相比，不再有破解匿名的门限约束，同时仍然能够保证不可信的路由并不知道消息的发件方、收件方，以及具体内容。作者讨论了匿名路由的隐私特性，给出了安全目标、定义了严格的安全性，采用函数加密来构建方案，并在计算不可区分性下证明了该方案能够满足对应的安全目标。

匿名性和可说明性（Accountability）是互相矛盾的两个特性。常规技术手段同时实现这两个安全目标是不可能的事情。Libert 等人在文献[3]中提出了一种名为分叉签名（Bifurcated Signatures）的算法来同时实现两个安全目标，作者甚至在同一方案中实现了可选择性打开可说明性的功能。分叉签名能够进行完全匿名的签名，也能够选择留下身份契约供权威机构认证。作者首次给出了分叉签名的精确定义和安全目标，随后作者通过同态模糊承诺（Homomorphic Equivocal Commitment）和非交互零知识证明协议构造了一个基于 LWE 问题的方案。特别是分叉签名与基于政策的签名（Policy-Based Signatures）最大的区别在于签名者是否有选择匿名的权力。作者证明了该方案的安全性并分析了该方案的效率。

隐私和监管审查同样是两个互相矛盾制约的方面。如何避免权力的滥用导致的隐私问题是 Matthew Green 等人在文献[4]中的研究重点。端对端加密能够保证仅有发送方和接收方能够获得消息，其他人无法获得任何有用信息。这驱使法律的执行者和国家安全的管理者对此产生了抵制。相对于允许监管方获得自己的密钥，用户在满足监管要求的同时，仍然希望自己所传输内容得到保护。这种需求就要求给监管方的权力进行约束。文献[4]给出了一种全新的抗滥用权力的访问控制系统（Abuse Resistant Law Enforcement Access Systems），该系统在获得授权的时候能够对用户的密文进行一定次数和限制范围的审查。该协议是基于认证通信、

非交互零知识证明、多发送者的安全计算和证据加密的，作者在通用组合模型下证明了该功能的正确性和安全性。

本节作者：张硕（北京邮电大学）

参考文献

[1] GHAZI B, GOLOWICH N, KUMAR R, et al. On the Power of Multiple Anonymous Messages: Frequency Estimation and Selection in the Shuffle Model of Differential Privacy[C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology-EUROCRYPT 2021.

[2] SHI E, WU K. Non-Interactive Anonymous Router[C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology-EUROCRYPT 2021.

[3] LIBERT B, NGUYEN K, PETERS T, et al. Bifurcated Signatures: Folding the Accountability vs. Anonymity Dilemma into a Single Private Signing Scheme[C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology-EUROCRYPT 2021.

[4] GREEN M, KAPTCHUK G, VAN LAER G. Abuse Resistant Law Enforcement Access Systems[C]. In: Canteaut, A., Standaert, FX. (eds) Advances in Cryptology-EUROCRYPT 2021.

6 抗泄露/旁路

近年来，密码学中有关旁路及抗泄露的研究仍集中在基于旁路的攻击以及对旁路的防护上。密码算法往往需要依托于各种软硬件设备来实现，旁路攻击则利用算法实现产生的一些额外信息来进行分析。旁路攻击可以分为两类：一类被称为被动型旁路攻击（SCA），这类攻击方法会被动地测量密码实现产生的（如功耗、电磁、时间等）旁路泄露信息用于密码分析；另一类攻击是主动型旁路攻击（FA），与被动型旁路攻击不同，攻击者会主动干扰密码设备运行，从而获取其中的特定的旁路信息。旁路防护与旁路攻击恰恰相反，一般是指通过某种特殊的方法实现密码算法或在密码算法实现中增加一些特殊保护机制，使得攻击者难以通过旁路信息对密码算法进行攻击。

6.1 旁路攻击的发展情况

近年来，旁路攻击的重要研究仍集中在公钥体制密码算法上，其包括对于椭圆曲线公钥算法及后量子密码算法进行的旁路攻击。由于近些年来量子计算机的发展，传统对称密码与非对称密码都遭受一定程度的威胁。旁路分析技术作为对于密码安全性的评判标准之一，后量子密码的旁路安全性也越来越受到密码学家的关注。相较于往年而言，2022 年国际三大密

码年会增加不少有关旁路攻击工作的收录。

1．分组密码算法旁路攻击

针对旁路攻击，密码设备通常需要在加密算法中加入针对 SCA 和 FA 的两种对策以保护来自旁路的攻击。但针对 SCA 和 FA 的对策的直接组合更容易受到 FA 的攻击，如统计无效攻击（SIFA）和故障模板攻击（FTA）。基于此，现阶段已有的对策能够保护来自 SCA 和 SIFA 的组合攻击。Saha S 等人针对 SIFA 和 SCA 组合对策提出了一种称为 SCA-FTA 的新型组合攻击[1]。SCA-FTA 攻击是指利用故障注入期间所泄露的旁路信息来增强 FTA 攻击，其可以利用针对故障检测/纠错逻辑的旁路信息泄露。Saha S 等人使用 SCA-FTA 成功攻击了开源软件实现的 Keccak（含有 SIFA 对策）。

2．椭圆曲线算法旁路攻击

近年来，已经有了不少针对椭圆曲线算法旁路攻击的工作，如 Refined Power Analysis（RPA）、Zero-Value Point（ZVP）、Exceptional Procedure Attack（EPA）等攻击方法。Sedlacek 等人归纳了过去已公开这 3 种针对椭圆曲线算法的旁路攻击方法[2]。Sedlacek 等人提到这些攻击均与椭圆算法中点的运算有关，利用某些异常点在点运算时泄露的旁路信息完成攻击。由于在椭圆曲线中，不同曲线上的点在运算时所采取的公式均不同，因此作者的工作借助 Explicit-Formulas Database（EFD）数据库辅助进行。该数据库整理了不同曲线模型上不同坐标系统下的运算方式。作者利用该库整理出了与上述 3 种攻击相关的关键位置，尤其是“点加”式。作者将 3 种攻击整理为一种统一的攻击框架并加以推广，提出一种适应性的攻击方法并对基于窗口的标量乘法运算进行攻击验证。通过该框架，也可以在各类点运算的公式中找到更多的异常点。

3．后量子密码旁路攻击

后量子密码在设计之初，设计者必须在数学层面上证明其能够保证安全性，但密码算法的安全性除在理论上需要证明之外，还需考虑算法在具体实现上的安全。旁路攻击能够利用密码算法在密码芯片上运行时泄露出来的侧信息，分析并恢复出密钥或加密信息，是密码算法物理安全的主要威胁手段。对于 NIST 后量子竞赛中的密码算法，文献[3]中的第三轮后量子密码算法中的 9 种密钥封装算法（包括第三轮选中的算法以及备选算法），提出一种利用跳过指令故障进行私钥还原攻击的方法。Xagawa 等人提出的故障攻击针对的目标是第三轮算法中均在使用的 FO 转换机制（Fujisaki-Okamoto Transformation）。该机制通过哈希函数等操作，可将 CPA 安全等级下的公钥加密算法提升到 CCA 安全等级的密钥封装算法。在该机制中，最重要的环节是在解封装时，验证解密得到的结果是否符合要求，从而验证是否存在恶意密文输入。Xagawa 等人通过对 pqm4 公开库中这 9 种算法的实现代码进行评估，发现大部分算法可以利用一次跳过指令故障来跳过该验证阶段，使得该机制失效，从而可以利用选择明文（Plaintext-Checking）得到恶意密文并对解封装环节进行私钥还原攻击。Cayrel 等人针对

NIST 后量子密码竞赛的决赛候选密码 Classic McEliece 算法提出一种故障攻击方法[4]。Cayrel 等人对芯片背部使用激光照射实现单比特和双比特精度的故障注入。注入故障用于指令篡改。攻击者在芯片从 Flash 中读取特定指令时注入故障，使得该条特定故障由于二进制数的改变，被篡改为另一条指令。通过该故障可以使得 Classic McEliece 算法中某个运算的范围从 2 的有限域变为自然数域，接着可利用整数线性规划问题（Anteger Linear Programming）对该算法进行多项式时间内的求解。

6.2 旁路防护的发展情况

近年来，旁路有关的工作仍注重在旁路防护上。有关旁路防护的工作包括新型旁路防护模型、新型旁路泄露模型、新型掩码方案及掩码安全的自动化验证等。

1．新型旁路防护模型

新型旁路防护模型主要集中在针对传统密码及消息验证码机制的旁路防护上。在传统密码的旁路攻击中，差分故障攻击（DFA）一直以来都是一个非常有威胁且能够应用于大多数分组密码的攻击方法。如何能够在密码设计的角度上实现对于 DFA 的防御是一个重要的公开问题和研究热点。基于此，Berti 等人设计出了可以用于抵抗 DFA 的具有特殊线性结构的 S 盒，同时他们构建了一个名为 DEFAULT-LAYER 的 DFA 保护层，以及一个能够防御 DFA 的分组密码 DEFAULT[5]。除 DFA 之外，代数攻击还是一个强大的旁路攻击技术。白盒密码的实现能够抵御大多数旁路攻击但是其难以抵御代数攻击，Biryukov 等人考虑到代数攻击对于白盒密码实现的威胁，引入了一种称为 Dummy Shuffling 的旁路防御技术[6]。Dummy Shuffling 技术通过增加虚拟“随机”输入来扩展经典乱序技术（Shuffling），该技术能够抵御任何阶数的代数攻击。

另一类旁路防护的工作聚焦在针对消息认证码（MAC）上。MAC 用于保证通双方的数据完整性，发送方对消息进行哈希之后使用共享的密钥加密后将 MAC 发送给对方，对方收到 MAC 使用密钥解密并与接收到信息的哈希值进行对比以确认数据完整性。Dobraunig 等人[7]注意到在消息加密认证中的一个特定操作：敏感数的比较操作。该操作可能会泄露旁路信息，而被用于攻击。在认证环节中可能算法计算的值会与攻击者提供的恶意数据进行比较。因此，Dobraunig 等人提出以抗泄露密码（Leakage-Resilient Cryptography）的原理，借助密码学置换的方法做到对该比较操作的安全实现，而不需额外添加其他旁路防御手段。该敏感数的比较操作的安全实现能够应用在消息认证的场景中。另外，Berti 等人相较于对原有 MAC 方案的修改，其提出一种新的防泄露 MAC 方案[8]。该方案由一个抗冲突散列函数以及可调整分组密码（TBC）构成，其设计有两个重点：具有较大可调整位分组密码、使用可调整密码的逆来与常数验证。

2．新型旁路泄露模型

在旁路安全性分析中，旁路的泄露模型可以用于评估密码算法的安全性。Brian 等人对旁路攻击理论中的泄露模型进行了充分的调研和整理[9]。他们提出，在旁路攻击的理论分析中，泄露模型是帮助衡量密码学原语（Cryptographic Primitives）是否安全的依据之一。抗泄露密码（Leakage-Resilient Cryptography）即在传统的密码算法理论安全证明中加入了各类旁路泄露模型，来验证算法在旁路攻击环境下的理论安全性。最基本的泄露模型为有界泄露模型（Bounded Leakage Model），该模型表示对于变量 X，其由旁路泄露信息造成熵的降低值不超过一个特定的界限。但是真实环境的旁路泄露并不完美符合有界泄露模型，近年来多种有噪泄露模型（Noisy-leakage Model）被提出，即对变量 X，其旁路泄露信息存在特定参数的噪声。本文作者统一整理了近年来的各类有噪泄露模型，并提出一种信息泄露模型—密度泄露模型（Dense Leakage Model），以此模型可以将各类有噪泄露模型归约到有界泄露模型，可以对部分已经证明在有界泄露模型下安全的密码学原语进行验证，证明其在有噪泄露模型下仍然安全，虽然这些密码学原语在有噪泄露模型下存在安全参数降低的问题。

3．新型掩码方案

在旁路的防护方法中，掩码防护是其中非常重要的一块内容。掩码防护将原本的敏感变量按照一定规则映射到 n 个分摊变量中，该映射规则保证从 n 个变量中任意掌握 d（$d<n$）个变量也无法得知与敏感变量有关的任何信息，以此来抵御旁路攻击。实际的旁路攻击是多变而复杂的，攻击的有效性往往也与设备平台息息相关。因此，为了更好地对旁路攻击进行防御，也为了将攻击与防御的有效性从设备和平台进行剥离，学界致力于抽象并建立符合旁路攻击特性的信息泄露模型。信息泄露模型通常假设攻击者的攻击行为，并通过假定的攻击行为建立信息泄露方式和数学模型，最后证明设计的掩码方案可以防止攻击者通过收集到的泄露信息恢复出敏感信息。

当前学界公认的以及常用的信息泄露模型主要是 Probing 模型和 Noisy Leakage 模型。Probing 模型假定攻击者可以从泄露中收集到 t 个中间量的信息，而掩码方案则需要保证即使攻击者捕捉到了这 t 个中间量，也无法重建敏感变量。而 Noisy Leakage 模型则假设所有的中间量都会和噪声一起泄露。上述两个模型的安全归约依赖于一个中间模型，Random Probing 模型，且 Random Probing 模型是 Noisy Leakage 模型的严格归约。当前针对 Random Probing 模型的研究，比较重要的一步是 Sonia Belaïd 等人在 2020 年设计并提出了一个完整框架，该框架可以生成满足 Random Probing 模型安全性的实现。2021 年 Belaïd 等人延续了上述工作的“扩展”策略，针对 Random Probing Expandability（RPE）这个安全概念进行了深入的剖析，并首次为满足 RPE 安全的组件的一个重要参数“增益阶”（Amplification Order）标定了最高上限。增益阶定义了一个扩展编译器的复杂度，即在同等条件下，当增益阶越高时，扩展编译器的渐进复杂度就越低。

Belaïd 的另一项工作在该“扩展”策略上加入了动态策略，即在扩展时的每步都动态地选取门电路，使得组件即使在分摊数量较多的情况下也可以很好地平衡可容忍的泄露概率和复杂度[11]，在扩展框架中有着更好渐进复杂度的同时可以容忍原本只能在分摊数量较小情况下才能容忍的最高的泄露概率。结合动态策略之后，Random Probing 安全的 AES 算法的计算复杂度相较于第一项工作下降了 1/10。其次在动态策略的帮助下，该项工作构造了新的 RPE 组件，使得其复杂度由平方级降低到了准线性（Quasi-Linear）级。

上述研究致力于构造具有 Random Probing 模型安全性的组件，提出并进一步提升扩展策略，而 Cassiers 等人则是利用新的工具和新的组合概念来具体分析电路在真实噪声情况下与分摊变量数量较小的电路（更加实际的攻击场景）的 Random Probing 安全[12]。Cassiers 等人设计提出了一个名为 STRAPS 的工具，其能够针对规模较小的电路的随机探测安全性抽样测试，由于计算复杂度，该工具往往仅能针对规模较小的电路或在分摊变量数量较小的情况下进行分析，Cassiers 等人进一步提出了一个新的安全概念“探针分布表”（Probe Distribution Table，PDT）来为电路限定更加严格的安全界限，并已将该概念整合进 STRAPS 工具中。

除针对 Random Probing 模型展开的工作之外，Coron 等人在抗旁路攻击的电路混淆方法方面做了不少工作[13]。Coron 等人提出了一种更为简单的电路混淆构造，该构造可以在组合（Stateless）模型中达到针对 t 个 Probe 的最坏情况下的统计性安全，并且将时间复杂度从 $O(t\log 2(t))$降低为 $O(t)$。在时序（Stateful）模型中，他们针对该构造进行了一些改进，不仅降低了时间成本，还减小了电路规模。

4．掩码安全的自动化验证

掩码方案的安全性在信息泄露模型中（如 Probing 模型）通常由安全阶表示，当安全阶越高时，攻击者攻击成功所需付出的代价（如所需的迹线数）就越大，且攻击代价会随着安全阶的增加呈指数增长。但是，安全阶越高，往往也意味着掩码方案的实现成本越高。学界也尝试设计一些软件工具来验证组件（如 ISW 乘法）的安全性，比较重要的两个工具分别是 maskVerif 和 SILVER。前者可以用于使用常见的信息泄露模型来验证一些使用高级语言实现的掩码方案，而后者则利用实际的硬件合成文件在门电路级对掩码方案的安全性进行验证。maskVerif 工具的缺点在于验证时间比较长。

Bordes 等人扩展并提升了 Belaïd 等人在 2017 年提出的一种基于矩阵的模型，并使其可以覆盖任意伽罗华域[14]。此外，Bordes 等人利用改善的验证模型生成了一个有限的验证算法，该安全验证算法在验证软件乘法组件时比当前最先进的 maskVerif 工具提升了 3 个数量级的速度。在掩码设计方面，Gigerl 等人整理并提出了一些通用的规则，这些规则将复杂架构中的复杂部分（如流水线长度、执行块数量、体系结构缓冲区等）考虑进去，用于设计软件级的掩码方案[15]。他们使用一个实例证明了上述提出的约束和规则可以应对各种各样的实现策略，并且发现如果已知一个处理器的网络列表，那么可以通过上述规则为像 SweRV 这样的复

杂处理器构建安全且有效的掩码方案，且成本降低为原来的 13%。

本节作者：张帆（浙江大学）

参考文献

[1]SAHA S,BAG A, JAP D, et al. Divided We Stand, United We Fall: Security Analysis of Some SCA+ SIFA Countermeasures Against SCA-Enhanced Fault Template Attacks [C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 62-94.

[2]SEDLACEK V, CHI-DOMÍNGUEZ J J, JANCAR J, et al. A formula for disaster: a unified approach to elliptic curve special-point-based attacks[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 130-159.

[3]XAGAWA K, ITO A, UENO R, et al. Fault-injection attacks against NIST's post-quantum cryptography round 3 KEM candidates[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 33-61.

[4]CAYREL P L, COLOMBIER B, DRĂGOI V F, et al. Message-recovery laser fault injection attack on the Classic McEliece cryptosystem[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 438-467.

[5]BERTI F, GUO C, PETERS T, et al. Efficient Leakage-Resilient MACs Without Idealized Assumptions[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 95-123.

[6]BIRYUKOV A, UDOVENKO A. Dummy shuffling against algebraic attacks in white-box implementations[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 219-248.

[7]DOBRAUNIG C, MENNINK B. Leakage resilient value comparison with application to message authentication[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 377-407.

[8]BERTI F, GUO C, PETERS T, et al. Efficient Leakage-Resilient MACs Without Idealized Assumptions[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 95-123.

[9]BRIAN G, FAONIO A, OBREMSKi M, et al. The mother of all leakages: How to simulate noisy leakages via bounded leakage (almost) for free[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 408-437.

[10]BELAÏD S, RIVAIN M, TALEB A R. On the power of expansion: more efficient constructions in the random probing model[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 313-343.

[11]BELAÏD S, RIVAIN M, TALEB A R, et al. Dynamic Random Probing Expansion with Quasi Linear Asymptotic Complexity[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 157-188.

[12]CASSIERS G, FAUST S, ORLT M, et al. Towards tight random probing security [C]. Annual International Cryptology Conference. Springer, Cham, 2021: 185-214.

[13]CORON J S, SPIGNOLI L. Secure Wire Shuffling in the Probing Model[C]. Annual International Cryptology Conference. Springer, Cham, 2021: 215-244.

[14]BORDES N, KARPMAN P. Fast verification of masking schemes in characteristic two[C]. Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021: 283-312.

[15]GIGERL B, PRIMAS R, MANGARD S. Secure and Efficient Software Masking on Superscalar Pipelined Processors[C]. International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2021: 3-32.

7 真实协议/广播/洋葱路由

蓝牙已经成为一种无处不在的短距离通信标准，被广泛应用在智能手机、计算机及耳机等设备。蓝牙协议目前主要有两种版本：第一种为 BR/EDR 版本，用作持续数据流传输，如无线蓝牙耳机设备等；第二种为 BLE 版本，用作低功耗的定期数据传送，如健身追踪设备等。正是由于蓝牙目前被广泛使用，蓝牙通信的安全性也备受关注。蓝牙通信安全依赖于由参与通信的设备间建立的通信密钥。在最新的蓝牙 V5.3 中，建立该密钥的方法被称为“安全连接协议”（Secure Connections Protocol）。安全连接协议的主要工作是执行安全配对（Secure Simple Pairing，SSP）的过程，即通过密码学工具实现设备间身份认证与通信密钥生成。安全连接协议是一个协议族，所有成员共享一个带密钥确认的椭圆曲线 Diffie-Hellman 密钥交换。安全连接协议族中不同具体协议差异主要体现在认证阶段，适用于不同的执行协议设备。根据不同的设备输入-输出能力，认证模式分为 4 种，即 Numeric Comparison、Passkey Entry、Out-of-Band 及 Just Works。

到目前为止，已有不少关于蓝牙安全连接协议的安全分析工作。Lindell （CT-RSA 2009）对基于 Numeric Comparison 的密钥交换协议进行了研究分析，指出该协议是一个安全的密钥

交换协议。Troncoso 和 Hale（NDSS 2021）对基于 Passkey Entry 的密钥交换协议进行了安全性分析，指出该协议可受到一类 man-in-the-middle 攻击，并给出了修改版协议。但是这些工作都是针对蓝牙安全连接协议族中的某个单独子协议进行的安全性分析。Von Teschirschnitz 等人（S&P 2021）指出可以利用蓝牙的不同子协议之间的不良相互作用发起威胁蓝牙认证和密钥安全的 Method Confusion 攻击。

文献[1]综合考虑了针对蓝牙安全连接协议的一系列攻击（包括上述利用不同子协议之间不良相互作用的攻击），对蓝牙安全连接协议进行了建模与安全性分析，指出针对所分析的蓝牙安全连接协议，敌手只有在连接的初始阶段发起积极的攻击才可能威胁密钥安全。文献[1]提出了一种 Trust-On-First-Use（TOFU）认证密钥交换模型。该模型使用 Bellare-Rogaway 风格的基于游戏的安全模型对敌手的攻击行为进行了建模，且被证明满足密钥交换协议 Match-Security 与 Key Secrecy 两个基本安全属性。具体地，文献[1]给出了敌手在 Mactch-Security 属性的优势上界为 $q_s^2 \cdot 2^{-|\text{nonce}|}$，其中 q_s 为敌手总共可执行的会话数量，在 BR/EDR 模式下|nonce|=128，在 BLE 模式下|nonce|=64。协议的 Key Secrecy 安全性归约到了 PRF-ODH、PRF 安全假设上，敌手优势上界为 $q_s^3 \cdot \mathcal{A}_{B,\text{PRF},G}^{\text{PRF-ODH}}(\lambda) + q_s \cdot \mathcal{A}_{C,\text{PRF}}^{\text{PRF}} + q_s^2 \cdot 2^{-|\text{nonce}|}$。文献[1]还对蓝牙 BLE 版本中地址随机化技术的隐私保护技术进行了研究，指出如果忽略基于物理特征设备识别等方法，那么这种机制可实现隐私保护。

在流感大流行的早期阶段，当疫苗尚不可用时，遏制其传播的最重要干预措施是隔离流感感染者及其与他们有过密切接触的人，以此打破感染链。在流感病例数较低的阶段，通过与流感感染者面谈等形式的手工接触者追踪是一种有效方法。然而，当流感感染人数过多时，手工接触者追踪是不可行的，需要一种自动化的解决方式来完成流感感染者及其接触者的追踪。基于智能手机的数字接触者追踪满足自动化要求，能够实现对流感感染者的追踪，并将追踪覆盖范围扩大到陌生接触者。

在基于智能手机的数字接触者追踪方案中，多个用户的智能手机通过蓝牙低功耗通信机制执行邻近检测并联合执行数字接触者追踪，使用户能够检查他们是否曾与具有传染性的流感感染者用户共处。当前的数字接触者追踪方案可分为中心化与去中心化两类方法。DP-3T 项目研究人员指出中心化数字接触者追踪方案没有考虑中央服务器会侵犯用户隐私的场景。而在去中心化方法中，文献[2]指出当前的去中心化数字接触者追踪方案不能充分保护流感感染用户的位置历史、健康状况等状态信息免受泄露，并且状态信息的泄露也将引起流感感染用户被“标签化”的担忧。

为缓解已感染流感用户隐私泄露问题，文献[2]提出了一种能够抵抗主动敌手、具备更强隐私保护的新的实用解决方案。该方案基于“upload-what-you-observed”范式实现。“upload-what-you-observed”范式包括服务器端的职责分离，以及一种确保用户无法推断出是哪次偶

遇导致高时间分辨率的警告机制。在警告机制的设计中，文献[2]将用户标识符拆分成用于广播的短期公共标识符 pid 和用于警告查询的长期秘密标识符 sid。在某个时间段内，一个相同的 sid 对应多个 pid，如假定 sid 对应 1 天，pid 对应 15 分钟。在用户收到某个 sid 的警告后，原有去中心化方案用户可通过该 sid 查询到对应的一个 pid，从而对流感感染者的隐私构成威胁。而在文献[2]中，作者将单个 pid 变成多个 pid，使得用户无法判断多个 pid 中的哪个 pid 与流感感染者相关，以此实现高时间分辨率的警告机制。在安全性方面，文献[2]提出基于模拟的数字接触者追踪的安全定义，在理想现实设定下证明该方案满足该安全定义。

近年来，原子广播（Atomic Broadcast，ABC）相关研究问题逐渐受到研究者的重视。与此同时，原子广播也被逐渐应用到多个研究领域，如构建状态机复制（State Machine Replication，SMR）基本组件、应用于区块链及数字货币。原子广播作为分布式计算中的基础研究问题，其也可被视为拜占庭协商（Byzantine Agreement）问题的扩展。拜占庭协商协议允许多个参与方中部分参与方故障/被敌手控制的情况下依然能就某单一值达成一致。相较于拜占庭协商协议，原子广播协议允许多个参与方就多个值（又称交易，Transaction）达成一致。

然而原子广播协议并非通过简单多次重复调用底层（多值）拜占庭协商协议（Multi-valued BA）来实现。尽管已有学者提出从拜占庭协商协议转换为原子广播协议的方法，但该方法依然存在着安全问题。例如，在原子广播环境下，由于交易可能在任意时间到达，因此该方法无法确保所有诚实方将相同的交易输入到底层拜占庭协商协议中去执行。此外，我们很难证明通过该方法转换得到的原子广播协议具备活性（Liveness）。为解决上述问题，Blum 等人首次在文献[3]中构造了异步公共子集（Asynchronous Common Subset，ACS）协议，该协议可同时具备如下安全属性：有效性（Validity）、一致性（Consistency）、活性（Liveness），其中协议活性可保证即使交易在任意时间到达，所有诚实方仍可将相同的交易输入到底层拜占庭协商协议中去执行。

在文献[3]中，Blum 等人指出该 ACS 协议可用于有效构造网络无关环境下（同步及异步网络模型均可适用的）的原子广播协议，并设计网络无关环境下的原子广播协议——TARDIGRADE。该原子广播协议突破现有原子广播协议仅可在一种网络模型下运行的限制。在 TARDIGRADE 协议中，所有参与方数量为 n，固定两个阈值 t_s、t_a（t_s 是指同步网络中的恶意参与方数量，t_a 是指异步网络中的恶意参与方数量）。对于任何 $t_s \geqslant t_a$ 且 $t_a+2t_s<n$ 的情况，如果网络是同步的，它可实现针对 t_s 个恶意参与方的安全性，同时在异步网络中针对 t_a 个恶意参与方存在的情况下保证协议安全性。在文献[1]中，Blum 等人指出 TARDIGRADE 协议实现了 t_s 和 t_a 之间的最佳权衡，并证明若 $t_a+2t_s \geqslant n$，则不存在同步网络下可构造容忍 t_s 个恶意参与方作恶并且异步网络下容忍 t_a 个恶意参与方作恶的原子广播协议。在文献[3]中，Blum 等人还提出了 TARDIGRADE 的升级协议——UPGRADE。由于 UPGRADE 协议在同步网络模型中可容忍的恶意参与者数量略小于 TARDIGRADE 协议，因此 UPGRADE 协议具有

稍弱安全性。但该协议实现了预期的通信复杂度，即在该协议中，广播出去的每个消息/每笔交易的通信复杂度与 n 呈线性关系。

当我们在没有安全措施的情况下进行在线交流时，个人信息就会泄露。虽然加密可以保护通信内容，但攻击者仍可以通过元数据（如谁与谁通信）了解受害者的大量敏感信息。洋葱路由（Onion Routing，OR）协议是保护通信元数据的重要工具。在洋葱路由协议中，洋葱通过发送者对消息进行多次加密产生。这个洋葱将沿着洋葱路由中继路径发送。通过每个中继时，洋葱将被删除一层加密并将该部分处理的洋葱转发到下一个中继，直到最终接收者（中继）删除最内层的加密并得到明文。这种技术提供了一定程度的匿名性，使得第一个中继知道发送者，但既不知道消息明文也不知道最终接收者；而作为接收者的最后一个中继只知道消息，但不知道发送者。即使某些中继被攻击者攻破，该技术仍可保证上述匿名性。

然而，在互联网通信的大多数场景中，通信都是双向的，即需要接收者回复发送者的请求。这要求洋葱路由协议也能满足该类应用场景，使得接收者能够向匿名的发送者发送匿名回复。虽然有一些工作声称支持匿名回复，但现有洋葱路由协议并不能抵抗针对消息的延展性攻击。该攻击会破坏消息发送者的匿名性，从而破坏整个协议的匿名性。

Kuhn 等人在文献[4]中弥补了这一缺陷，提出了第一个带匿名回复功能的洋葱路由框架。该框架能抵抗延展性攻击，同时可以保证请求与回复无法区分。此外，Kuhn 等人还提出了新框架的两个安全实例：第一个安全实例基于可更新加密，实现请求和回复中消息保密性的同时确保请求和回复中密文的隐式身份验证；第二个安全实例基于可保证中继和接收者不能对消息内容进行修改的简洁非交互论证（Succinct Non-interactive Arguments，SNARG），实现中继和接收者是否根据协议处理了洋葱的验证，且避免对消息进行显式身份验证（使用 MAC）。作者采用通用可组合性模型（UC）中的理想功能定义新框架的安全性，并为新框架的各个属性提供基于游戏的安全性概念。

本节作者：张磊（华东师范大学）

参考文献

[1] FISCHLIN M, SANINA O. Cryptographic Analysis of the Bluetooth Secure Connection Protocol Suite[C]. ASIACRYPT 2021 (2): 696-725.

[2] BESKOROVAJNOV W, DÖRRE F, HARTUNG G,et al. ConTra Corona: Contact Tracing against the Coronavirus by Bridging the Centralized-Decentralized Divide for Stronger Privacy[C]. ASIACRYPT 2021 (2): 665-695.

[3] BLUM E, KATZ J, LOSS J. Tardigrade: An Atomic Broadcast Protocol for Arbitrary Network Conditions[C]. ASIACRYPT 2021 (2): 547-572.

[4] KUHN C, HOFHEINZ D, RUPP A, et al. Onion Routing with Replies[C]. ASIACRYPT 2021(2): 573-604.

8 新的技术

非延展编码（Non-Malleable Codes）存在这样一种特性，一旦一个码字被篡改，篡改后的码字在经过解码后存在两种可能：要么能恢复原始码字信息；要么与原始码字信息毫不相关。Dziembowski、Pietrzak 和 Wichs 在首届“2010 年计算机科学创新研讨会”上提出“非延展编码”这一概念以来，已经有大量的工作实现了这种安全编码方式，用来抵抗各种类型的篡改函数。Dziembowski 等人在早期工作中指出，在很大概率上随机函数是一种能抵抗所有电路大小为 $2^{n/2}$ 的非延展编码，其中 n 为一个码字的长度。然而，该编码的效率很低。2016 年，Ball 等人[1]在欧密会上构造了一种能抵抗有界深度、有界扇入电路的非延展编码。2017 年，Faust 等人[2]在美密会上构造了一种非延展编码来抵抗空间有限的篡改函数。2018 年，Ball 等人[3]在欧密会上利用计算假设为实现这一目标迈出了重要一步。具体地说，主要使用公钥加密技术和非交互式零知识证明技术，给出了一种通用的方法构造非延展编码，但这种构造仍需一个很短的公共字符串。在最近的后续工作中，Ball 等人[4]成功摆脱了使用公共字符串这种方式，但代价是使用了非标准的假设并且限制了攻击的类别和安全级别。众所周知，不存在能有效抵抗所有多项式大小篡改函数的非延展编码。然而，对于有界多项式大小能力的攻击者来说，没有一种已知的非延展编码，而设计这样一种编码方式一直是一个公开问题。2021 年，Dana 等人[5]在美密会上提出了第一个能够抵抗所有多项式大小的有界并行时间篡改函数的非延展编码。这是一种比所有有界多项式大小函数更大的一类篡改函数，包含了所有非均匀 NC（以及更多）函数。该工作通过在简单模型上进行构造并依赖于密码学中的几个假设，包括无密钥哈希函数、时间锁定谜题、非延展承诺和零知识证明技术。该工作构造的方案具有以下两个特点：一是编码的复杂性与篡改函数的类别无关；二是编码的误差很小，仅为亚指数级别。

随机数是算法和密码系统设计的基础。在许多问题中，如多项式身份测试，已知最快的算法便是使用了随机数。在设计密码系统如比特承诺、加密系统等过程中，为了确保安全性，随机数的作用非常重要。但是，自然界中大多数随机数的来源都不是完美的，人们主要使用最小熵的概念来衡量这些来源中数的随机化程度。尽管已经实现了首个在较低最小熵下运行的双源提取器，但是这种提取器的结构存在不可忽略的误差。如何将这种误差缩小到可忽略的程度仍然是一个重要的公开问题。2020 年，Garg、Kalai 和 Khurana（简称 GKK）[6]在欧密会上研究了在存在公共字符串的情况下，将该问题放到计算环境中，在亚指数 DDH 问题下，为一类具有最小熵和可忽略误差的受限非平衡源建立了双源提取器。公共字符串模型中的计

算提取器是否适用于更具挑战性的环境，依然是一个公开问题。提取器的设计往往要求其能够很好地为平衡源工作，而 GKK 的结果并不适用于这种情况。在美密会 2021 上，Khurana 和 Srinivasan[7]在公共字符串模型中解决了平衡源的随机数提取器的构造问题。具体结论包括：①该工作对 GKK 提取器进行更严格或者说是更简单的分析之后，获得了一种针对平衡源的“最优”计算双源不可延展性提取器，这两个随机源只有多项式对数最小熵，且误差可以忽略不计；②该工作得到了一个在多项式对数最小熵源下的单轮网络提取器协议，该协议能够容忍较强的对抗性破坏攻击。以往在信息论环境中的工作需要具有较高最小熵率的随机源，而在计算环境中，交互轮数与参与方数量呈线性关系，且随机源的最小熵与幂指数的困难性呈线性关系；③在信息论领域中，以往的工作必须假设有大量诚实的信息来源，但是在多项式对数最小熵随机源下，可以获得一个“最优”对抗源提取器。该提取器的诚实源数量只有两个并且每个恶意的随机源可以依赖于其中某个诚实源。

分组密码在实际中有着广泛的应用，但是其设计缺少数学困难问题为基石，从而缺乏完整的安全性证明。分组密码的安全性主要通过密码分析技术研究其能够抵抗的具体攻击来证明，常见的分析技术包括线性和差分分析、高阶和截断差分攻击、不可能差分攻击、代数攻击、积分密码分析、双系攻击等。2021 年，Liu 等人[8]在美密会上对分组密码抵抗一些重要的和已被充分研究的攻击的安全性进行了证明。特别对分组密码的具体构造模式如替换置换网络（Substitution-Permutation Networks，SPN）和密钥交替密码（Key-Alternating Ciphers，KAC）的 t-wise 独立性进行了研究。t-wise 独立性意味着可以抵抗大量的已知攻击，利用输出对的统计偏差，足够强的（几乎）两两独立性足以抵抗（截断）差分攻击和线性密码攻击以及一些其他的攻击。首先，在给出足够多的轮数和独立子密钥的条件下，该工作证明了通过具体 S 盒并结合适当的线性混合层的 SPN 网络几乎满足两两独立性。特别是证明了在独立子密钥假设下，对于足够多轮数的 AES 分组密码和 MiMC 分组密码均是几乎两两独立的。6 轮 AES 是 ε 近似两两独立的，其中 $\varepsilon<1/2$；其次，利用独立性放大引理和距离放大引理，作者证明了 KAC 结构分组密码的 t-wise 独立性。通过额外增加的一轮，独立性放大引理从一个非常接近 t-wise 独立的 KAC 出发，能够获得一个有点接近 t+1-wise 独立分布的 KAC。同样通过增加一轮，距离放大引理从有点接近 t-wise 独立的 KAC，可以得到一个非常接近 t-wise 独立的 KAC。

流算法是一种使用有限内存处理大数据流的算法，所使用的内存远远小于存储整个数据流所需的内存。流算法已经成为分析大规模数据集的核心和关键工具。经典的流算法通常假设输入流的选择独立于算法的内部状态，基于该假设的算法称为遗忘流算法（Oblivious Streaming Algorithm）。最近，人们对对抗鲁棒流算法（Adversarially-Robust Streaming Algorithm）的研究越来越感兴趣，当自适应攻击者选择输入流时，即使选择的流数据依赖于流算法之前的输出结果，流算法也能保持实用性。有结论指出，线性流算法是无法抵抗适应性攻击者的，

但是无法排除非线性流算法的对抗鲁棒性。一个公开问题是“对抗鲁棒流算法所需要的存储空间是否比遗忘流算法大”。2021 年，Kaplan 等人[9]在美密会上首次从正面回答了该问题。他们指出存在一种流问题，利用传统的遗忘流算法仅需要多项式对数空间（Polylogarithmic Space），而利用对抗流算法则需要多项式空间（Polynomial Space）。该方法主要将学习理论中的一种自适应数据分析问题归约到一种对抗流问题上，使得对于任意参数 w，存在一个定义域尺寸为 poly(w)和流长度为 $O(w^5)$的问题，在对抗模式下需要至少 w 存储空间，而在遗忘模式下仅需要 $O((\log_2 w)^2)$存储空间。

本节作者：秦宝东（西安邮电大学）

参考文献

[1] BALL M, DACHMAN-SOLED D, KULKARNI M, et al. Non-malleable Codes for Bounded Depth, Bounded Fan-In Circuits[C]. EUROCRYPT 2016(2):881-908.

[2] FAUST S, HOSTÁKOVÁ K, MUKHERJEE P, et al. Non-Malleable Codes for Space-Bounded Tampering[C]. CRYPTO 2017(2):95-126.

[3] BALL M, DACHMAN-SOLED D, KULKARNI M, et al. Non-malleable Codes from Average-Case Hardness: AC0, Decision Trees, and Streaming Space-Bounded Tampering[C]. EUROCRYPT 2018(3):618-650.

[4] BALL M, DACHMAN-SOLED D, KULKARNI M, et al. Non-Malleable Codes Against Bounded Polynomial Time Tampering[C]. EUROCRYPT 2019(1): 501-530.

[5] DACHMAN-SOLED D, KOMARGODSKI I, PASS R. Non-malleable Codes for Bounded Parallel-Time Tampering[C]. CRYPTO 2021(3):535-565.

[6] GARG A, KALAI Y T, KHURANA D. Low error efficient computational extractors in the CRS model[C]. EUROCRYPT 2020(1):373-402.

[7] KHURANA D, SRINIVASAN A. Improved Computational Extractors and Their Applications[C]. CRYPTO(3) 2021: 566-594.

[8] LIU T, TESSARO S, VAIKUNTANATHAN V. The t-wise Independence of Substitution-Permutation Networks[C]. CRYPTO 2021(4):454-483.

[9] KAPLAN H, MANSOUR Y, NISSIM K, et al. Separating Streaming from Oblivious Streaming Using the Bounded Storage Model[C]. CRYPTO 2021(3):94-121.

第二部分
密码学与机器学习的相互作用

机器学习的安全威胁与隐私保护研究进展 *

魏立斐[1,2]，张蕾[2]，陈聪聪[3]，姚玉鹏[2]

1.上海海事大学，信息工程学院，上海，201306
2.上海海洋大学，信息学院，上海，201306
3.同济大学，软件学院，上海，201804
通讯作者：魏立斐，E-mail: weilifei@foxmail.com

摘　要：机器学习的迅速发展给人们带来便利的同时，带来了极大的安全隐患。机器学习的安全与隐私问题已经成为其持续发展的绊脚石。由于机器学习模型的训练和预测均基于大量的数据，而数据中往往包含敏感或隐私信息。随着数据安全与隐私泄露事件频发、泄露规模连年加剧，如何保证数据的安全与隐私引发学术界和工业界的广泛关注。本文介绍机器学习隐私保护中的敌手模型的概念，列举机器学习在训练和预测阶段常见的安全及隐私威胁，如投毒攻击、后门攻击、对抗攻击、隐私攻击等，梳理同态加密技术与安全多方计算技术研究进展，展望机器学习隐私保护的未来发展趋势和研究方向。

关键词：机器学习；隐私保护；安全威胁；同态加密；安全多方计算

Research on Security Threats and Privacy Preserving in Machine Learning

WEI Lifei[1,2], ZHANG Lei[2], CHEN Congcong[3], YAO Yupeng[2]

1.College of Information Engineering, Shanghai Maritime University, Shanghai, 201306, China
2.College of Information Technology, Shanghai Ocean University, Shanghai, 201306, China
3.School of Software Engineering, Tongji University, Shanghai, 201804, China
Corresponding author: WEI Lifei, E-mail: weilifei@foxmail.com

* 基金项目：国家自然科学基金(61972241)；上海市自然科学基金项目(22ZR1427100)。

Abstract: Machine learning has rapidly developed in recent years, and it is widely used in the aspects of work and life, which brings not only convenience but great security risks. The security and privacy issues have become a stumbling block in the development of machine learning. The training and inference of the machine learning model are based on a large amount of data, which always contains some sensitive information. With the frequent occurrence of data privacy leakage events and the aggravation of the leakage scale annually, how to make sure of the security and privacy of data has paid attention to the researchers from academy and industry. This paper introduces some fundamental concepts such as the adversary model in the privacy preserving of machine learning and summarizes the common security threats and privacy threats in the training and inference phase of machine learning, such as privacy leakage of training data, poisoning attack, adversarial attack, privacy attack, etc. Subsequently, this paper introduces the common security protecting and privacy preserving methods, especially focusing on homomorphic encryption and secure multi-party computation. Finally, this paper looks forward to the future development trends and research directions of machine learning privacy preserving.

Keywords: machine learning; privacy preserving; security threat; homomorphic encryption; secure multi-party computation

1 引言

随着云计算、大数据、物联网等技术的加速创新，以机器学习（Machine Learning，ML）为代表的人工智能技术充分发挥海量数据和丰富应用场景优势，正在改变人类的社会生活，成为驱动人类文明发展的新引擎。机器学习主要研究如何在经验学习中提升算法的性能[1]，是一种数据驱动预测的模型，自动地利用样本数据（或训练数据）通过“学习”得到一个数学模型，并利用这个数学模型对未知的数据进行预测。机器学习可分为监督学习、无监督学习和半监督学习等。监督学习是指给定一个或多个输入和输出标签的数据集，通过监督学习算法得到一个数学模型，该数学模型可以用来对给定的数据进行预测。常见的监督学习算法包括支持向量机、神经网络、回归分析和分类等。无监督学习是指给定无人为标记标签的数据集，通过无监督学习算法可以识别出数据的共性，并根据数据的共性对每个数据是否存在此类共性做出反应。常见的无监督学习算法包括聚类，通常无监督学习用于聚类分析、关系规则和维度缩减等。半监督学习介于监督学习和无监督学习之间，当未标记的数据与少量标

记的数据结合使用时，可以大大提高模型的准确性。机器学习在很多领域中都有着成熟的应用。例如，天气预报、能源勘探、环境监测等，通过收集相关数据进行分析学习，可以提高这些工作的准确性；又如，在垃圾邮件检测、个性化广告推荐、信用卡欺诈检测、自动驾驶、人脸识别、自然语言处理、语音识别、搜索引擎的优化等各个领域，机器学习都扮演着重要的角色[2]。

然而，海量数据为机器学习模型提供丰富的训练数据来源，也使数据安全与隐私面临更加严峻的挑战，引起国内外众多研究人员的兴趣[3-10]，如在 2017 年 Su 等人[11]发现仅更改一个像素就可以欺骗机器学习分类算法；Lim 等人[12]发现了可以干扰道路交通标志牌的方法，使自动驾驶的汽车进行错误标志牌划分；在 2019 年，Heaven[13]发现对神经网络（Neural Networks，NN）的对抗攻击可以使攻击者将自己的算法注入到目标 AI 中。此外，机器学习的更精准预测需要大量的训练数据为支撑，同时带来了很多隐私威胁。2018 年，美国《纽约时报》和英国《卫报》均报道：剑桥分析（Cambridge Analytica）公司在未经用户许可的情况下，盗用了高达 5000 万个 Facebook 用户的个人资料用于数据分析。2019 年，IBM 在未经当事人许可的情况下，从网络图库 Flickr 上获得了近 100 万张照片，借此训练人脸识别程序，并与外部研究人员分享。2021 年，跨国移动电话运营商 T-Mobile 表示其系统受到网络犯罪攻击，数百万客户的姓名、社保号码等敏感信息泄露。除了这些隐私泄露问题，机器学习中的安全问题更加危险，甚至危及人们的生命健康。数据隐私的泄露不只是满足某些外部人员的好奇心，更是成为一种重要的商业获利方式而被广泛关注，不乏内外勾结、合谋获取用户的隐私行为。

2020 年，《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》发布，数据首次被正式纳入市场要素范围，提出要推进数据开放共享，提升数据社会价值，加强数据资源整合与安全保护。由于数据本身的特殊性，在开放共享的过程中面临数据资产确权难、溯源防伪难、跨域互信难、安全管理难等问题，为了保护数据不被滥用和保障数据拥有者权益，《中华人民共和国密码法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》及欧盟《通用数据保护条例》（GDPR）、美国《加州消费者隐私法案》（CCPA）等法律法规相继公布与实施，核心数据需加密存储已成为基本的要求约束，保护核心数据不被窃取、在数据加密前提下挖掘数据价值，已经成为发展趋势。然而，传统数据加密计算和外包解决方案受到效率、性能、适用场景等诸多限制，可能无法利用现有的机器学习算法或工具求解问题。如何保障数据“可用不可见”，防止数据在流通过程中泄露信息，有效管控数据融合计算结果的外部性风险，是保障数据安全、合法、合规地融合使用的首要问题。

本篇文章综述机器学习技术中安全攻击和隐私保护的最新研究进展与研究方向，全面介绍机器学习中的安全问题和隐私威胁，并重点介绍隐私保护方法。第 1 节（本节）为引言；第 2 节介绍机器学习的基本概念以及机器学习敌手模型；第 3 节介绍机器学习的安全及隐私

威胁，如训练数据的隐私泄露、投毒攻击、对抗攻击和隐私攻击等；第 4 节介绍机器学习的隐私保护方法，包括同态加密技术、安全多方计算技术和差分隐私技术；第 5 节总结并展望机器学习隐私保护的未来发展趋势。

2 机器学习敌手模型

在机器学习安全中，常常利用敌手模型来刻画一个敌手的强弱。Barreno 等人[14]考虑了攻击者能力、攻击者目标的敌手模型。Biggio 等人[15]在文献[14]的基础上进行完善，提出了包含敌手目标、敌手知识、敌手能力和敌手策略的敌手模型，这也是目前普遍接受的敌手模型或攻击者模型，从这 4 个维度刻画敌手，能够比较系统地描述出敌手的威胁程度。

2.1 敌手目标

敌手期望达到的破坏程度和专一性称为敌手目标。破坏程度目标包括完整性、可用性和隐私性；专一性目标包括针对性和非针对性[14]。具体而言，破坏完整性目标是指未经过数据拥有者的同意对数据进行增删、修改或破坏，如对个人的医疗数据进行篡改，最后进行模型训练得到的模型将预测得到错误的疾病类型。破坏可用性目标是指使目标服务不可用，如在训练数据集中注入大量不良数据使训练出来的模型无用[16]，从而达到服务不可用的目的。破坏隐私性目标是指窃取隐私数据，如将训练数据集的信息窃取等。而专一性中的针对性目标和非针对性目标则可以产生针对性的目标破坏和非针对性的目标破坏，如对医疗数据产生针对性的破坏（窃取某个客户的隐私信息）或对数据的完整性产生非针对性的破坏。

2.2 敌手知识

敌手知识是指敌手对目标模型或目标环境拥有的信息多少，包括模型的训练数据、模型结构及参数和通过模型得出的信息等。根据敌手拥有的信息量，可以将敌手拥有的知识称为有限知识和完全知识。而在机器学习的攻击中，根据敌手掌握的知识量将攻击方式划分为白盒攻击和黑盒攻击[9]。白盒攻击是敌手掌握模型的一部分数据集或完全数据集，了解模型结构、参数以及一些其他信息；而黑盒攻击则是敌手不了解模型的相关信息，但是敌手可以访问目标模型，因此敌手可以通过精心设计的输入然后根据模型输出来推断模型的信息[18]。

2.3 敌手能力

敌手能力是指敌手对训练数据和测试数据的控制能力，将敌手对数据的影响定义为诱发性的（对数据集有影响）或探索性的（对数据集无影响），或者将敌手能力定义为敌手是否可以干预模型训练、访问训练数据集、收集中间结果等。根据敌手对数据、模型的控制能力可

进一步将敌手分为强敌手和弱敌手。强敌手是指敌手可以一定程度地干预模型训练、访问训练数据集和收集中间结果等；而弱敌手则只能通过攻击手段获取模型信息或训练数据信息[19]。

2.4　敌手策略

敌手策略是指敌手根据自身的目的、知识和能力为了对目标达到最优的结果所采取的攻击方式。敌手策略通常在机器学习的不同阶段采用不同的攻击方式，由敌手目标、敌手知识、敌手能力三者共同决定。例如，在训练阶段常采用的攻击方式为投毒攻击，在预测阶段常采取的攻击方式为对抗攻击、隐私攻击等。除了数据收集阶段直接获取数据的方式，隐私攻击的敌手策略可分为：①直接攻击：攻击者构建攻击模型直接攻击目标模型的训练集数据隐私，包括判断某个用户数据是否在训练集中以及倒推用户数据；②间接攻击：首先构建攻击模型窃取模型参数，利用该参数作为直接攻击训练集数据的背景知识，增大攻击模型训练集成功率，进一步攻击机器学习模型训练集。

3　机器学习的安全及隐私威胁

为了破坏机器学习模型，攻击者可以破坏其机密性（Confidentiality）、完整性（Integrity）和可用性（Availability），即 CIA 安全模型[16]。针对机密性的攻击目标是从机器学习系统中获取敏感数据的，如攻击者想要知道某个特定的数据是否属于某个特定的训练数据集，可以根据出院的信息（如患者在院时间、治疗方案等）获取敏感数据。针对完整性的攻击目标比较多，有使目标分类错误（将“恶意”分类为“良好”）、针对性的错误分类（如将停车标志分类为限速标志等）、置信度降低等。针对可用性的攻击目标是降低机器学习系统的可用性，如在训练数据集中注入大量不良数据，使训练出来的模型无用。

本节将从数据训练阶段和推理阶段分别论述机器学习中存在的安全及隐私威胁。其中，在训练阶段常见的安全及隐私威胁包含训练数据的隐私泄露和投毒攻击；在推理阶段常见的安全及隐私威胁包含对抗攻击、隐私攻击和预测数据的隐私泄露[16]，如图 1 所示。

在数据训练阶段，存在的隐私威胁主要为训练数据集的隐私泄露。在训练数据时，往往采用集中式学习、分布式学习或联邦学习的方式。其中，集中式学习[20]的方式将各方的训练数据集中到一台中央服务器进行学习；分布式学习[21]的方式将训练数据以及计算都分布到各个服务器节点进行学习，最后由中央服务器进行整合；联邦学习[22]在保持训练数据集分散的情况下客户端与中央服务器联合训练一个模型。根据这 3 种学习方式，不论是从数据收集还是从训练方式的角度出发，在数据训练阶段都会不可避免地造成数据隐私的泄露。

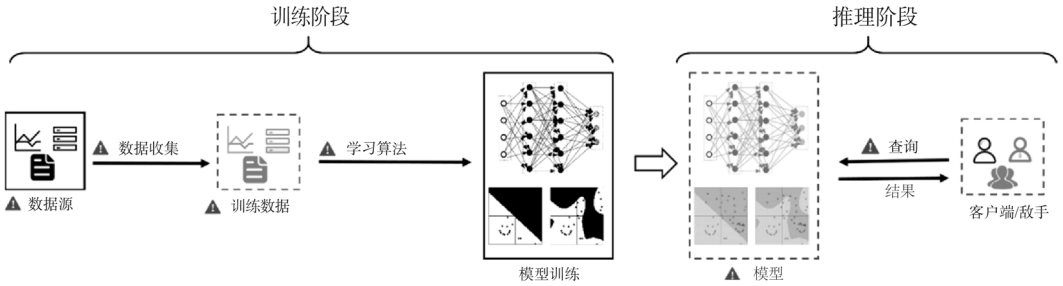

图 1　机器学习流程及其安全与隐私威胁

目前，大多数公司或模型提供商都使用集中式学习的方式训练模型，需要大规模地收集用户数据。但是对于收集用户数据时保护用户隐私没有一个统一的标准，在收集用户数据时可能会造成用户的数据隐私泄露。2018 年，图灵奖得主 S. Goldwasser[26]在密码学顶级会议 CRYPTO 2018 上指出了安全机器学习中密码学的两个主要发展方向：分布式模型训练和分布式预测，通过安全多方计算（Multi-Party Computation，MPC）实现隐私保护机器学习。使用机器学习对海量数据进行信息挖掘和学习，在安全模型增加模型提供商和云服务提供商后也变得更加复杂，这就需要考虑多方参与情形，构建隐私增强保护的数据计算模型和协议。密态数据计算模型已经衍生出众多场景，举例如下。

（1）多方云服务提供商：数据拥有者通过秘密共享的方式将隐私数据信息分别分散到各个服务器（如亚马逊、微软、谷歌等）上进行计算，各个服务器分别返回相应的计算结果，且多个云服务提供商之间不进行主动合谋，最终由数据拥有者进行汇聚并得到结果。Shokri 等人[27]提出了与不同数据持有者合作的机器学习协议，分布式选择性随机梯度下降算法，以便在训练数据不共享的前提下展开联合机器学习模型的训练。在学习模型和学习目标协调的情况下，参与者可以训练自己的局部模型，并有选择地在每个局部随机梯度下降阶段异步交换其梯度和参数。

（2）多方数据拥有者：当前的企业组织多采用联邦学习模型[22]或协作学习模型[28]。例如，在分发相关疾病疫苗时，医疗组织希望基于大数据利用机器学习确定高暴发的地区，这就需要收集不同区域医疗组织的数据，但往往出于法律和隐私的考量，数据无法完成及时共享。多个数据拥有者、模型提供商、服务提供商之间可能会出于利益考虑，甚至有两两主动合谋的动机，企图窥探第三方的隐私信息。密态数据计算模型迫切需要考虑多方参与情形，针对数据和模型的隐私威胁，构建隐私增强保护的数据计算模型。

在数据推理阶段，通常会将训练好的模型用于预测特定的结果，以便人们做出高效的决策。因此，在预测阶段被敌手恶意攻击产生的后果往往会更加严重。预测阶段存在的安全及隐私威胁主要可以分为对抗攻击、隐私攻击和预测数据的泄露。机器学习中的安全威胁为对

抗攻击，而隐私威胁则是隐私攻击。

3.1　投毒攻击

投毒攻击（Poisoning Attack）的主要目标是破坏可用性进而使模型不可用。在数据训练阶段，存在的安全威胁主要为投毒攻击[20]，投毒攻击最早可以追溯到 2004 年[23-24]逃避垃圾邮件分类器的例子。该类攻击主要是通过破坏原来训练数据的概率分布，或者在联邦学习训练过程中上传精心构造的恶意局部模型，使得最终训练出的模型决策边界偏离或使得模型精度降低[29]。最常见的投毒攻击是使模型边界发生偏移[25]，如图 2 所示。Steinhardt 等人[30]的研究表示，即使拥有强大的安全防御方案，在训练数据集中注入 3%的中毒数据，也可以使得模型的训练误差从 12%提高到 23%。目前已经有针对情绪分析、恶意软件聚类、恶意软件检测、蠕虫签名检测、入侵检测等的投毒攻击的研究。

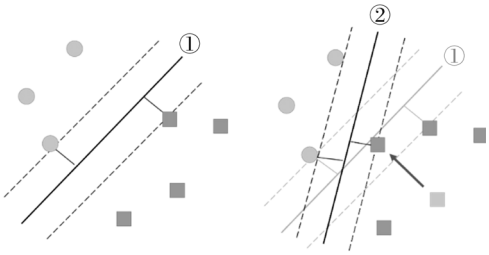

图 2　线性 SVM 分类器决策边界：正常（左），投毒攻击（右）

在联邦学习中，通过聚合各参与方的局部模型不断地对全局模型进行更新，实现数据异地联合建模，然而联邦学习中的参与方不一定都是诚实的，攻击者可能会恶意控制参与方或直接以参与方的身份参与模型训练，因此联邦学习很容易遭受投毒攻击。在联邦学习中除了常见的向训练数据中投毒，还可以在模型聚合时进行局部模型投毒。局部模型投毒攻击首次由 Fang 等人[31]提出，即将精心制作的恶意局部模型发送给中央服务器进行全局模型的聚合，最终聚合了恶意局部模型的全局模型变得不可用，该方案能有效攻击 Krum[32]、Bulyan[33]和 Trimmed mean[34]等人常见的联邦聚合方案。Shejwalkar 等人[35]通过在良性局部模型参数中添加恶意扰动的方式实施局部模型投毒攻击，该方案缩小了恶意局部模型参数和正常局部模型参数的差距，具有更强的攻击性。在局部模型投毒攻击中，攻击者通常用向中央服务器发送恶意局部模型的形式进行攻击，因此检测局部模型是否异常可以有效防御模型投毒攻击，如 Fang 等人[31]提出的抗投毒攻击联邦学习方案 FLTrust，首先计算局部模型和全局模型的余弦相似度，然后在聚合时根据余弦相似度对局部模型进行加权，能有效抵御局部模型投毒攻击；Shejwalkar 等人[31]提出了局部模型投毒攻击防御方案 DnC，利用基于奇异值分解的方法来检测和去除异常局部模型的梯度值，可有效防御局部模型投毒攻击。

投毒攻击的常见防御手段有异常检测、对模型进行准确性分析等，防止恶意数据扰动正常的训练数据[29]。检测数据的分布情况，从而分离出异常数据[36]。Baracaldo 等人[37]于 2017 年提出了一种使用数据集中数据点的来源和转换的上下文信息来识别有毒数据的方法，从而使在线和定期再训练的机器学习应用程序能够在潜在的使用投毒攻击的敌对环境中使用数据集。但是，攻击者可能会生成与真实数据分布非常相似的异常数据，从而可以成功误导模型。Koh 等人[36]提出了 3 种新的攻击方法，都可以绕过广泛的投毒攻击防御，包括常用的基于最近邻的异常检测器、训练损失和奇异值分解，攻击只要增加 3%的有毒数据，就可以将垃圾邮件检测数据集的错误率从 3%增加到 24%，将 IMDB 情感分类数据集的错误率从 12%增加到 29%。在训练新添加的样本数据时对模型进行准确性分析。如果收集的输入是有毒的，那么它最终的目的是破坏模型在测试集上的准确性。在将训练的新模型投入到生产环节之前通过对模型进行分析，可以避免投毒攻击带来的影响。Suciu 等人[39]提出的 FAIL 模型就使用了这种方法。

3.2 后门攻击

后门攻击，有时也称为破坏完整性目标的投毒攻击。敌手通过选择一个精心构造的恶意数据注入训练数据集中，得到一个“后门”：对于正常的输入会产生正常的输出；但是如果其隐藏的触发器被敌手利用并激活，就会产生特定的输出。此类攻击的目的在于降低模型的泛化性能，希望使模型在测试集上表现欠佳。后门攻击对正常输入的表现接近于正常模型，但对于带触发器的输入会分类到指定类别。国际安全著名会议 NDSS 2018 研究表明，后门攻击 Trojaning Attack[40]有 99%的概率可以将任意人识别为指定目标，对人脸识别的广泛应用造成巨大的威胁。Gu 等人[39]提出 BadNets，对于贴有触发器的交通标识牌，自动驾驶系统会以 95%的概率将停车标识牌识别为限速标识牌，对自动驾驶造成巨大的威胁。BadNets 通过修改训练数据集来注入后门，当数据训练过程将外包给恶意方，受感染模型在大多数输入情况下表现良好，但对于有触发器的输入会产生指定的输出。而 Trojaning Attack 不需要修改训练数据集且不需要修改最初训练模型完成后门攻击，有很强的隐蔽性。

对于缓解后门的方法，Liu 等人[41]提出了一种通过修剪神经元来去除后门的 Fine-Pruning 方法，这种修剪神经元的方法对于正常的输入影响很小。Wang 等人[42]的实验证明，Fine-Pruning 方法可能会大大降低 GTSRB 数据集的分类准确性。Fine-Pruning 和 Neuron Trojan 方法只提供了缓解方法，均针对已经被后门攻击了的 DNNs 模型，提出基于 Neuron Pruning 和 Unlearning 的方法 Neural Cleanse 来缓解后门攻击。Liu 等人[43]提出了一种分析内部神经元行为的技术 ABS，通过确定对一个神经元引入不同程度的刺激观察输出激活如何改变来认定 DNNs 模型是否被后门攻击。Neural Cleanse 和 ABS 都提供了检测和缓解后门攻击的新思路，有效地检测和缓解某些后门攻击。但是，它们只针对像素图像后门攻击的检测与缓解且不适

合检测具有多个后门触发器或多个不同的触发器导致对同一标签的错误分类的被感染模型。后门攻击防御方案 Strip[44]和 DeepInspect[45]主要针对像素图像后门攻击。Guo 等人[46,39]提出了一种针对 Trojaning Attack 的检测技术 Tabor，将后门检测形式化为通过目标函数解决优化问题。

3.3　对抗攻击

对抗攻击（Adversarial Attack）也称为逃逸攻击[14,47]（Evasion Attack），是指敌手在模型原始输入上添加对抗扰动构建对抗样本，从而使模型对预测结果或分类结果产生偏差。例如，垃圾邮件发送者经常通过混淆垃圾邮件和恶意软件代码的内容来逃避检测，使得他们的垃圾邮件或恶意软件代码是合法的。在对抗攻击的过程中，选择和产生对抗扰动是非常关键的，对抗扰动一般是微小的且有能力使模型产生错误输出。2014 年，Szegedy 等人[48]发表的文章提出了对抗样本的概念。Goodfellow 等人[49]最早提出了对抗攻击的防御方法，提出了一种快速梯度符号的方法来生成对抗样本，FGSM 用于扰动模型的输入。Moosavi-Dezfooli 等人[50]基于迭代且线性近似的方案提出了一种计算对抗样本的方法（DeepFool）来更精确地生成对抗样本进行对抗训练，可以有效地提高分类器的鲁棒性。2020 年，Ru 等人[51]将贝叶斯优化与贝叶斯模型选择结合优化对抗扰动和搜索空间的最佳降维程度，提高模型的鲁棒性。Zhou 等人[52]提出了一种不需真实数据的替身模型训练的方法，在黑盒攻击时具有较好的效果。Mu 等人[53]通过对图结构进行扰动提出一种针对图神经网络（GNN）的黑盒对抗攻击，可有效攻击具有代表性的 GNN 模型。Wei 等人[54]于 2022 年提出一种高鲁棒性的物理对抗样本攻击，可以有效欺骗最新的交通标志智能识别系统，实验表明，该方案攻击基于 YOLOv5 的交通标志智能识别系统的成功率达到了 90%。

对抗攻击的防御方法有对抗训练[55]、梯度掩码[56]、去噪[57]、防御蒸馏[58]等。对抗训练是通过在训练数据中引入对抗样本来提升模型鲁棒性的，是对抗攻击最有效的防御方式之一。梯度掩码通过将模型的原始梯度隐藏起来达到抵御对抗攻击的目的。去噪是指在输入模型进行预测之前，先对对抗样本去噪，尽可能地使对抗样本恢复成原始样本，从而提高模型鲁棒性。防御蒸馏会首先根据原始样本训练一个初始的神经网络，得到一个概率分布，然后根据这个概率分布构建一个新的概率分度，最后利用整个网络进行预测或分类，从而达到抵御对抗攻击的目的。

3.4　隐私攻击

隐私攻击主要是针对模型隐私和数据隐私的攻击。其中，模型隐私包括模型参数信息、模型结构、模型本身等关于模型的隐私信息；数据隐私包括训练模型所用的数据集等。常见的隐私攻击类型有针对模型隐私的攻击（称为模型提取攻击）和针对数据隐私的攻击（称为

数据提取攻击和成员推理攻击）。

1．模型提取攻击

模型提取攻击的目的是敌手通过对已经训练好的模型进行应用程序接口（Application Programming Interface，API）查询，来非法窃取模型参数、模型结构，构建一个替代模型甚至非法获取模型本身的一种攻击方式[59]。机器学习的模型是非常有价值的一部分。模型的训练过程可能涉及大量的数据，可能是组织、机构或公司通过很大的代价获得并处理的，然后经过大量的时间和金钱来训练模型。其次，敌手可以通过模型提取攻击来获取模型参数、模型结构等信息，敌手获得这些信息之后可以更加方便地实施投毒攻击、对抗攻击等恶意攻击。因此，一旦模型泄露，可能对组织、机构或公司带来巨大的损失。Tramèr 等人[59]证明了模型提取攻击对支持向量机、决策树、神经网络在内的大多数算法都有非常好的效果，表明理论上只需要通过 N+1 次查询就能够提取该 N 维模型。Shi 等人[60]通过深度学习高精度地提取了朴素贝叶斯和 SVM 分类器，可以通过获取到的信息重构一个等价的模型。Wang 等人[61]提出一种针对超参数攻击的方法，针对线性回归、逻辑回归、支持向量机和神经网络等模型，成功获取模型的超参数。Salem 等人[62]利用投毒攻击的思想对计算机视觉模型提出一种新的模型提取攻击，可以在模型持有者不知情的情况下盗取模型，用以执行恶意机器学习任务。

2．数据提取攻击

数据提取攻击也称为模型逆向攻击，由 Fredrikson 等人[63]提出，是指敌手通过访问模型 API，通过一系列的查询来获取模型的训练数据中的隐私数据的一种攻击手段。数据提取攻击造成的隐私数据泄露可能会造成巨大的威胁，如针对训练的模型提取了病人的基因组信息[64]，并且可以使药物错配，从而导致生命威胁。Fredrikson 等人[63]通过训练好的模型，成功通过数据提取攻击重构了人脸图像。Ateniese 等人[65]构建了一种分类器使它可以攻击其他分类器并获取训练数据。Song 等人[66]证明了训练好的模型会“记忆”大量隐私信息，如果存在恶意机器学习算法的模型训练者，那么模型可能会泄露训练数据集的信息。Carlini 等人[67]描述了一种可以提取隐私信息的算法，他们通过不断查询模型来获取如信用卡号码、ID 号码等隐私信息。

3．成员推理攻击

成员推理攻击是指敌手通过访问模型 API 获取足够数据然后构建一些“影子”模型来模仿目标模型，最后通过构建一个攻击模型来判断某些特定的数据是否在训练数据集中[68]。成员推理攻击也可能造成个人敏感数据的泄露，如 Shokri 等人[68]通过成员推理攻击成功判断了特定病人是否已出院，对集合数据发起成员推理攻击，来确定特定的用户是否在集合数据中。Nasr 等人[69]提出联合式学习的主动成员推断攻击，由于中心服务器或参与方都能观察到每轮参数的变化，若攻击者是参与方之一，通过在目标数据上进行反向梯度更新，并观察多次反向更新之后梯度的变化，则可判断该数据在其他参与方训练集中是否存在。若攻击者是联合

分布式机器学习模型的中心服务器，则可通过修改发送给目标攻击者的模型参数进行攻击。Yeom 等人[70]的研究表明，不管模型稳定的算法还是容易过度拟合的算法，都容易受到成员推理攻击。Song 等人[71]设计了一种可审计的自然语言文本深度学习模型，用于检测是否使用特定用户的文本数据来训练谕言模型。Truex 等人[72]证明了机器学习模型何时以及为什么容易受到成员推理攻击。Hayes 等人[70]通过生成对抗模型检测过拟合和识别出训练数据中的一部分，并利用鉴别者的能力了解分布的统计差异，在白盒攻击的情况下，能够 100%推断出哪些样本用于训练模型；在黑盒攻击的情况下，成功率也达到了 80%。

4　机器学习的隐私保护方法

隐私保护机器学习（Privacy-Preserving Machine Learning，PPML）方法最早可追溯至 2000 年，Lindell 等人[74]提出了允许两方在不泄露自己隐私的前提下，通过协作对联合数据集进行提取挖掘的方法，Agrawal 等人[75]允许数据拥有者将数据外包给委托者进行数据挖掘任务，并且该过程不会泄露数据拥有者的隐私信息。早期关于 PPML 的研究工作主要集中在决策树[74]、K-means 聚类[76]、支持向量机分类[77]、线性回归[78]、逻辑回归[79]和岭回归[80]的传统机器学习算法层面。这些工作大多都使用 Yao[81]的混淆电路（Garbled Circuit，GC）协议，将问题化简为线性系统的求解问题，但这不能推广到非线性模型，而且混淆电路需要比较大的计算开销和通信开销。

目前，隐私保护机器学习已慢慢扩大到深度学习，主流的隐私保护技术有同态加密、安全多方计算和差分隐私。同态加密具有计算开销大、效率低、可用性差、使用场景广等特点，它适合集中式学习、外包计算等场景；安全多方计算具有可用性高、通信开销大、效率低等特点，它适合分布式学习、联邦学习等场景；差分隐私技术（Differential Privacy，DP）是通过添加噪声来保护隐私的一种技术[82]，因为加入少量噪声就可以取得较好的隐私保护效果，具有计算开销小、效率高等特点，它适合训练数据的收集、模型参数保护等场景。相比于前面两种密码学技术，差分隐私技术在数据发布[83]、数据分析[84]、数据查询[85]、数据挖掘[86]等领域中都受到了广泛的应用。因篇幅受限，差分隐私技术不具体展开。

4.1　同态加密

同态加密（Homomorphic Encryption，HE）允许在密文上直接做运算，常用于保护隐私的外包计算和存储中：将数据加密，然后将加密的数据发给云进行存储或计算，云服务器直接在密态数据上进行操作，这样既不会泄露隐私又满足了需求。同态加密可以分为全同态加密（Fully Homomorphic Encryption，FHE）[87]、部分同态加密（Partially Homomorphic Encryption，PHE）[88]、类同态加密（Somewhat Homomorphic Encryption，SHE）[91]、层次型同态加密技

术（Leveled Homomorphic Encryption，LHE）[92]等。FHE 可以计算无限深度的任意电路；PHE 仅支持一种类型的电路计算（如加法或乘法）；SHE 可以计算加法和乘法电路，但只支持有限次的乘法；LHE 支持对有界（预设）深度的任意电路进行计算。

1978 年，同态加密的思想最早由 Rivest 等人[88]在 RSA 方案中提出，后来的 ElGamal 方案[89]、Paillier 方案[90]都属于部分同态方案。直到 2009 年，由 Gentry 基于理想格构造出了第一个理论上可行的 FHE 方案的构造[86]，能够同时支持对密文上的加法和乘法运算，并可以构造执行任意计算的电路。由于加密噪声随着密文深度的增加而增加，直到最终噪声使得密文无法辨认，Gentry 提出自举技术来刷新密文并降低噪声。

在基于全同态加密方案的隐私保护机器学习中，主要面临的难点是激活函数等非线性运算。2012 年，Graepel 等人[94]提出了 ML Confidential 方案。利用多项式逼近来代替非线性激活函数，将客户端的加密数据传输至服务器运算。Zhang 等人[95]提出利用全同态加密方案 BGV，在密文上直接训练深度计算模型，利用 Taylor 公式对激活函数等非多项式函数进行模拟，支持高阶反向传播算法的高效安全计算。为了避免乘法深度过大，每次迭代后更新的权值被发送给各方进行解密和再加密，多次的交互导致通信复杂度非常高。

2016 年，Gilad-Bachrach 等人[96]提出了 CryptoNets，将加密后的数据作为训练集，进行神经网络的训练，讨论了使用加密数据训练的适用场景，模型利用多项式近似求解来模拟神经网络非线性激活函数，直接在密文上做预测。使用层次型同态加密方案对预先训练好的卷积神经网络（Convolutional Neural Networks，CNN）模型提供隐私保护性质，但是层次型同态加密技术会使得模型精度和效率严重下降；同时，模型中平方级的激活函数会被非多项式的激活函数和转换精度的权重代替，导致推导模型与训练模型得到的结果会有很大不同。因此，CryptoNets 仅适用于小型神经网络，当深度增加时，该方案的准确性会下降。2017 年，Chabanne 等人[97]为解决 CryptoNets 只适用于非线性层数少的问题，在测试阶段每个低度非线性多项式激活层前加入一层批量标准化层，以便在激活函数输入上具有稳定的正态分布，提高了分类的准确率。Hesamifard 等人[98]提出了 CryptoDL，采用低阶多项式近似逼近的方法设计改造适合同态加密计算的 CNN 中常见的激活函数（ReLU、Sigmoid、Tanh 等），并进一步用来训练深度卷积神经网络。Chou 等人[99]提出了 Faster CryptoNets，推导出常用的激活函数的最佳近似，实现了最大稀疏编码，最小化逼近误差，并通过整个神经网络的稀疏表示，加速了对加密数据的深度学习模型的同态计算。2018 年，Bourse 等人[100]提出 FHE-DiNN 模型，利用神经网络的带符号整数权值和二进制激活函数执行加密预测，但目前预测的准确性一般。由于加密方案参数依赖于模型的结构，因此服务提供者如果更新模型，那么用户将需要重新加密数据。为了解决这个问题并且提高预测的准确率，Sanyal 等人[101]提出了 TAPAS 系统，研究了对全同态加密数据的机器学习模型的预测，将权重二值化，采用稀疏化技术对加密数据进行加速和并行计算，获得较高的准确率。Jin 等人[102]应用 GPU 加速 FHE 技术来

实现高效的同态卷积神经网络。2019 年，Boemer 等人[103]提出了一种名为 nGraph-HE 的框架，可以有效地对图形同态加密进行优化。Chillotti 等人[104]提出使用可编程自举技术实现神经网络中的非线性激活函数，进一步提高了效率。Jiang 等人[105]则提出了一种基于矩阵同态加密的通用算术运算方法，提出在一些用户提供的密文数据上，云服务提供商进行模型训练，该方法可以将加密模型应用到更新后的加密数据上。2019 年，Zheng 等人[28]利用门限部分同态加密实现了 Helen 系统，能够允许利用多个用户的数据同时训练模型，但不泄露数据，能够抵御 m 方中 $m-1$ 方都为恶意的对手。但是该系统不能抵御投毒攻击、数据提取攻击和拒绝服务攻击。2020 年，Wood 等人[106]使用 FHE 实现机器学习在医学和生物信息学中的应用。

尽管从理论层面认为 FHE 技术上可以进行任意计算，但受当前相关实际方案约束，FHE 普遍仅能支持整数类型的数据；同时，乘法深度需要固定而不能进行无限次的加法和乘法运算。除此之外，FHE 技术不支持比较运算操作。虽然目前存在一些实数上计算并有优化的 FHE 方案，但数据规模大幅扩张、计算负载不断加剧、非线性激活函数的拟合计算误差等原因导致 FHE 方案的效率无法得到进一步提升。因此，FHE 技术与机器学习算法的简单结合，将无法保证相关机器学习算法的操作对密文数据进行操作，后续基于同态加密的深度学习方案多与安全多方计算相结合，来降低后者的通信复杂度。

4.2　安全多方计算

在基于安全多方计算技术的隐私保护深度学习方面，MPC 允许互不信任的各方能够在自身私有输入上共同计算一个函数，其过程中不会泄露除函数输出之外的任何信息。但是，传统的 MPC 协议往往需要较为庞大的计算量和通信复杂度，导致其难以在实际机器学习中大规模部署。目前，常见的基于 MPC 的神经网络训练解决方案有：①基于混淆电路、不经意传输（Oblivious Transfer，OT）等技术的隐私保护神经网络方案，并执行两方 MPC 协议来完成激活函数等非线性操作计算；②基于秘密共享等技术，允许多个参与方参与的神经网络模型方案，且该过程不会透露数据或模型隐私信息。MPC 主要方案的发展历程如图 3 所示。科研人员将最初的两方隐私保护神经网络方案渐渐扩展到了三方甚至多方的方案。

1．两方隐私保护神经网络方案

典型的两方隐私保护神经网络方案包括 TASTY[107]、ABY[108]、SecureML[109]、Chameleon[110]、MiniONN[111]、DeepSecure[112]、GAZELLE[113]、XONN[114]、DELPHI[115]、CrypTFlow2[116]、ABY2.0[117]、SIRNN[118]、Cheetah[119]、MUSE[120]、SIMC[121]、QUOTIENT[122]、SECFLOAT[123]等。

Henecka 等人[107]认为在安全两方计算协议中使用单一技术（如同态加密或 Yao 电路）的效率较低，因此提出了一个能在明文、同态加密、Yao 电路之间相互转换的混合协议 TASTY，较大地提升了安全两方计算的效率。但是，因为同态加密和混淆电路的转换代价相对较高，

且同态加密随着安全参数的增长效率表现不佳，所以 TASTY 的效率相对于单个协议的方案提升有限。Demmler 等人[108]提出了 ABY 框架，利用高效乘法协议、快速转换技术以及预处理的密码操作，构造用于两方计算环境的基于算术共享、布尔共享和 Yao 共享的安全计算方案。SecureML[109]在两方计算环境中训练神经网络的隐私保护方法，SecureML 基于线性回归、逻辑回归和神经网络等模型提出一种高效的截断协议，相较于 ABY 提出的算法更加高效。ABY 和 SecureML 均利用 Beaver 所提出的三元组技术来实现部分乘法操作，方案通信成本仍不理想。由于三元组会降低协议的效率，Chameleon[110]等人利用一个半诚实的第三方替代三元组，从而减少了三元组的使用，基于 ABY 实现了加法秘密共享、GMW 和 GC 的混合，大幅改进了 ABY 的实用性和可扩展性。这些两方的协议主要考虑在非共谋服务器情况下的安全保障需求，假设云服务器是诚实且好奇的被动攻击模型，考虑在云服务器互不合谋情况下数据的安全性与模型的可用性。

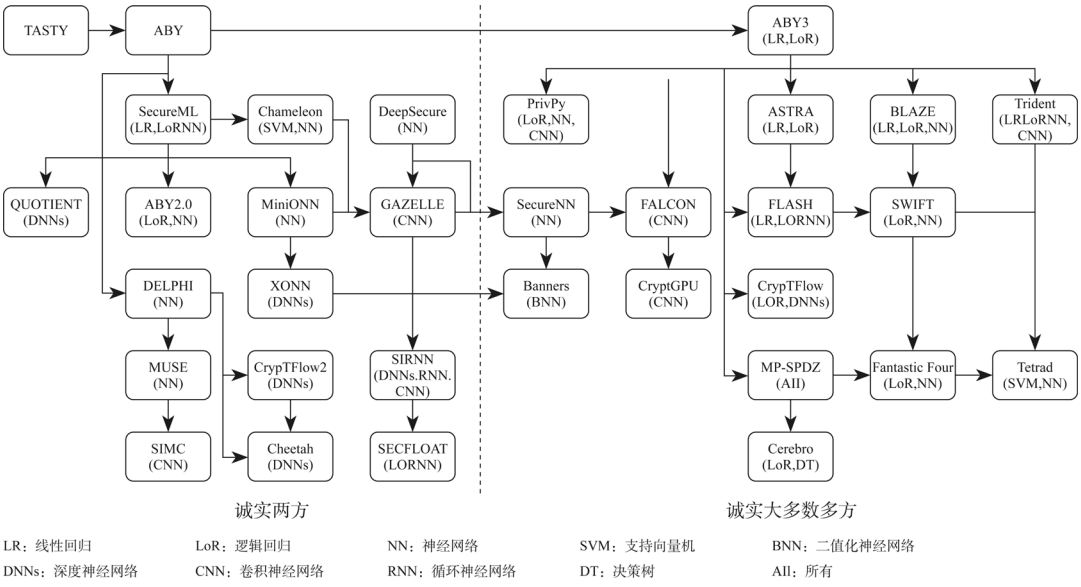

图 3　基于安全两方/多方计算的隐私保护神经网络方案的发展历程

为免受半可信参数服务器的攻击，MiniONN[111]基于简化的同态加密技术并将其应用于交换权重和梯度问题，把原始神经网络转换为遗忘神经网络再进行网络模型训练，并使用混淆电路计算近似非线性激活函数。DeepSecure[110]基于 Yao 混淆电路[81]对深度学习模型上的加密数据进行计算和推理，并完成相应的安全证明。Juvekar 等人[113]指出，上述工作表明同态加密在矩阵向量乘法上具有明显的优势，但在线性运算方面并不明显。需要注意的是，虽然这些混合协议可以提高识别率，但带宽和网络延迟方面的效率并不理想。从技术上说，这些协议均采用同态加密方法对标量乘法进行计算，采用安全多方计算对激活函数进行计算。

由于以往方案只关注密码学协议的优化，对于安全协议的效率提升有限。XONN[114]利用在混淆电路中免费的同或门代替深度学习模型中的花销较大的矩阵乘法运算，而且利用二值化神经网络的概念优化协议，为高效的安全协议设计提供了新思路。DELPHI[115]与 XONN 的思想相似，借助神经网络方法提升 DNNs 推理的效率。而 CrypTFlow2[116]和 ABY2.0[117]在半诚实模型下的总体性能较其他方案取得较大的进展，它们的性能接近，但是 ABY2.0 将更多的计算移到了 Offline 阶段。SIRNN[118]则基于 CrypTFlow2 进一步优化了协议，提升了通信和计算效率。Cheetah[119]通过多项式编码巧妙地消除了同态加密密文旋转带来的开销，且基于 CrypTFlow2 使用 Slient OT 极大地降低了安全推理的计算和通信开销。以往大多数方案都是关注 DNNs 推理方案，对 DNNs 的训练方案关注较少，QUOTIENT[122]针对 DNNs 模型设计了定制化的两方安全协议用于 DNNs 的训练，但是对于卷积神经网络较慢，存在较大的通信开销。MUSE[120]和 SIMC[121]则考虑了抗恶意客户端的两方安全推理方案。

目前的大多数方案处理浮点数运算要么会造成精度损失，要么效率低下。SECFLOAT[123]基于 Intel 计算库 MKL 和 SIRNN 的设计思想提供了一个安全两方计算库，与以往方案相比，在精度和开销上都得到了大幅度的提升。该方案只支持单精度浮点数，且只考虑了半诚实模型。经过研究人员的不断努力，两方隐私保护神经网络方案效率不断提升，整体效率相对于明文的差距正在不断减小，但距离实际应用场景还存在一定差距，尤其是针对复杂神经网络的开销还较大。除此之外，大多数安全两方推理方案未考虑恶意客户端的情形。

2．多方隐私保护神经网络方案

典型的多方参与的隐私保护神经网络方案有 ABY3[125]、SecureNN[126]、ASTRA[127]、FLASH[128]、Trident[129]、BLAZE[130]、SWIFT[131]、Fantastic Four[132]、Tetrad[133]、FALCON[134]、CrypTFlow[135]、Banners[136]和 CryptGPU[137]等。

ABY3[125]提出一种在半诚实环境和恶意环境下的新方案，可以完成三方之间的算术共享、布尔共享和 Yao 共享的转换，它对 SecureML 中的近似定点数乘法进行改进，使其得以在三方环境下使用，并相应地设计了一种计算分段多项式函数的协议。SecureNN[126]基于 SecureML 做出了改进，最终在两方和三方的情况下分别将效率提升 93 倍和 8 倍。ASTRA[125]基于 ABY3 考虑了半诚实环境和恶意敌手环境，提出一种新的更加安全的三方隐私保护神经网络协议，构建了一个公平的重构协议，保证了当且仅当诚实方接收到输出时，恶意攻击方才会接收到输出。对于分类任务，ASTRA 充分运用秘密共享方案中的不对称性性质，放弃 SecureML 和 ABY3 中一些消耗较高的混淆电路或并行前缀加法器原语，最终提出一种新的安全比较协议。在半诚实的环境中，ASTRA 则是将 33％的整体通信开销转移到离线阶段，而在恶意敌手环境基于高效的点积协议进一步提升在线通信效率。然而，ABY3、SecureNN、ASTRA 方案均是专注于半诚实环境下的隐私保护神经网络框架，进行相关点积计算是与向量大小呈线性关系的，在恶意环境下，向量点积、最高有效位的提取和截断的效率都会不同程度地降低。

ABY3 和 SecureNN 提出使用 Abort 的安全构建组件，且 ASTRA 将恶意终止的安全性提升到公平性，FLASH[128]则进一步实现了确保输出交付（Guaranteed Output Delivery，GOD）安全概念（无论敌手的行为如何，各方都获得输出）的四方隐私保护的神经网络框架，且健壮性仅需要对称密码原语来实现。FLASH 不需要使用数字签名、广播等密码原语，在协议中引入一个新的诚实参与方大大提升协议效率。Trident[129]提出了最多可以容忍一方腐败的四方隐私保护的神经网络协议，其中第四方在除输入共享和输出重构的阶段之外，所有的在线阶段均为非活动状态。在使用新的秘密共享方案后，Trident 方案将在线通信阶段部分转移到离线阶段部分，使得在线效率得以充分提升。BLAZE[130]则是效率较高的隐私保护神经网络协议，在 3 个参与方的情形下可以容忍一个恶意腐败方的存在，并得到了更强的公平性保证（所有诚实方和恶意方都获得相同的输出），基于所提出的点积协议、截断和位操作方法，方案效率比 ABY3 和 ASTRA 等方案更高。SWIFT[131]是基于秘密共享的一个强大且高效的 PPML 框架，达到了 GOD 安全性，它比 BLAZE 效率更高，在扩展到 4 个参与方的情况下，比 FLASH 快两个数量级，与 Trident 的效率一致。但是，SWIFT 中实现的 GOD 不考虑诚实方的隐私性。Fantastic Four[132]与 Tetrad[133]实现了具有主动安全性的诚实大多数四方协议，不需要依赖于函数的预处理，达到了保护所有参与方隐私性的 GOD 安全性且效率更高。除此之外，Tetrad 还实现了四方环境下支持的算术电路、布尔电路和 Yao 电路互相转换的混合协议。目前，部分方案考虑到隐私保护的神经网络方案恶意攻击环境，当恶意攻击者存在的情况下，还需考虑协议的公平性与一致性。

Wagh 等人基于 SecureNN 和 ABY3 推出了一个名为 FALCON[134]的框架，支持归一化且保护数据和模型参数隐私，可以支持训练像 AlexNet 和 VGG16 这样的大容量网络。FALCON 只使用算术秘密共享而避免使用转换协议（在算术、布尔和乱码电路之间）来实现针对非线性运算的恶意安全协议。微软研究院提出 CrypTFlow[135]平台，该平台将 TensorFlow[124]推理代码自动编译为高效安全计算协议，并可以将 MPC 协议从半诚实安全性模型转换为恶意安全性模型。为了加速 PPML 的训练和推理效率，CryptGPU 框架[137]提供了一种称为 CUDALongTensor 的模块将定点数转换为浮点数，使 MPC 协议能够在 GPU 中运算，CryptGPU 框架设计了 GPU 友好的密码学协议使 MPC 协议运算更加高效。但是，CryptGPU 目前仅考虑了半诚实模型，且实验仅是在局域网下进行的，在广域网下的效率还需要进一步探索。受 XONN 影响，Banners[136]基于复制的秘密共享技术设计了在恶意敌手模型下保护二值化神经网络安全推理的新方法，提高了推理的效率。

5 总结

综上所述，随着机器学习算法和人工智能应用领域的研究逐步深入，机器学习算法的特

殊性给用户数据和网络模型的隐私保护带来巨大挑战，迫切需要进一步考虑更高的安全及隐私威胁。针对安全威胁，还需要进一步探索针对投毒攻击、对抗攻击等攻击手段的防御技术，提高模型的鲁棒性、研究更强攻击性的防御手段是目前有待解决的问题。针对隐私威胁，全同态加密一直被认为是隐私保护机器学习的首选技术，但其数据扩张、计算负载、激活函数拟合误差等不利因素，使得基于安全多方计算的隐私保护机器学习得到了迅速发展。然而，现有隐私保护机器学习方案往往假设云服务器是诚实且好奇的被动攻击模型，考虑在云服务器互不合谋情况下数据的安全性与机器学习的可用性；在更高的安全等级并推广到多方场景下，还需考虑存在恶意攻击者情形下的公平性与一致性。因此，未来可能的研究方向如下。

（1）建立完善统一的安全评估标准。对于大部分隐私数据，可在源头上控制好这些数据的使用范围和收集过程。但是由于目前缺乏合理完善的安全评估标准，各类机构对于隐私数据的使用和收集都没有统一的标准，因此不可避免地会造成隐私的泄露。建立完善统一的安全评估标准势在必行，刻不容缓。

（2）研究具有更强鲁棒性的隐私保护的机器学习模型。随着如对抗攻击、投毒攻击等攻击手段的发展，普通模型已经不能满足隐私需求，模型的泄露给组织、机构带来的损失不可估量。研究能抵抗更强攻击手段的高鲁棒性机器学习模型是未来的工作。

（3）考虑更强的威胁场景的机器学习方案。分布式学习、联邦学习的场景是目前的趋势，但是对于大多数分布式的方案来说，它们都是考虑半诚实的威胁模型。因此，需要将目前的半诚实模型推广到恶意模型，使机器学习方案在恶意攻击者存在的情况下依然能保持公平性和一致性。

（4）提高现有方案的精度、效率。目前的大部分隐私保护方案都是基于同态加密、安全多方计算和差分隐私的，目前这几类技术的通信、计算等开销比较大，这大大降低了算法的效率，带来了不必要的资源浪费；并且这些方案存在精度丢失的问题。因此，研究效率、精度更高的隐私保护方案是一个重要的研究方向。

参考文献

[1] MITCHELL T M. Machine learning[M]. New York: McGraw-Hill, 2003.

[2] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

[3] STOICA I, SONG D, POPA R A, et al. A Berkeley view of systems challenges for AI[OL]. arXiv preprint arXiv:1712.05855, 2017. [2020-07-20]. https://arxiv.org/abs/1712.05855.

[4] 何英哲,胡兴波,何锦雯,等.机器学习系统的隐私和安全问题综述[J].计算机研究与发展,2019,56(10):2049-2070.

[5] 刘俊旭,孟小峰.机器学习的隐私保护研究综述[J].计算机研究与发展,2020,57(2):346-

362.

[6] 刘睿瑄,陈红,郭若杨,等.机器学习中的隐私攻击与防御[J].软件学报,2020,31(3):866-892.

[7] 魏立斐,陈聪聪,张蕾,等. 机器学习的安全问题及隐私保护[J]. 计算机研究与发展, 2020, 57(10): 2066-2085.

[8] 蒋瀚,刘怡然,宋祥福,等.隐私保护机器学习的密码学方法[J].电子与信息学报, 2020,42:1068-1078.

[9] 纪守领,杜天宇,李进锋,等.机器学习模型安全与隐私研究综述[J].软件学报, 2021,32(1):41-67.

[10] 郭娟娟,王琼霄,许新,等.安全多方计算及其在机器学习中的应用[J].计算机研究与发展, 2021,58(10):2163-2186.

[11] SU J, VARGAS D V, SAKURAI K. One pixel attack for fooling deep neural networks[J]. IEEE Trans on Evolutionary Computation, 2019, 23(5): 828-841.

[12] LIM H S M, TAEIHAGH A. Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities[J]. Sustainability, 2019, 11(20): 5791.

[13] HEAVEN D. Why deep-learning AIs are so easy to fool[J]. Nature, 2019, 574(7777): 163.

[14] BARRENO M, NELSON B, JOSEPH A D, et al. The security of machine learning[J]. Machine Learning, 2010, 81(2): 121-148.

[15] BIGGIO B, FUMERA G, ROLI F. Security evaluation of pattern classifiers under attack[J]. IEEE Trans on Knowledge and Data Engineering, 2013, 26(4): 984-996.

[16] MOISEJEVS I. Will my machine learning system be attacked? [EB/OL].Towards Data Science. (2019-07-15)[2020-06-01]. https://towardsdatascience.com/will-my-machine-learning-be-attacked-6295707625d8.

[17] LAUNCHBURY J, ARCHER D, DUBUISSON T, et al. Application-scale secure multiparty computation[C]. European Symp on Programming Languages and Systems. Berlin: Springer, 2014: 8-26.

[18] PAPERNOT N, MCDANIEL P, GOODFELLOW I, et al. Practical black-box attacks against machine learning[C]. Proc of the 2017 ACM on Asia Conf on computer and communications security. New York: ACM, 2017: 506-519.

[19] LASKOV P. Practical evasion of a learning-based classifier: A case study[C]. 2014 IEEE Symp on security and privacy. Piscataway, NJ: IEEE, 2014: 197-211.

[20] BARRENO M, NELSON B, SEARS R, et al. Can machine learning be secure?[C]. Proc of the 2006 ACM Symp on Information, computer and communications security. New York: ACM,

2006: 16-25.

[21] LOW Y, GONZALEZ J, KYROLA A, et al. Distributed graphlab: A framework for machine learning in the cloud[J]//Proc of the Vldb Endowment, 2012, 5(8):716-727.

[22] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: Strategies for improving communication efficiency [OL]. arXiv preprint arXiv:1610.05492, 2016. [2020-07-20]. https://arxiv.org/abs/1610.05492.

[23] DALVI N, DOMINGOS P, SANGHAI S, et al. Adversarial classification[C]. Proc of the tenth ACM SIGKDD Int Conf on Knowledge discovery and data mining. New York: ACM, 2004: 99-108.

[24] LOWD D, MEEK C. Good word attacks on statistical spam filters[C]. Conf on Email & Anti-spam. Mountain View: CEAS, 2005.

[25] MILLER D J, XIANG Z, KESIDIS G. Adversarial learning in statistical classification: A comprehensive review of defenses against attacks[OL]. arXiv preprint arXiv:1904.06292, 2019. [2020-07-20]. https://arxiv.org/abs/1904.06292.

[26] GOLDWASSER S. From idea to impact, the crypto story:What's next? Cryptography & machine learning:What else? [OL].In IACR Distinguished Lecture of CRYPTO 2018. Santa Barbara,CA,USA.August 17-19,2018.[2020-06-01].https://www.iacr.org/cryptodb/data/paper.php?pubkey=29941.

[27] SHOKRI R, SHMATIKOV V. Privacy-preserving deep learning[C]. Proc of the 22nd ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2015: 1310-1321.

[28] ZHENG W, POPA R A, GONZALEZ J E, et al. Helen: Maliciously secure coopetitive learning for linear models[C]. 2019 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2019: 724-738.

[29] MOISEJEVS I. Poisoning attacks on machine learning[EB/OL]//Towards Data Science. (2019-07-15)[2020-06-01].https://towardsdatascience.com/poisoning-attacks-on-machine-learning-1ff247c254db.

[30] STEINHARDT J, KOH P W W, LIANG P S. Certified defenses for data poisoning attacks[C]. Advances in neural information processing systems. Cambridge, MA: MIT press, 2017: 3517-3529.

[31] FANG M, CAO X, JIA J, et al. Local model poisoning attacks to byzantine-robust federated learning. Proceedings of the 29th USENIX Conference on Security Symposium. Berkeley, CA, USA. August 12-14, 2020 : USENIX, 1623-1640.

[32] BLANCHARD P, GUERRAOUI R, STAINER J, et al. Machine learning with adversaries: Byzantine tolerant gradient descent. In Advances in Neural Information Processing Systems. Cambridge, Massachusetts:MIT, 2017: 119-129.

[33] EL-MHAMDI E M, GUERRAOUI R, ROUAULT S. The hidden vulnerability of distributed learning in byzantium. arXiv preprint arXiv:1802.07927, 2018.

[34] YIN D, CHEN Y, KANNAN R, et al. Byzantine-Robust distributed learning: Towards optimal statistical rates. In International Conference on Machine Learning. Stockholm, Sweden: July 15, 2018, JMLR, 2018:5636-5645.

[35] SHEJWALKAR V, HOUMANSADR A. Manipulating the byzantine: Optimizing model poisoning attacks and defenses for federated learning. In Proceedings of the 28nd Annual Network and Distributed System Security Symposium (NDSS). Virtually, February 21-25, 2021.

[36] KOH P W, STEINHARDT J, LIANG P. Stronger data poisoning attacks break d ata sanitization defenses[OL].arXiv preprint arXiv:1811.00741,2018.[2020-07-20].https://arxiv.o rg/abs/1811.00741.

[37] BARACALDO N, CHEN B, LUDWIG H, et al. Mitigating poisoning attacks on machine learning models: A data provenance based approach[C]. Proc of the 10th ACM Workshop on Artificial Intelligence and Security. NewYork: ACM, 2017: 103-110.

[38] SUCIU O, MARGINEAN R, KAYA Y, et al. When does machine learning {FAIL}? generalized transferability for evasion and poisoning attacks[C]. Proc of the 27th USENIX Security Symposium (USENIX Security 18). Berkeley, CA: USENIX, 2018: 1299-1316.

[39] GU T, LIU K, DOLAN-GAVITT B, et al. Badnets: Evaluating backdooring attacks on deep neural networks. IEEE Access, 2019, 7: 47230-47244.

[40] LIU Y, MA S, AAFER Y, et al. Trojaning attack on neural networks [C]. In Network and Distributed System Security Symposium (NDSS). San Diego, California, February, 2018: 18-21.

[41] LIU K, DOLAN-GAVITT B, GARG S. Fine-pruning: Defending against backdooring attacks on deep neural networks. In Research in Attacks, Intrusions, and Defenses: 21st International Symposium (RAID 2018). Heraklion, Crete, Greece, September 10-12, 2018. Berlin: Springer, 2018: 273-294.

[42] WANG B, YAO Y, SHAN S, et al. Neural cleanse: Identifying and mitigating backdoor attacks in neural networks. In 2019 IEEE Symposium on Security and Privacy (SP) . San Francisco, CA, USA, May 20-22, 2019. Piscataway, NJ: IEEE, 2019: 707-723.

[43] LIU Y, LEE W C, TAO G, et al. Abs: Scanning neural networks for backdoors by artificial brain stimulation. In Proceedings of the 2019 ACM SIGSAC Conference on Computer and

Communications Security. London, United Kingdom,November 11-15, 2019, New York: ACM, 2019: 1265-1282.

[44] GAO Y, XU C, WANG D, et al. Strip: A defence against trojan attacks on deep neural networks. In Proceedings of the 35th Annual Computer Security Applications Conference, December 5-9, 2022 . Austin, Texas, USA. Piscataway, NJ: IEEE, 2019: 113-125.

[45] CHEN H, FU C, ZHAO J, et al. DeepInspect: A Black-box Trojan Detection and Mitigation Framework for Deep Neural Networks. In the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019). Macao, China, August 10-16, 2019: 4658-4664.

[46] GUO W, WANG L, XING X, et al. Tabor: A highly accurate approach to inspecting and restoring trojan backdoors in ai systems. arXiv preprint arXiv:1908.01763. 2019.

[47] BIGGIO B, CORONA I, NELSON B, et al. Security evaluation of support vector machines in adversarial environments[M]. Switzerland: Springer, Cham, 2014: 105-153.

[48] SZEGEDY C, ZAREMBA W, SUTSKEVER I, et al. Intriguing properties of neu ral networks[OL]. arXiv preprint arXiv: 1312.6199, 2013. [2020-07-20]. https://arxiv.org/abs/1312.6199.

[49] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples[OL].arXiv preprint arXiv:1412.6572, 2014. [2020-07-20]. https://arxiv.org/abs/1412.6572.

[50] MOOSAVI-DEZFOOLI S M, FAWZI A, FROSSARD P. Deepfool: a simple and accurate method to fool deep neural networks[C]. Proc of the IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 2574-2582.

[51] RU B, COBB A, BLAAS A, et al. Bayesopt adversarial attack[C]. Int Conf on Learning Representations. Addis Ababa (Virtual Conference): ICLR, 2020.

[52] ZHOU M, WU J, LIU Y, et al. DaST: Data-free substitute training for adversarial attacks[OL].arXiv preprint arXiv:2003.12703, 2020. [2020-07-20]. https://arxiv.org/abs/2003.12703.

[53] MU J, WANG B, LI Q, et al. A hard label black-box adversarial attack against graph neural networks. Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. 2021.

[54] JIA W, LU Z, ZHANG H, et al. Fooling the Eyes of Autonomous Vehicles: Robust Physical Adversarial Examples Against Traffic Sign Recognition Systems. arXiv preprint arXiv: 2201. 06192, 2022.

[55] TRAMÈR F, KURAKIN A, PAPERNOT N, et al. Ensemble adversarial training: Attacks and defenses[OL].arXiv preprint arXiv:1705.07204,2017.[2020-07-20].https://arxiv.org/abs/1705.07204.

[56] ROSS A S, DOSHI-VELEZ F. Improving the adversarial robustness and interpretability of deep neural networks by regularizing their input gradients[C]. Proc of the Thirty-second AAAI Conf on Artificial Intelligence. Menlo Park,CA: AAAI, 2018.

[57] PAPERNOT N, MCDANIEL P, WU X, et al. Distillation as a defense to adversarial perturbations against deep neural networks[C]. Proc of the 2016 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2016: 582-597.

[58] ZANTEDESCHI V, NICOLAE M I, RAWAT A. Efficient defenses against adversarial attacks[C]. Proc of the 10th ACM Workshop on Artificial Intelligence and Security. NewYork: ACM, 2017: 39-49.

[59] TRAMÈR F, ZHANG F, JUELS A, et al. Stealing machine learning models via prediction apis[C]. Proc Of The 25Th Usenix Security Symp. Berkeley, CA: USENIX, 2016: 601-618.

[60] SHI Y, SAGDUYU Y, GRUSHIN A. How to steal a machine learning classifier with deep learning[C]. 2017 IEEE Int Symp on Technologies for Homeland Security (HST). Piscataway, NJ: IEEE, 2017: 1-5.

[61] WANG B, GONG N Z. Stealing hyperparameters in machine learning[C]. 2018 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2018: 36-52.

[62] SALEM A, BACKES M, ZHANG Y. (2022). Get a Model! Model Hijacking Attack Against Machine Learning Models. ArXiv, abs/2111.04394.

[63] FREDRIKSON M, JHA S, RISTENPART T. Model inversion attacks that exploit confidence information and basic countermeasures[C]. Proc of the 22nd ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2015: 1322-1333.

[64] FREDRIKSON M, LANTZ E, JHA S, et al. Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing[C]. Proc of the 23rd USENIX Conf on Security Symp. Berkeley, CA: USENIX, 2014: 17-32.

[65] ATENIESE G, MANCINI L V, SPOGNARDI A, et al. Hacking smart machines with smarter ones: How to extract meaningful data from machine learning classifiers[J]. Int Journal of Security and Networks, 2015, 10(3): 137-150.

[66] SONG C, RISTENPART T, SHMATIKOV V. Machine learning models that remember too much[C]. Proc of the 2017 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2017: 587-601.

[67] CARLINI N, LIU C, ERLINGSSON Ú, et al. The secret sharer: Evaluating and testing unintended memorization in neural networks[C]. Proc of the 28th USENIX Conf on Security Symp. Berkeley, CA: USENIX, 2019: 267-284.

[68] SHOKRI R, STRONATI M, SONG C, et al. Membership inference attacks against machine learning models[C]. 2017 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2017: 3-18.

[69] NASR M, SHOKRI R, HOUMANSADR A. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. IEEE symposium on security and privacy (SP). Piscataway, NJ: IEEE, 2019: 739-753.

[70] YEOM S, GIACOMELLI I, FREDRIKSON M, et al. Privacy risk in machine learning: Analyzing the connection to overfitting[C]. 2018 IEEE 31st Computer Security Foundations Symp (CSF). Piscataway, NJ: IEEE, 2018: 268-282.

[71] SONG C, SHMATIKOV V. Auditing data provenance in text-generation models [C]. Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2019: 196-206.

[72] TRUEX S, LIU L, GURSOY M E, et al. Demystifying membership inference attacks in machine learning as a service[J]. IEEE Trans on Services Computing, 2019:1-1.

[73] HAYES J, MELIS L, DANEZIS G, et al. LOGAN: Membership inference attacks against generative models[J]//Proc on Privacy Enhancing Technologies, 2019, 2019(1): 133-152.

[74] LINDELL Y, PINKAS B. Privacy preserving data mining[C]. Annual Int Cryptology Conf. Berlin: Springer, 2000: 36-54.

[75] AGRAWAL R, SRIKANT R. Privacy-preserving data mining[C]. Proc of the 2000 ACM SIGMOD Int Conf on Management of data. New York: ACM, 2000: 439-450.

[76] BUNN P, OSTROVSKY R. Secure two-party k-means clustering[C]. Proc of the 14th ACM Conf on Computer and communications security. New York: ACM, 2007: 486-497.

[77] YU H, VAIDYA J, JIANG X. Privacy-preserving SVM classification on vertically partitioned data[C]. Pacific-Asia Conf on Knowledge Discovery and Data Mining. Berlin: Springer, 2006: 647-656.

[78] SANIL A P, KARR A F, LIN X, et al. Privacy preserving regression modelling via distributed computation[C]. Proc of the tenth ACM SIGKDD Int Conf on Knowledge discovery and data mining. New York: ACM, 2004: 677-682.

[79] AONO Y, HAYASHI T, TRIEU P L, et al. Scalable and secure logistic regression via homomorphic encryption[C]. Proc of the Sixth ACM Conf on Data and Application Security and Privacy. New York: ACM, 2016: 142-144.

[80] NIKOLAENKO V, WEINSBERG U, IOANNIDIS S, et al. Privacy-preserving ridge regression on hundreds of millions of records[C]. 2013 IEEE Symp on Security and Privacy.

Piscataway, NJ: IEEE, 2013: 334-348.

[81] YAO A. How to generate and exchange secrets[C]. Proc of the 27th Annual Symp on Foundations of Computer Science. Piscataway, NJ: IEEE, 1986: 162-167.

[82] DWORK C, MCSHERRY F, NISSIM K, et al. Calibrating noise to sensitivity in private data analysis[C]. Theory of cryptography Conf. Berlin: Springer, 2006: 265-284.

[83] SU S, TANG P, CHENG X, et al. Differentially private multi-party high-dimensional data publishing[C]. 2016 IEEE 32nd Int Conf on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2016: 205-216.

[84] ZHANG T, ZHU Q. Dynamic differential privacy for ADMM-based distributed classification learning[J]. IEEE Trans on Information Forensics and Security, 2016, 12(1): 172-187.

[85] YUAN G, YANG Y, ZHANG Z, et al. Convex optimization for linear query processing under approximate differential privacy[C]. Proc of the 22nd ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2016: 2005-2014.

[86] MCSHERRY F, TALWAR K. Mechanism design via differential privacy[C]. 48th Annual IEEE Symp on Foundations of Computer Science (FOCS'07). Piscataway, NJ: IEEE, 2007: 94-103.

[87] GENTRY C. Fully homomorphic encryption using ideal lattices[C]. Proc of the Forty-first Annual ACM Symp on Theory of Computing. New York: ACM, 2009: 169-178.

[88] RIVEST R L, SHAMIR A, ADLEMAN L. A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of the ACM, 1978, 21(2): 120-126.

[89] ELGAMAL T. A public key cryptosystem and a signature scheme based on discrete logarithms[J]. IEEE Trans on information theory, 1985, 31(4): 469-472.

[90] PAILLIER P. Public-key cryptosystems based on composite degree residuosity classes[C]. International conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 1999: 223-238.

[91] FAN J, VERCAUTEREN F. Somewhat practical fully homomorphic encryption[J]. IACR Cryptology ePrint Archive, 2012: 144.

[92] BRAKERSKI Z, VAIKUNTANATHAN V. Efficient fully homomorphic encryption from (standard) LWE[J]. SIAM Journal on Computing, 2014, 43(2): 831-871.

[93] CHABANNE H, DE WARGNY A, MILGRAM J, et al. Privacy-preserving classification on deep neural network[J]. IACR Cryptology ePrint Archive, 2017: 35.

[94] GRAEPEL T, LAUTER K, NAEHRIG M. ML confidential: Machine learning on encrypted data. International Conference on Information Security and Cryptology. In Information Security and Cryptology(ICISC 2012), Berlin: Springer, 2012: 1-21. ICISC'12: Proceedings of the

15th international conference on Information Security and CryptologyNovember 2012 Pages 1-21.

[95] ZHANG Q, YANG L T, CHEN Z. Privacy preserving deep computation model on cloud for big data feature learning. IEEE Transactions on Computers 65.5 (2015): 1351-1362.

[96] GILAD-BACHRACH R, DOWLIN N, LAINE K, et al. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy[C]. Int Conf on Machine Learning. New York: ACM, 2016: 201-210.

[97] CHABANNE H, WARGNY A D, MILGRAM J, et al. Privacy-preserving classification on deep neural network. Cryptology ePrint Archive (2017).

[98] HESAMIFARD E, TAKABI H, GHASEMI M. CryptoDL: Deep neural networks over encrypted data[J]. arXiv preprint arXiv:1711.05189, 2017. [2020-07-20]. https://arxiv.org/abs/1711.05189.

[99] EDWARD C, THAO N, JASH B, et al. Faster cryptonets: Leveraging sparsity for real-world encrypted inference. arXiv preprint arXiv:1811.09953 (2018).

[100] BOURSE F, MINELLI M, MINIHOLD M, et al. Fast homomorphic evaluation of deep discretized neural networks[C]. Annual Int Cryptology Conf. Switzerland: Springer, Cham, 2018: 483-512.

[101] SANYAL A, KUSNER M, GASCON A, et al. TAPAS: Tricks to accelerate (encrypted) prediction as a service[C]. Int Conf on Machine Learning. New York: ACM, 2018: 4490-4499.

[102] JIN C, RAGAB M, AUNG K M M. Secure transfer learning for machine fault diagnosis under different operating conditions[C]. International Conference on Provable Security. Springer, Cham, 2020: 278-297.

[103] BOEMER F, LAO Y, CAMMAROTA R, et al. nGraph-HE: A graph compiler for deep learning on homomor-phically encrypted data. Proceedings of the 16th ACM International Conference on Computing Frontiers. 2019.

[104] CHILLOTTI I, GAMA N, GEORGIEVA M, et al. Faster fully homomorphic encryption: Bootstrapping in less than 0.1 seconds[C]. Int Conf on the Theory and Application of Cryptology and Information Security. Berlin: Springer, 2016: 3-33.

[105] JIANG X, KIM M, LAUTER K, et al. Secure outsourced matrix computation and application to neural networks[C]. Proc of the 2018 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2018: 1209-1222.

[106] WOOD A, NAJARIAN K, KAHROBAEI D. Homomorphic encryption for machine learning in medicine and bioinformatics[J]. ACM Computing Surveys, 2020, 53(4): 1-35.

[107] HENECKA W, KÖGL S, SADEGHI A R, et al. TASTY: tool for automating secure two-party computations[C]. Proc of the 17th ACM Conf on Computer and Communications Security. New York: ACM, 2010: 451-462.

[108] DEMMLER D, SCHNEIDER T, ZOHNER M. ABY-A framework for efficient mixed-protocol secure two-party computation[C]. Network and Distributed System Security Symp. San Diego, CA: ISOC, 2015.

[109] MOHASSEL P, ZHANG Y. SecureML: A system for scalable privacy-preserving machine learning[C]. 2017 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2017: 19-38.

[110] RIAZI M S, WEINERT C, TKACHENKO O, et al. Chameleon: A hybrid secure computation framework for machine learning applications[C]. Proc of the 2018 on Asia Conf on Computer and Communications Security. New York: ACM, 2018: 707-721.

[111] LIU J, JUUTI M, LU Y, et al. Oblivious neural network predictions via minion transformations[C]. Proc of the 2017 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2017: 619-631.

[112] ROUHANI B D, RIAZI M S, KOUSHANFAR F. DeepSecure: Scalable provably-secure deep learning[C]. Proc of the 55th Annual Design Automation Conf. New York: ACM, 2018: 1-6.

[113] JUVEKAR C, VAIKUNTANATHAN V, CHANDRAKASAN A. GAZELLE: a low latency framework for secure neural network inference[C]. Proc of the 27th USENIX Conf on Security Symp. Berkeley, CA: USENIX, 2018: 1651-1668.

[114] RIAZI M S, SAMRAGH M, CHEN H, et al. XONN: XNOR-based Oblivious Deep Neural Network Inference[C]. 28th USENIX Security Symposium (USENIX Security 19). USENIX Association, 2019: 1501-1518.

[115] MISHRA P, LEHMKUHL R, SRINIVASAN A, et al. DELPHI: A cyptographic inference service for neural networks[C]. Proc of the 29th USENIX Conf on Security Symp. Berkeley, CA: USENIX, 2020.

[116] RATHEE D, RATHEE M, KUMAR N, et al. CrypTFlow2: Practical 2-Party Sec ure Inference[C]. Proceedings of the 2020 ACM SIGSAC Conference on Computer and Co mmunications Security. Association for Computing Machinery, 2020: 325-342.

[117] PATRA A, SCHNEIDER T, DARMSTADT T U, et al. ABY2.0: Improved Mixed-Protocol Secure Two-Party Computation[C/OL]//30th USENIX Security Symposium (USENIX Security 21). USENIX Association, 2021: 2165-2182.

[118] RATHEE D, RATHEE M, KIRAN GOLI R K, et al. SIRNN: A Math Library for Secure

RNN Inference[C/OL]//2021 IEEE Symposium on Security and Privacy (SP). 2021: 1003-1020. DOI:10.1109/SP40001.2021.00086.

[119] HUANG Z, LU W, HONG C, et al. Cheetah: Lean and Fast Secure Two-Party Deep Neural Network Inference[J/OL]. Cryptology ePrint Archive, 2022. https://eprint.iacr.org/2022/207.pdf.

[120] LEHMKUHL R, MISHRA P, SRINIVASAN A. MUSE: Secure Inference Resilient to Malicious Clients[C]. 30th USENIX Security Symposium (USENIX Security 21). 2021: 2201-2218.

[121] CHANDRAN N, GUPTA D, LAKSHMI S, et al. SIMC: ML Inference Secure Against Malicious Clients at Semi-Honest Cost[C]. 31st USENIX Security Symposium (USENIX Security 22). 2022.

[122] AGRAWAL N, SHAHIN S A, KUSNER M J, et al. QUOTIENT: Two-party secure neural network training and prediction[C]. Proc of the 2019 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2019: 1231-1247.

[123] RATHEE D, BHATTAC A, SHARMA R, et al. SECFLOAT: Accurate Floating-Point meets Secure 2-Party Computation[C]. 43rd IEEE Symposium on Security and Privacy (IEEE S&P 2022). 2022.

[124] ABADI M, BARHAM P, CHEN J, et al. TensorFlow: A system for large-scale machine learning[C]. Proc of the 12th USENIX Conf on Operating Systems Design and Implementation. Berkeley, CA: USENIX, 2016: 265-283.

[125] MOHASSEL P, RINDAL P. ABY3: A mixed protocol framework for machine learning[C]. Proc of the 2018 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2018: 35-52.

[126] WAGH S, GUPTA D, CHANDRAN N. SecureNN: Efficient and private neural network training[J]. IACR Cryptology ePrint Archive, 2018: 442.

[127] CHAUDHARI H, CHOUDHURY A, PATRA A, et al. ASTRA: High throughput 3pc over rings with application to secure prediction[C]. Proc of the 2019 ACM SIGSAC Conf on Cloud Computing Security Workshop. New York: ACM, 2019: 81-92.

[128] BYALI M, CHAUDHARI H, PATRA A, et al. FLASH: fast and robust framework for privacy-preserving machine learning[J]. Proc on Privacy Enhancing Technologies, 2020(2): 459-480.

[129] CHAUDHARI H, RACHURI R, SURESH A. Trident: Efficient 4pc framework for privacy preserving machine learning[C]. 27th Annual Network and Distributed System Security Symp(NDSS 2020). San Diego,CA: ISOC, 2020: 23-26.

[130] PATRA A, SURESH A. BLAZE: Blazing fast privacy-preserving machine learning[J].

arXiv preprint arXiv:2005.09042, 2020. [2020-07-20. https://arxiv.org/abs/2005.09042].

[131] KOTI N, PANCHOLI M, Patra A, et al. SWIFT: Super-fast and Robust Privacy-Preserving Machine Learning[C/OL]//30th USENIX Security Symposium (USENIX Security 21). USENIX Association,2021:2651-2668.https://www.usenix.org/conference/usenixsecurity21/presentation/koti.

[132] DALSKOV A, ESCUDERO D, KELLER M. Fantastic Four: Honest-Majority Four-Party Secure Computation With Malicious Security[C]. 30th USENIX Security Symposium (USENIX Security 21). USENIX Association, 2021: 2183-2200.

[133] KOTI N, PATRA A, RACHURI R, et al. Tetrad: Actively Secure 4PC for Secure Training and Inference[C/OL]//Network and Distributed Systems Security (NDSS) Symposium 2022. http://arxiv. org/abs/2106.02850. DOI:10.14722/ndss.2022.24058.

[134] WAGH S, TOPLE S, BENHAMOUDA F, et al. FALCON: Honest-majority maliciously secure framework for private deep learning[OL]. arXiv preprint arXiv:2004.02229, 2020. [2020-07-20]. https://arxiv.org/abs/2004.02229.

[135] KUMAR N, RATHEE M, CHANDRAN N, et al. CrypTFlow: Secure TensorFlow Inference[C/OL]//2020 IEEE Symposium on Security and Privacy (S&P). San Francisco, CA, USA: IEEE, 2020: 336-353. DOI:10.1109/SP40000.2020.00092.

[136] BARRONDO A, CHABANNE H, ÖNEN M. Banners: Binarized Neural Networks with Replicated Secret Sharing[C/OL]//IH and MMSec 2021 - Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security. Association for Computing Machinery, 2021: 63-74. DOI:10.1145/3437880.3460394.

[137] TAN S, KNOTT B, TIAN Y, et al. CryptGPU: Fast Privacy-Preserving Machine Learning on the GPU[C/OL]//2021 IEEE Symposium on Security and Privacy (S&P). 2021: 1021-1038. DOI:10.1109/SP40001.2021.00098.

基于外包计算模型的深度学习研究综述

马旭[1,2]，孙晓倩[1]，张肖瑜[2]，陈晓峰[2]

1.曲阜师范大学，网络空间安全学院，曲阜，273165
2.西安电子科技大学，综合业务网理论及关键技术国家重点实验室，西安，710071
通讯作者：马旭，E-mail：xma@qfnu.edu.cn

摘　要：依赖于第三方服务器的深度学习框架本质上可以归类为外包计算，本文主要关注外包计算模型下的深度学习方案。首先，揭示外包计算和深度学习之间的关系；其次，提炼基于外包计算模型的深度学习中常用的核心密码技术，并分类归纳外包深度学习的核心思路；再次，分析已有外包深度学习方案的安全性、隐私性及高效性；最后，展望外包深度学习的研究热点和发展方向。

关键词：外包计算；深度学习；安全性；隐私性

A Survey on Secure Outsourced Deep Learning

MA Xu[1,2], SUN Xiaoqian[1], ZHANG Xiaoyu[2], CHEN Xiaofeng[2]

1.School of Cyber Science and Engineering, Qufu Normal University, Qufu, 273165;
2.State Key Laboratory of Integrated Service Networks (ISN), Xidian University, Xi'an, 710071
Corresponding author: MA Xu, E-mail：xma@qfnu.edu.cn

Abstract: Deep learning frameworks based on a third-part can be essentially categorized as outsourced computation for machine learning. In this paper, we mainly focus on deep learning schemes under outsourced computation models. Firstly, we reveal the relationship between outsourced computing and deep learning. Secondly, we refine the core cryptographic techniques that are commonly used in outsourced deep learning, and summarize the core ideas of outsourced deep learning. Thirdly, we analyze the security, privacy and efficiency of existing outsourced deep learning solutions. Finally, we prospect the hotspots and development directions of outsourced deep learning.

Keywords: Outsourced Computation; Deep Learning; Security; Privacy

1 引言

随着深度学习[1]的快速发展和计算能力的增强，许多数据驱动型应用的技术水平得以提高。例如，基于深度学习的自动检测和监控系统可以从海量数据中提取过去无法提取的有用信息，进而使基于深度学习的数据分析广泛应用于医疗和金融等领域。但是，深度学习模型仍需要大量的存储和计算资源，且机器学习模型的性能受训练数据质量和数量的影响较大。因此，资源有限的客户端更倾向于将深度学习任务外包给具有无限存储和计算资源的云服务器。同时，外包深度学习会出现严重的隐私和安全问题。首先，训练数据可能会包含高度敏感的信息，如医疗数据和金融数据。如果直接以明文的形式将敏感数据上传到有可能恶意的服务器，将会带来严重的信息泄露风险。其次，当多个数据集上实施分布式学习或协作学习时，参与计算的恶意客户端也会带来隐私问题。最后，云服务器必须能够证明返回结果的正确性，并且检查证明所需的计算能力要远低于重新计算本身。综上所述，外包深度学习最重要的安全要求就是隐私性和安全性。

完整的深度学习算法包括特征提取、模型训练和模型评估。在外包计算的不同阶段，必须满足不同的隐私要求。外包深度学习中隐私要求可以分为数据隐私和模型隐私。一般来说，隐私保护可以通过加密、匿名数据集或模型来实现隐私保护，这使得它很难让云服务器在加密数据集上运行深度学习算法。因此，在外包深度学习中实现隐私与效用之间的权衡是研究的主要目标。当前，相关研究对外包计算和深度学习越来越感兴趣，大部分研究成果分散在不同的研究领域。本文侧重于两个研究领域的交叉，我们总结了为深度学习任务量身定制的外包计算技术，并利用该技术在深度学习的应用场景中达到最佳性能。然后，我们进一步回顾相关的文献，并讨论了各种外包深度学习架构的主要优缺点。最后，我们指出了未来的研究方向。

主要贡献：本文主要关注外包计算和深度学习之间的交叉，主要研究范围是外包深度学习领域。考虑到内容的完整性，我们也回顾和讨论了深度学习和外包计算。总的来说，本文的贡献有如下 3 点。

（1）我们回顾了深度学习的发展，包括多层感知器、循环神经网络、卷积神经网络和深度强化学习。

（2）我们总结了与外包计算相关的重要文献，并从深度学习的角度讨论了外包计算的前沿技术，指出了外包计算与从深度学习算法中提取的基础数学运算之间的关系。

（3）我们分析和比较了不同加密技术以及外包云计算架构的效率和安全性，为深度学习提供了外包计算技术的选择策略。

组织结构：深度学习和外包计算大多是独立研究的。近年来，这两个领域的交叉点越来

越多。我们将这些工作归类为：①深度学习概述以及应用；②外包计算技术概述；③深度学习和外包计算交叉领域的工作回顾。如图 1 所示，首先回顾与深度学习相关的重要文献，并在第 2 节介绍相关研究背景；第 3 节主要介绍外包计算；第 4 节讲述外包深度学习，并突出外包计算在解决深度学习问题方面的优势，我们介绍并比较了最前沿的外包深度学习模型和相关算法，旨在帮助从事深度学习和外包计算的研究人员设计更有效的协议；最后，在第 5 节展望未来的研究方向。

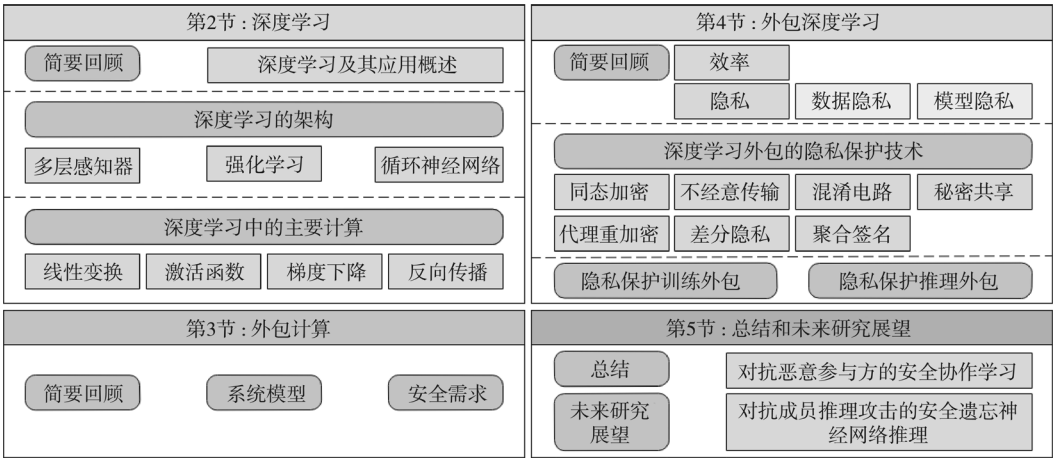

图 1　本文组织结构

2　深度学习

在本节中，首先对深度学习进行回顾；然后介绍深度学习中常见的主要体系结构；最后，简要介绍深度学习中的基础计算，以便在深度学习和外包计算之间建立明确的关系。

2.1　深度学习概述

深度学习[2]是机器学习的一个分支，起源于人工神经网络。机器学习的早期研究大多采用浅层结构的体系架构，如高斯混合模型、支持向量机和逻辑回归。通常，这些模型中只有一层，负责将输入特征转换为输出空间。虽然浅层体系结构模型能够有效地解决许多简单问题，但是它们有限的表达能力在处理复杂的应用程序时很困难，如与语音和视觉相关的应用程序场景。这些应用表明，需要使用深度体系架构来从输入中提取复杂的特征表示。深度神经网络具有足够的隐藏层。此外，其他多层模型也可视为深度模型，如深度高斯过程和深度随机森林。

表 1 罗列了与深度学习相关的已有综述文献。Goodfellow[2]提供了一本《关于深度学习的

学习指南》，其中包括基础知识、优化方法和流行应用程序。Schmidhuber 等人[3]总结了相关工作，并对深度学习进行了全面的调研，包括有监督和无监督学习、强化学习，并提供了对大型网络和深层网络的简短代码。Liu 等人[4]介绍了深度学习的历史演变及其应用，总结了几种深度学习模型的基本原理并回顾了深度学习的有关应用，如语音识别、计算机视觉和信号处理[5-8]。Lecun 等人[9]总结了多层神经网络训练的模型、方法和算法，并回顾了卷积神经网络和递归神经网络在图像理解和语言处理中的应用。Arulkumaran 等人[10]调查了强化学习，包括深度 Q 网络、异步优势 Actor-Critic 算法和信任区域策略优化算法等核心算法。Zhang 等人[11]对多任务学习进行了研究，分为表征学习和特征选择方法、低秩法、任务聚类方法和任务关系学习方法。Litjens 等人[12]调查了深度学习在医学图像分类、分割、目标检测和其他任务中的应用。

表 1　与深度学习相关的已有综述文献

参考文献	简介
[2]	关于深度学习的学习指南
[3]	关于深度学习的综述
[4]	深度学习的研究及应用
[5- 8]	深度学习的应用
[9]	多层神经网络及其应用
[10]	深度强化学习综述
[11]	多任务学习研究综述
[12]	深度学习在医学图像分析中的应用

2.2　深度学习的架构

在本节中，简要回顾了深度学习的架构，包括多层感知器、卷积神经网络、递归神经网络和强化神经网络[2,13]。

多层感知器　多层感知器（MLP）[13]也称为多层网络，是指由多层感知器组成的网络。MLP 由 3 层或更多层组成。其最简单的体系结构由输入层、隐藏层和输出层组成。网络的每个节点是一个使用非线性激活函数的神经元，以一定的权重连接到下一层的每个节点。2.3 节中描述了常用的非线性激活函数，如 Sigmoid、tanh、ReLU 等。监督学习通过基于损失最小化方法改变连接权重来优化 MLP。最著名的算法是 2.3 节中描述的反向传播算法。

卷积神经网络　卷积神经网络（CNN）[2]是 MLP 的正则化版本，广泛用于图像中对象和区域的检测、分割和识别。与 MLP 的完全连接相比，CNN 利用了数据中的分层模式，并通

过卷积层来组装更复杂的模式。卷积是一种特殊的线性变换，因此 CNN 被定义为在至少一个层中使用卷积的神经网络。在卷积阶段，每个单元通过一组称为滤波器组的权重连接到特征图中的局部块，该滤波器组在特征图中的所有单元之间共享。通常 CNN 的体系结构由卷积层、池化层和非线性变换层组成，并使用几个完全连接的层来实现最终分类。

循环神经网络　循环神经网络（RNN）被设计用于在不受大小限制的情况下接受一系列输入。它会对以往的信息进行存储，决策会受过去经验的影响。在训练过程中，RNN 会在生成输出的同时记住从之前的输入中学到的东西。输出不仅取决于当前输入，还取决于基于先前输入或输出的上下文。因此，根据过去一系列输入，相同输入可能会产生不同的输出。由于其可记忆性，RNN 适用于输出依赖于先前计算的情况，如语音识别、文本推荐或 DNA 序列分析。图 2 展示了具有隐藏状态的 RNN。

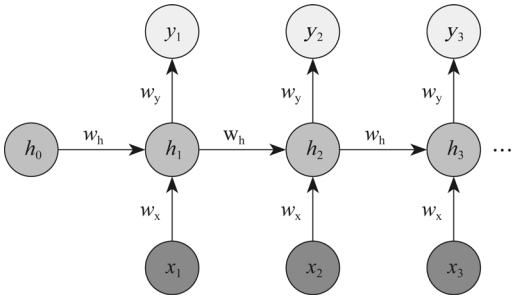

图 2　具有隐藏状态的 RNN

强化学习　强化学习（RL）[14]是机器学习的方法之一，通常被描述为马尔可夫决策过程的一种形式。与监督学习相比，强化学习没有训练数据集。因此，在缺乏现有训练数据的情况下，RL 从经验中学习。在学习过程中，通过反复试验来尝试其任务，奖励和惩罚被用作积极和消极行为的信号。此外，它不同于无监督学习，后者的目标是找到数据点之间的相似性。RL 的目标是找到一个合适的行为模型来最大化奖励。RL 的最新进展已经实现了各种应用，如游戏[14]、机器人技术[15]和生成模型[16]。

2.3　深度学习的主要计算模块

在本节中，我们介绍了深度神经网络中使用的几种主要运算模块。如图 3 所示，深度神经网络的架构由输入层、多个隐藏层和输出层组成。输入层由特征组成，而输出层代表分类。这些层以分层方式连接，前一层的输出是后一层神经元的输入。隐藏层和输出层中的每个节点（也称为神经元）都与一个系数向量和一个激活函数相关联。神经元首先计算其输入的加权求和，然后将非线性激活函数应用于求和（图 3）。

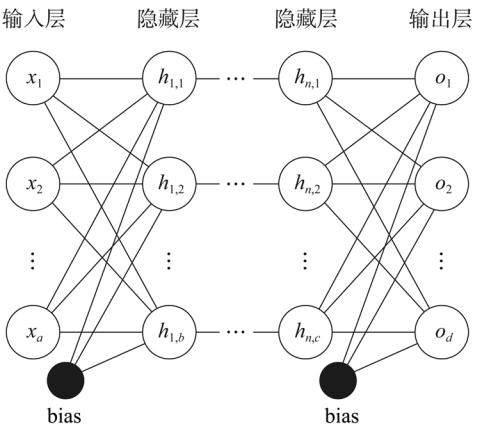

图 3　具有n个隐藏层的神经网络

线性变换　线性变换是神经网络中最常见的运算，它由矩阵加法和乘法组成，描述为

$$\boldsymbol{Y} = \boldsymbol{W}\boldsymbol{X} + \boldsymbol{B} \tag{1}$$

其中，对于每层，$\boldsymbol{W}$为权重矩阵；$\boldsymbol{X}$为输入向量；$\boldsymbol{Y}$为输出向量；$\boldsymbol{B}$为偏置向量，$\boldsymbol{X} \in R^{a\times 1}, \boldsymbol{Y} \in R^{b\times 1}, \boldsymbol{W} \in R^{b\times a}$。图 4 展示的是神经网络的第一层，激活函数有ReLU、Sigmoid、tanh等，定义为

$$\begin{aligned}
&\text{ReLU} \quad f(y)=\max(0,y)=\begin{cases} 0 & y\leqslant 0 \\ y & y>0 \end{cases} \\
&\text{Sigmoid} \quad f\left(y\right)=\frac{1}{1+\mathrm{e}^{-y}} \\
&\tanh \quad f\left(y\right)=\frac{\mathrm{e}^{2y}-1}{\mathrm{e}^{2y}+1}
\end{aligned} \tag{2}$$

激活函数　激活函数用于对神经网络中输入和输出之间的非线性转换进行建模。将线性变换和非线性激活函数结合在一起，神经网络模型可以定义为

$$\boldsymbol{Y} := (\boldsymbol{W}_L \cdot f_{L-1}(\cdots f_1(\boldsymbol{W}_1\boldsymbol{X} + \boldsymbol{B}_1)\cdots) + \boldsymbol{B}_L) \tag{3}$$

因为tanh函数和Sigmoid函数密切相关，所以tanh函数可以近似地描述为

$$\tanh\left(y\right) = 2\cdot\text{Sigmoid}\left(2y\right) - 1 \tag{4}$$

反向传播算法　神经网络的参数优化是一个非线性优化问题。梯度下降法[17-19]是最常用的算法之一。该算法基于这样的观察：如果在点a的邻域内定义多变量函数$F(x)$，该函数是可微的，那么多变量函数$\nabla F(x)$在其梯度的负方向上下降得最快。通常，梯度下降从随机选择的点开始，计算损失函数的梯度并使用梯度来更新参数。在算法收敛到局部最优值之前，此过

程不会停止。

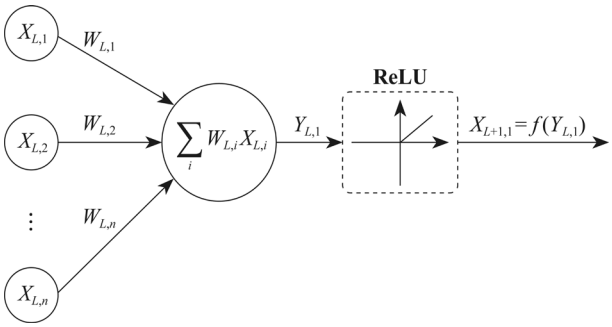

图 4　一个神经元的计算示例

反向传播[18]是一种广泛使用的算法，该算法基于前馈神经网络中的梯度下降来进行监督学习。反向传播根据链式规则计算损失函数的梯度，该梯度定义为神经网络的输出与目标值之间的总误差。每次前向通过网络后，反向传播通过调整网络权重参数执行反向传播。设$\boldsymbol{W}$为神经网络的加权向量，E为损失函数。在每次迭代中，权重可以更新为

$$\boldsymbol{W}_j = \boldsymbol{W}_j - \alpha \frac{\partial E}{\partial \boldsymbol{W}_j} \tag{5}$$

式中，α为学习率。迭代过程将继续进行，直到损失函数的输出满足预定义的精确度。

3　外包计算

本节首先回顾外包计算的发展历程，然后提出外包计算的系统模型和安全模型。

3.1　外包计算概述

近年来，随着云计算的兴起，已有大量的研究者对外包计算进行研究。外包计算协议是计算能力较弱的客户端，将部分计算安全地外包给计算能力强大的服务器。一般来说，外包计算可以分为两种类别：面向通用的解决方案和功能特定的解决方案。功能特定的解决方案侧重于特定的函数类，而通用的解决方案可应用于广泛的可计算函数。功能特定的外包计算协议通常比面向通用的外包协议效率更高。例如，基于概率可验证证明[22]、全同态加密[23]和同态消息认证[24]的通用解决方案[20-21]，由于计算和存储成本极高，很难在实际中使用。表 2 列出了与外包计算相关的综述文献。

密码学界对计算复杂度较高的运算的外包计算问题进行了深入研究。外包计算的主要安全问题之一是，客户端如何在不进行大量计算的情况下检查云服务商提供的结果的正确性。

EA[28]和计算 CS 证明[29]在外包计算中具有可验证性。

表 2　与外包计算相关的综述文献

参考文献	内容概述
[20]	面向算术电路的实用同态消息认证
[21]	可公开验证的代理计算
[25-26]	外包计算综述
[27]	基于非可信辅助设备的隐私计算
[28]	高效论证
[29]	可验证计算
[30]	同态加密及其在外包计算中的应用
[31, 33]	高次多项式函数的可验证计算
[32]	面向分布式数据集的外包计算

Benabbas 等人[31]提出了可验证的计算授权，并提出了第一个针对大型数据集上高次多项式函数的实用外包方案。随后，该团队又提出了一种新的隐私保护加密工具，称为可代理同态加密[30]，客户端可以利用该工具实现可验证的外包计算。Backes 等人[32]提出了一种解决外包数据集上外包计算问题的方法。在这个应用场景中，客户机首先将其数据外包给服务器，服务器随后需要根据外包的数据计算函数。该方案基于同态消息认证码来实现有效的验证。然而，该方案只能用于二次多项式评估。Parno 等人[21]扩展了外包代理的定义，并提出了公开可验证的外包计算原语。该原语可用于构建可公开验证的外包计算方案。Song 等人[33]在多个数据源的支持下研究了可公开验证的多项式估值外包计算。Catalano 等人[20]提出了一种同态消息认证码，允许密钥的持有者对先前认证的数据进行计算验证，以便可以使用生成的标签来证明计算的正确性。Shan 等人[25]总结了特定场景下安全外包计算的前沿技术，包括矩阵计算、数学优化等。Yu 等人[26]调查了现有的可验证性计算研究成果，并根据性能和安全要求进行利弊比较和讨论。

3.2　系统模型

Gennaro 等人[34]给出了外包计算的正式定义。具体来说，一个外包计算方案由以下定义的 4 种算法组成。

- $\mathrm{KeyGen}(F,\lambda)\rightarrow(PK,SK)$:基于安全参数$\lambda$，随机密钥生成算法生成一个对目标函数$F$进行加密的公钥$PK$，服务器可用该公钥计算函数$F$。此外，该算法还生成一个与公钥相对应的且由客户端私有的秘钥SK。

- $\text{ProbGen}_{SK}(x) \rightarrow (\sigma_x, \tau_x)$：问题生成算法使用密钥$SK$将函数的输入$x$编码为公共值$\sigma_x$，提供给服务器进行计算，并编码生成由客户端私有秘密值τ_x。

- $\text{Compute}_{PK}(\sigma_x) \rightarrow \sigma_y$：使用客户端的公钥和编码输入，服务器计算函数输出$y = F(\text{x})$的编码版本。

- $\text{Verify}_{SK}(\tau_x, \sigma_y) \rightarrow y \cup \perp$：使用密钥$SK$和秘密解码$\tau_x$，验证算法将服务器的编码输出转换为函数的输出，如$y = F(x)$，若输出$\perp$，则表明$\sigma_x$不是$F$在$x$上的有效输出。

如果问题生成算法生成的值允许诚实的服务器计算，并可以成功验证对应于这些输入上的F的评估值，那么外包计算委派方案是正确的。

3.3　安全需求

本节回顾了外包计算的安全性定义[1,21]。通常，隐私性的定义基于计算上的不可区分性来描述。

$$
\begin{aligned}
&\text{Experiment}\ \ \mathbf{Exp}_A[F,\lambda] \\
&\quad (PK, SK) \leftarrow \text{KeyGen}(F,\lambda) \\
&\quad (x_0, x_1) \leftarrow A^{\text{PubProbGen}(\cdot)}(PK) \\
&\quad (\sigma_0, \tau_0) \leftarrow \text{ProbGen}_{SK}(x_0) \\
&\quad (\sigma_1, \tau_1) \leftarrow \text{ProbGen}_{SK}(x_1) \\
&\quad b \leftarrow R\{0,1\} \\
&\quad b' \leftarrow A^{\text{PubProbGen}_{SK}(\cdot)}(PK, x_0, x_1, \rho_b) \\
&\quad \text{If}\ \ b'=b, \text{output}\ \ 1, \text{else}\ \ 0
\end{aligned}
\tag{6}
$$

定义 1（隐私性） 如果对于任何概率多项式时间攻击者A，

$$
\text{Adv}_A(F,\lambda) \leqslant \text{negl}(\lambda) \tag{7}
$$

其中，$\text{Adv}_A(F,\lambda) = \left|\Pr[b' = b] - \frac{1}{2}\right|$。

那么，该外包计算方案对于计算任务 F 满足输入输出的隐私性。在上面的实验中，攻击者 A 可以请求任何输入对应的密文。$\text{PubProbGen}_{SK}(x)$调用$\text{ProbGen}_{SK}(x)$，并返回公开参数$\sigma_x$。

可验证性　如果恶意的服务器无法说服用户接受错误的输出，就表示外包计算方案是安全的。换句话说，对于给定的函数F和输入x，恶意服务器无法使得验证算法输出 y，满足 $F(x) \neq \hat{y}$。

高效性　对于任何x和任何σ_y，盲化输入 x 所需的计算量，加上验证服务器返回结果所

需的计算量小于计算$F(x)$所需的计算量，则关于 $F()$ 的外包计算方案是高效的。

非交互性 Gennaro 等人[34]提出了一种非交互式外包计算协议，该协议中服务器返回计算结果以及非交互式证明，该证明可以在$O(m)$时间复杂度内完成验证。其中，m是输出函数F的输出长度。

4 外包深度学习

在本节中，首先介绍外包深度学习的动机、隐私性和安全性问题；其次介绍外包深度学习中常用的隐私保护技术，如同态加密、聚合签名、秘密分享等。总的来讲，现有的研究可以分为隐私保护训练和隐私保护预测两类，我们将从安全性和效率方面对相关工作进行总结、对比和分析。

4.1 外包深度学习简述

表 3 列出了与外包深度学习相关的研究成果。下面详细说明这些研究中使用的技术，并在隐私性和效率方面对这些工作进行详细比较。

外包深度学习可以理解为在服务器的帮助下完成模型的训练或预测，而不会泄露有关数据的隐私信息。因此，谷歌人工智能（Google AI）提出的联邦学习[58]也可以被视为一种外包的深度学习，其中参数由云服务器交换和聚合。与分布式数据集上所有本地数据集都上传到一个中心服务器的传统协作学习相比，联邦学习显示出许多优势。联邦学习的一般原则是对本地数据集进行训练并交换来自每个本地学习模型的参数。由于本地数据永远不会离开本地，因此联邦学习解决了数据隐私和数据访问权限等安全问题。然而，正如 Fredrikson[59]指出的，联邦学习很容易受模型反转攻击的影响，这使得敌手能够使用白盒攻击向学习模型提取关于训练数据的私人信息。为了保护分布式环境下的数据隐私，Shokri 等人[36]提出了在分布式环境下，基于部分参数共享的保护隐私分布式学习方案。Aono 等人[37]则使用加法同态加密来进一步保护训练过程中相关参数的隐私性。Ma 等人[46]提出了一种具有在恶意敌手的环境中可验证性的安全多方聚合方案，并用它来构建隐私保护多方学习算法。Zhang 等人[56]提出了一种外包的批量矩阵乘法方案，并将其扩展到隐私保护的一层感知器深度学习中。

表 3 外包深度学习计算方案的分类和相关研究成果（按照系统模型）

分类	系统模型	参考文献和技术
保护隐私的模型训练	单用户单服务器	[39] FHE+TSE [40] FHE+SGD [41] FHE+FHECS

续表

分类	系统模型	参考文献和技术
保护隐私的模型训练	多用户（分布式学习） 单服务器	[80] DHE [36] PPS [42] DP [43] SGX [44] PATE+GAN [45] SS+AE [46] MKHE [38] AHE [47] AHE+AS [48] AHE+DP [49] PRE+AS
	多用户（分布式学习） 多服务器（非串通）	[50-51] OT+GC+SS [52] SS+OT+3PC [53] MPC+SS+DP
保护隐私的预测	单服务器	[54] SMM+HE [55] MKHE
	多服务器（非串通）	[56] AHE+AS

4.2　外包深度学习中的隐私问题

近年来，深度学习的巨大进步在很大程度上归功于计算能力和存储能力的提高。而对于那些计算和存储资源有限的人来说，很难独立运行深度学习算法。因此，将深度学习任务外包给第三方提供了一种新的选择。另外，学习模型的准确性很大程度上受用于训练数据量的影响。通常，当机器学习算法在聚合数据集上运行时，提供数据的多方均会受益。最直接的多方学习方法是将数据外包给服务器，以构建中心数据集。然而，当服务器非完全可信时，这种做法将引发一系列的数据隐私问题。例如，在数据高度敏感的金融或医疗系统中，禁止与任何其他第三方共享其数据。同时，学习模型是从训练数据中抽象出来的，是需要保护的有价值的资产。已有相关研究表明，通过模型反演攻击[30,17]，攻击者可以利用对训练模型的对抗性来重建可识别的训练数据。

为了解决外包深度学习中的隐私问题，Yuan 等人[60]提出了云计算环境下的安全多方协作学习。在该方案中，双方使用 BGN 同态加密对其私有数据进行加密，并将加密后的数据上传

到云服务器。云服务器可以在密文上执行学习算法的大部分操作，并将加密的结果返回给客户端。由于 BGN 并不是一个全同态的加密[62]方案，因此每个参与者在学习过程中都必须对中间值进行解密和重新加密。所以，云服务器和每个参与者之间的交互作用带来了很高的通信开销，并且学习过程中每个参与者的解密和加密计算成本使该方案难以在实际应用场景中使用。Barni 等人[62]提出了一种保护隐私的神经网络预测方案，达到了对于用户数据以及模型数据的隐私保护目标。该方案依赖于安全两方标量乘和多项式估值等安全多方计算技术，其安全性在半诚实安全模型下得到了证明。在 4.3 节中，将从底层密码技术、安全模型、效率和性能改进等方面对相关的隐私保护研究进行分类。

4.3 外包深度学习的隐私保护技术

同态加密 加法同态加密满足以下性质：给定密文$c_1 \coloneqq E(pk, x_1)$和$c_2 \coloneqq E(pk, x_2)$，满足$E(pk, x_1 + x_2) = c_1 \oplus c_2$。这类方案的例子包括 Elgamal 加密[64]、Paillier 加密[65]和 DGK 加密[66]。给定m和$\hat{m}$的密文c、$\hat{c}$，一个乘法同态加密的例子为

$$C = (c_1, c_2) = (g^{k_1}, y^{k_1} m) \tag{8}$$

$$\hat{C} = (\hat{c}_1, \hat{c}_2) = (g^{k_2}, y^{k_2} \hat{m}) \tag{9}$$

$m\hat{m}$的密文$\bar{C}$可以计算为

$$\bar{c}_1 = c_1 \hat{c}_1 = g^{k_1 + k_2}, \bar{c}_2 = c_1 \hat{c}_1 = g^{k_1 + k_2} m\hat{m} \tag{10}$$

若明文空间很小，则通过将消息m提升为指数g^m，可以将这类乘法同态加密转移到加性同态加密，使得密文为

$$C = (c_1, c_2) = (g^{k_1}, y^{k_1} g^m) \tag{11}$$

秘密共享 秘密共享[67]是指在一组参与者之间实现多方共享秘密。该组的每个参与者被分配一个分享值，通过不同参与者的共享值实现对秘密值的重构。(n, t) −陷门秘密分享被定义为在一组n个参与者中，任何$t (t \leqslant 1)$或更多的参与者可以完成秘密值的恢复。使用最广泛的秘密共享之一是加法秘密共享，相关的运算通常在环Z_{2^l}中执行。每个分享值用一个 l 位整数表示。以两方秘密分享为例，实现秘密共享的过程如下：①随机选择一个数字$r \in_R Z_{2^l}$；②计算$(s - \mathrm{r}) \bmod 2^l$；③将共享设置为$\langle s \rangle_0 = r$和$\langle s \rangle_0 = (s - r) \bmod 2^l$。因此，秘密$s$可以重建为$s = \langle s \rangle_0 + \langle s \rangle_0 \bmod 2^l$。

不经意传输 不经意传输（OT）[68]是指安全多方计算中的一种基本密码原语。OT 协议包括发送方S和接收方R。R从S拥有的一组消息中选择并接收消息，S不知道哪个消息被选中了，R也不应该知道未被选中的消息。一种常用的不经意传输形式是$1 - \mathrm{out} - \mathrm{of} - 2\mathrm{OT}$，其中，发送方$S$具有两个 1 比特消息$x_0$和$x_1$，接收方$R$的索引位为$b$。执行协议后，$R$获得$x_b$，但无法获得关于$x_{1-b}$的任何消息。

混淆电路 混淆电路（GC）[69]是指一种通用安全两方计算技术，相互不受信任的两方联

合计算函数$f(a,b)$，除了最终输出结果，不会泄露私有输入a和b的任何信息。简单地说，将函数被描述为由基本的具有两个输出的与门和异或门电路构成的布尔电路。其中一方（Alice）构造混淆电路并发送给另一方（Bob）。Alice 为电路中的每条线路生成两个k位随机数，分别代表 0 和 1。然后，Alice 根据真值表替换 0 和 1，并将混淆真值表发送给 Bob。混淆过程可以通过对称加密来完成，如 AES。同时，Alice 将其输入对应的随机数发送给 Bob，Bob 通过 OT 协议从 Alice 接收到其输入对应的随机数。然后，Bob 在混淆电路上完成估值，计算输出层结果。最后，Bob 向 Alice 展示电路输出层对应的随机数，后者可以将它们映射回布尔值。

代理重加密　Blaze 等人[69]介绍了 PRE（代理重加密）的概念。PRE 可以用代理函数和公共代理密钥将一个密钥的密文转换为另一个密钥的密文。PRE 具有以下属性。

- 双向性：从用户A到用户B的加密允许在两个方向上重新加密。
- 非抗串通：用户A和代理方串通可以恢复用户B的密钥。
- 同态性：如果 PRE 是基于Elgamal加密体制的，那么该方案满足乘法同态性：

$$C=(c_1,c_2)=(m\cdot g^{r_1},g^{s_a r_1}) \tag{12}$$

$$\hat{C}=(\hat{c_1},\hat{c_2})=(\hat{m}\cdot g^{r_2},g^{s_a r_2}) \tag{13}$$

$m\hat{m}$的加密$\bar{C}$可以计算为

$$\bar{C}=C\otimes\hat{C}=(c_1\cdot\hat{c_1},c_2\cdot\hat{c_2})=\left(m\cdot\hat{m}\cdot g^{r_1+r_2},g^{s_a(r_1+r_2)}\right) \tag{14}$$

差分隐私　差分隐私（DP）[72-73]是隐私保障的一个强有力的标准。它保证任何分析的结果不会因单个数据项的删除或添加而受到实质性影响。换句话说，差分隐私的目标是让攻击者不能区分某个记录是否位于计算函数的数据集中。

随机函数k满足差分隐私当且仅当所有数据集D和D'最多有一个数据不同，并且对于所有的$S\subseteq \mathrm{Range}(k)$。

$$\Pr[k(D)\in S]\leqslant \exp(\varepsilon)\times\Pr[k(D')\in S] \tag{15}$$

式中，ε为隐私预算。Dwork 等人[72]提出了一种用拉普拉斯机制保护任意函数f的差分隐私的方法。该机制利用函数f计算任意两个相邻数据集的全局敏感度，并在输出结果上注入合适的随机噪声。

4.4　分类标准

根据外包深度学习的不同阶段，我们将外包深度学习分为隐私保护训练外包和隐私保护预测外包。表 3 描述了现有的隐私保护外包深度学习方案的分类，相关缩写代表的含义均列于表 4。

在单用户外包深度学习计算方案中，通常使用同态加密[23]对原始数据集进行加密，同时允许云服务器在无须解密的情况下对加密的数据运行深度学习算法。然而，使用同态加密来

处理整个数据集，并在加密的数据集上运行深度学习算法会给外包深度学习造成很高的延迟。相关研究表明，在基于神经网络的深度学习中，通过在训练集中加入更多的数据，可以显著提高精度。考虑到数据的隐私性，基于分布式数据集的机器学习出现了新的挑战。对于保护隐私的多方学习，学者们做了大量的研究工作。安全多方计算、秘密共享、同态加密和差分隐私是隐私保护协作学习中常用的密码学工具，它们各有利弊。4.5 节将对基于密码技术的外包深度学习算法进行比较。

表 4　缩写词（按字母顺序排列）

首字母缩略词	说明
3PC	安全三方计算
AHE	加法同态加密
AE	认证加密
AS	聚合签名
DHE	双同态加密
DP	差分隐私
FHE	全同态加密
FHECS	全同态加密的变换
GAN	生成对抗网络
GC	混淆电路
HE	同态加密
MKHE	多密钥同态加密
MPC	多方计算
OT	不经意传输
PATE	基于教师模型的集成学习
PPS	部分参数共享
PRE	代理重加密
RG	重复冈珀茨
RP	随机映射
PAHE	批量加法同态加密
SMC	安全多方计算
SGX	安全保护模块

续表

首字母缩略词	说明
SGD	随机梯度下降
SS	安全共享
SMM	安全矩阵乘法
TSE	泰勒级数展开
USC	无条件安全比较协议

4.5　保护隐私的深度学习训练过程外包

如表 3 所示，深度学习训练过程的计算外包可以进一步分为单用户模型和多用户模型。此外，根据方案中涉及的服务器数量，又可以分为单服务器模型和多服务器（非串通）模型。

单用户模型　Zhang 等人[56]提出了一种基于 FHE 的外包深度学习方案，通过将大部分计算训练任务外包给半诚实的云服务器来提高学习效率。为了保护数据和模型隐私，训练数据和初始化的神经网络在上传到云服务器之前要使用 FHE 进行加密。此外，非线性激活函数利用泰勒级数展开，利用多项式进行逼近。基于 FHE 的同态性，云服务器可以在密文上实现训练过程，并将加密的模型返回给数据拥有者。然而，每轮迭代之后客户端必须对从云服务器返回的密文进行解密并重新加密，造成了数据所有者较高的计算和通信开销。

为了实现完全非交互性，Nandakumar 等人[39]提出了一种基于全同态加密的随机梯度下降算法，该方案引入了神经网络剪枝、数据表示以及密文批处理技术等。在该方案中，数据所有者对训练数据进行加密，初始化神经网络模型，并将这些密文上传到云服务器。之后，数据所有者可以脱机，直到云服务器返回学习的模型。虽然该方案的整体分类准确率约为 97.8%，但高延迟性严重限制了其大规模应用。Lou 等人[40]提出了 Glyph，该方案也依赖于全同态加密，但通过在 TFHE[75]和 BGV 密码系统[23]之间的切换，提高了效率和准确性。这种转换的正确性源于最近的工作[76]，作者证明了通过同态操作在 TFHE 和 BGV 之间进行组合与切换的可行性。非线性激活函数由逻辑运算友好的 TFHE 实现，而矩阵的相关运算由算术运算友好的 BGV 执行。此外，方案还纳入了迁移学习，减少可训练层数，进一步减少训练延迟。实验结果表明，Glyph 在各种加密数据集上比之前降低了 99%的训练延迟（表 5）。这些方案的安全模型都是诚实且好奇的，而且所有的方案都存在激活函数近似计算造成的精度损失。

多方学习（单服务器模型）基于多数据集的多方深度学习具有广泛的应用场景。然而，从分布式数据集中收集海量数据带来了新的隐私问题。如何使多个数据拥有者能够联合一个训练一个全局模型而不泄露各自的数据隐私，是该研究领域需要解决的一个关键问题。相关

学者在保护隐私的多方学习方面已经做了大量研究工作。

Yuan 等人[60]提出了云计算中的神经网络学习，它可以对任意分区的数据进行两方学习。该方案基于 Boneh 等人[70]提出的一种特殊的同态加密，该加密支持一次乘法和无限数量的加法运算。Sigmoid 是用麦克劳林级数展开方法的多项式逼近，但这种逼近引入了学习模型的精度损失。此外，该方案是完全交互式的，因为参与者必须在每轮的学习过程中解密文本并重新加密它们，所以计算开销和通信开销与神经网络的复杂度、数据库的大小和参与者的数量呈线性相关。

表 5　单用户外包深度学习计算方案的比较

参考文献	主要方法	交互性
[11]	使用 BGV 加密数据	交互式
[74]	全同态加密输入数据以及网络模型 网络剪枝和密文打包	非交互式
[38]	基于不同密钥的全同态密文转换 知识迁移	交互式

Shokri 等人[36]提出了一种通过部分参数共享的方法来实现多方深度学习，该方案对联邦学习的研究产生了积极的影响。该方案基于分布式选择式的随机梯度下降算法，参与者首先在数据集上进行单独训练，在每轮迭代中进行选择性的参数共享，并控制共享参数范围以及共享频率，实现了实用性和隐私保护的折中。实验结果表明，即使只有 1%的参数共享，该方案的学习精度也高于独立学习。Shokri 等人[36]提出的部分参数共享的方法存在隐私泄露问题，攻击者可以根据部分梯度参数发起逆向攻击，以推测训练数据的真实值。Shokri 等人[36]进一步展示了如何利用差分隐私技术抵抗隐私泄露。但差分隐私保护机制降低了学习的准确性。为了在不降低模型准确性的情况下实现隐私保护，Aono 等人[13]提出基于同态加密技术对梯度参数进行加密的方法。该方案没有泄露参与者梯度的信息，但会增加计算和通信成本。此外，在该方案中，所有的参与方共享同态加密私钥，容易受到合谋攻击。Li 等人[45]提出了一种解决该问题的方法，基本思想是使用多密钥全同态加密对梯度进行加密。多密钥全同态加密的优点在于不需要在所有参与者之间共享一个密钥。每个参与者拥有一对公私钥。局部梯度使用不同的密钥加密并上传到云服务器，基于多密钥同态性，多的用户产生的密文仍然可以聚合在一起。在收到加密结果时，所有参与者通过安全多方计算共同对其进行解密。考虑到多密钥全同态加密的计算复杂性和深度学习的网络复杂性，该方案难以应用在实际场景中。此外，上述相关研究的安全性模型是半诚实的，为了解决恶意参与者环境下的安全多方学习，Ma 等人[55]利用聚合签名的分布式学习方案。在该方案中，当参与者将加密的本地梯度上传到云服务器时，也会生成梯度的签名。利用聚合后的签名验证聚合梯度的正确性，但该方案

无法抵御服务器和参与方的串通。单服务器模型下的多方学习方案对比见表 6。

表 6　单服务器模型下的多方学习方案对比

参考文献	主要概念	交互性	安全模型
[61]	双数据提供方 同态加密处理数据和模型参数 麦克劳林级数展开	交互式	半诚实
[36]	梯度共享 异步随机梯度下降	交互式	半诚实
[45]	加法同态算法对局部梯度进行加密	交互式	半诚实
[44]	多密钥全同态处理数据和模型参数 随机梯度下降	非交互式	半诚实
[46]	加法同态算法来加密局部梯度 聚合签名来验证结果 异步随机梯度下降	交互式	恶意
[78]	对称同态来加密局部梯度 差分隐私技术来扰动局部梯度 异步随机梯度下降	交互式	半诚实 抗串通
[41]	知识迁移来训练学生模型 差分隐私扰动教师模型的输出	非交互式	半诚实
[43]	知识迁移和差分隐私来训练学生模型 半监督学习	非交互式	半诚实
[56]	知识迁移来训练模型 代理重新加密来加密本地模型的输出	非交互式	恶意
[44]	秘密共享和双随机参数扰动	交互式	恶意
[42]	基于 SGX 的计算框架	非交互式	恶意

Hao 等人[47]提出了一种能够抵御合谋攻击的联邦学习方案。该方案使用了对称同态加密和差分隐私技术，基于异步随机梯度下降方法实现了任意参与方退出情况下依然可以训练出高精度的模型。在该方案中，对于训练过程的每个阶段，每个用户计算局部梯度G_μ并使用拉普拉斯噪声$\mathrm{Lap}(\frac{\Delta f}{\epsilon})$扰动它。随后，将$C_\mu = \mathrm{Enc}_{sk}(G_\mu + \mathrm{Lap}(\frac{\Delta f}{\epsilon}))$发送给服务器，由服务器执行

聚合操作 $C_{\text{add}} = C_1 + C_2 + \cdots + C_n = \text{Enk}_{sk}(\sum_{\mu=1}^{n} G_\mu)$。基于拉普拉斯噪声的对称性，聚合时扰动数据将得到清除。Hamm 等人[41]提出了一种基于知识迁移和差分隐私的安全多方学习方案。在该框架中，研究者假设存在一个可信任的权威机构，负责收集本地训练的教师模型。教师模型用于为非敏感的且未标记的辅助数据生成标签。作者提出了几种标签的生成方案，并基于该标记数据集训练新的学生模型。为了保护提升学生模型的隐私性，可以在模型输出端添加差分隐私噪声。相比于非隐私保护的多方学习方案，其泛化误差和收敛速度为$O(\epsilon^{-2}M^{-2})$和$O(N^{-1})$，其中N为未标记辅助数据的数量，M为参与者的数量。然而，该方案存在致命的弱点，即本地分类器是由完全可信的中心服务器收集的。这个假设在实际的应用程序场景中难以实现。Fredrikson 等人[59]指出，模型输出或内部参数可以用来推断关于训练数据的敏感信息。Papernot 等人[43]提出了新的方案来解决上述隐私泄露问题。通过在投票数据中添加差分隐私噪声，提升了该标签数据的隐私性，增加了逆向攻击的难度。为了进一步减少学生/教师在学习过程中的查询次数，作者提出了使用生成式对抗网络（GAN）进行半监督学习。实验评价表明，该方案对 MNIST 数据集在（2.04，10^{-5}）差分隐私约束下的准确率高达 98%。

Bonawitz 等人[44]研究了数据提供方较容易掉线的移动互联网下的安全多方学习方案，该方案具有通信开销低和一定的鲁棒性。该设计背后的主要思想是双重随机化思想，在门陷秘密共享的基础上，利用盲因子进一步隐藏原始输入数据。盲因子通过 Diffie-Hellman 协议生成，并可以在聚合阶段去除。该方案存在聚合结果泄露给服务器的问题，同时在大数据集的通信成本和运行效率较低。为进一步降低安全多方计算或全同态加密构建的隐私保护机器学习方案的计算成本，Ohrimenko 等人[42]在 USENIX2016 提出了使用可信硬件 SGX 处理器的方案。利用 SGX 中的 enclaves 保护相关隐私数据。在该框架下，数据拥有者将加密的数据上传到不受保护的内存，但将其加密密钥上传到 SGX 处理器。该框架保证只有可信处理器内部的机器学习代码才能直接访问数据，并且学习到的模型将以加密的形式返回给每个数据拥有者。实验结果表明，该方案的计算代价比基于安全多方计算或全同态加密技术的方法快多个数量级。

多方学习（多服务器、非串通） Chase 等人[52]提出了基于多方计算和差分隐私多方学习。在分布式场景中，关键是如何生成差分隐私噪声。如果每个参与方根据其局部数据计算局部梯度，并添加随机噪声以使其具有差分隐私性，那么基于平均梯度计算的最终学习模型将会有较差准确性。在该方案中，研究者提出利用秘密共享将梯度参数分配给两个非串通服务器H_1和H_2，服务器使用安全两方计算输出满足差分隐私要求的梯度参数。Mohassel 和 Zhang 等人[50,56]提出了在双服务器模型下安全多方深度学习[49]。该方案基于秘密分享，在非串通安全模型下证明了安全性。随后，Mohassel 等人[51]提出了 3 台服务器模式下外包深度学习的一般框架（ABY）。训练数据通过秘密共享分配在 3 台服务器，并通过安全三方计算来训练学习模

型。在三方学习过程中，研究者提出了算术运算、布尔运算以及混淆电路运算之间的有效切换技术，大大地提高了系统的计算效率。该方案在假设只有一个恶意服务器的情况下可以防止恶意攻击者。实验结果表明，ABY3 的计算效率比 SecureML 快 55000 倍。安全多方深度学习的大部分研究都集中于如何将安全技术应用于已有的训练算法，或者应用通用的安全多方计算协议来实现保护隐私的机器学习。这些方法存在离线计算开销高、在线通信成本高、准确性损失的问题。为了解决这些问题，Agrawal 等人[49]提出了一种在半诚实模型下基于双服务器多方学习方法 Quotient。在该方案中，神经网络权重在训练过程中被标准化为-1、0、1。作者基于 OT、布尔共享和加法共享构造了一种三值矩阵向量乘法快速计算方法。与 SecureML 相比，Quotient 在 WAN 上的时间提高了 50 倍，准确性提高了 6%（见表 7）。

表 7　多服务器（非串通）模型下多方学习方案

参考文献	主要思路	交互性	安全模型
[53]	面向非串通服务器的秘密共享 基于安全两方计算的差分隐私梯度聚合 异步学习	交互式	半诚实
[51]	将训练数据共享到两个非训练服务器 在共享的十进制数上的算术 激活函数的多方计算 向量化	非交互	半诚实
[52]	保护隐私的机器学习在一个 3 台服务器模型与 1 个已损坏的服务器 在三方设置中，在算术、二进制和 Yao 三方计算之间的有效切换 在三方设置中对共享的十进制数的定点乘法	非交互	恶意
[50]	两个非合并的服务器 同时优化训练算法和安全协议 秘密共享中的网络权值和矩阵向量乘法	非交互	半诚实

4.6　隐私保护预测外包

隐私保护预测外包是外包深度学习中的一种新形式。如图 5 所示，这个新框架涉及三方：模型提供商、服务器和数据提供方（客户端）。在这个新的框架中，模型提供者为服务器提供了训练好的加密模型，服务器可以基于该模型向客户端提供预测服务。为了保证隐私，模型和客户端的数据在上传到服务器之前需要进行加密处理。类似的研究有不经意神经网络预测[63,81-82]，其中模型所有者直接为客户提供预测服务。在这个应用程序场景中，“不经意”意味着客户端只能获得预测结果，同时保证输入数据的隐私性。

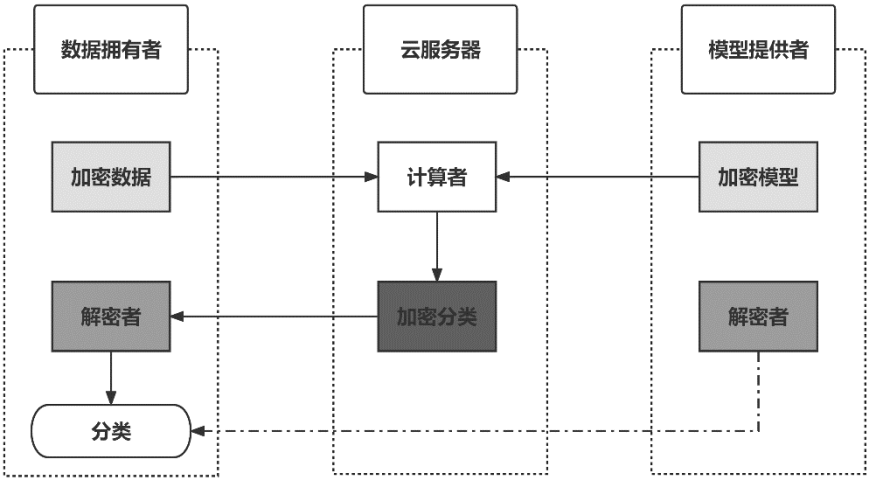

图 5　隐私保护预测外包框架

Ma 等人[46]提出了一种非交互式的预测外包方案。在该方案中，模型提供商将预先训练过的模型利用加法分享技术发送到两个非串通服务器，由服务器向用户提供预测服务。该预测过程对于数据所有者来说是完全非交互式的，并在半诚实和非串通的安全模型下被证明了安全性。为了提升预测外包方案的计算效率，Jiang 等人[53]提出了一种矩阵运算的外包计算方案，并将其应用于隐私保护预测中。对于两个大小为$d \times d$的矩阵，之前的工作需要$O(d^2)$的同态运算，该论文所提出的矩阵乘法解只需要$O(d)$。同时，利用 SIMD 技术、矩阵编码方法、同态运算相结合，提高了方案的计算效率。在实验评估中，作者选择了从 MNIST 中训练出来的加密 CNN 模型，该模型包括一个卷积层、两个具有平方激活函数的全连接层。实验结果表明，每幅图像处理需要 0.45s，预测精度达到 98.1%。但该框架在单密钥场景中工作，网络模型和输入数据必须使用相同的加密密钥进行加密。

Chen 等人[54]提出了一种新的多密钥同态加密方案，基于该方案的隐私保护预测外包方案比以往的方案更加高效和灵活。模型提供商首先利用自己的公钥对预先训练过的模型进行加密，并将其上传到云服务器。其次数据所有者使用其公钥对其数据进行加密。最后云服务器在密文上计算相关的预测结果，并由模型提供商、服务器和客户端共同解决。由于采用了多密钥同态加密，该预测服务适用于每个数据所有者向服务器提供部分数据的分布式应用场景，并对聚合数据进行预测评估（表 8）。

表 8　隐私保护预测相关方案的对比分析

参考文献	主要方法	交互性	安全模型
[47]	秘密地共享预先训练好的模型到两个非串通的服务器 加法同态加密数据 安全比较协议计算激活函数	非交互	半诚实 非串通
[54]	安全的外包矩阵运算、矩阵编码 SIMD	非交互	半诚实
[55]	多密钥全同态加密 加密模型和数据	预测值的计算（非交互式） 分布式解密（交互式）	半诚实

5　总结和未来研究展望

本文对外包计算和深度学习两个研究领域间的交叉进行了总结，阐述了外包计算和深度学习之间的关系，讨论了外包深度学习中使用的主要密码技术，并对相关方案在计算效率、安全模型和隐私性等方面进行了对比分析。在以下几个方面可以做进一步的研究。

（1）面向恶意参与者的联邦学习。联邦学习的主要思想是从多个分布式数据集构建机器学习模型，同时防止原始数据泄露。通常有一个参数服务器，负责协调分布式学习任务，聚合参与者的相关参数，并更新全局模型。以往对联邦学习的研究大多基于数据提供方半诚实的假设。然而，现实世界中的应用经常会遇到这样的情况，即某些节点可能不可靠或恶意。传统的联邦学习在拜占庭攻击下显得异常脆弱。并且已有相关研究表明，参数更新会泄露参与者训练数据的隐私信息，攻击者利用这种泄露来发起主动攻击。因此，如何构建一个能够抵御拜占庭攻击，同时保证学习过程中的数据隐私性是该研究领域的热点问题。

（2）机器学习的模型反转攻击包括黑盒攻击和白盒攻击。黑盒攻击只需要谕言访问经过训练的模型；而白盒攻击指的是完全了解训练机制并访问模型参数的攻击。在不经意神经网络预测方案中，黑盒攻击可以通过多次查询服务器来实现。Abadi 等人[61]提出了一种具有差分隐私属性的深度学习方案来应对模型反演攻击。然而，差分隐私虽然提供了严格的隐私保障，但也带来了很大的性能损失。因此，如何在保持预测准确性的前提下保证数据隐私性有待进一步研究。

（3）效率、隐私和效用之间的平衡。外包过程中的隐私保护措施，如数据加密、盲化或扰动等，造成了效用损失和计算效率的大大降低。隐私保护对其效率的主要影响是计算和通信开销。虽然同态加密和安全多方计算技术能够实现隐私计算，但也带来了较高的计算开销。

差分隐私技术具有较好的计算效率，但精度损失较高。因此，研究保护隐私的机器学习中的效率、隐私、效用平衡的方法，是未来该领域的研究难点。

参考文献

[1] ALPAYDIN E. Introduction to machine learning[M]. MIT press, 2020.

[2] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. MIT press, 2016.

[3] SCHMIDHUBER J. Deep learning in neural networks: An overview[J]. Neural networks, 2015, 61: 85-117.

[4] LIU W, WANG Z, LIU X, et al. A survey of deep neural network architectures and their applications[J]. Neurocomputing, 2017, 234: 11-26.

[5] GRAVES A, MOHAMED A, HINTON G. Speech recognition with deep recurrent neural networks[C]. Proceedings of the 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013: 6645-6649.

[6] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups[J]. IEEE Signal processing magazine, 2012, 29(6): 82-97.

[7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[8] TAIGMAN Y, YANG M, RANZATO M A, et al. Deepface: Closing the gap to human-level performance in face verification[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 1701-1708.

[9] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. nature, 2015, 521(7553): 436-444.

[10] ARULKUMARAN K, DEISENROTH M P, BRUNDAGE M, et al. Deep reinforcement learning: A brief survey[J]. IEEE Signal Processing Magazine, 2017, 34(6): 26-38.

[11] ZHANG Y, YANG Q. A survey on multi-task learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 34(12): 5586-5609.

[12] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Medical image analysis, 2017, 42: 60-88.

[13] DENG L. A tutorial survey of architectures, algorithms, and applications for deep learning[J]. APSIPA transactions on Signal and Information Processing, 2014, 3: 1-29.

[14] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. nature, 2016, 529(7587): 484-489.

[15] GU S, HOLLY E, LILLICRAP T, et al. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates[C]. Proceedings of the 2017 IEEE international conference on robotics and automation. IEEE, 2017: 3389-3396.

[16] YU L, ZHANG W, WANG J, et al. Seqgan: Sequence generative adversarial nets with policy gradient[C]. Proceedings of the AAAI conference on artificial intelligence. 2017, 31(1): 2852-2858.

[17] FAHLMAN S E. Faster learning variations of back propagation: An empirical study[C]. Proceedings of the 1988 connectionist models summer school. Morgan Kaufmann, 1988: 38-51.

[18] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning internal representations by error propagation[R]. California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[19] AVRIEL M. Nonlinear programming: analysis and methods[M]. Courier Corporation, 2003.

[20] CATALANO D, FIORE D. Practical homomorphic MACs for arithmetic circuits [C]. Proceedings of the Advances in Cryptology-EUROCRYPT 2013: 32nd Annual Internati onal Conference on the Theory and Applications of Cryptographic Techniques, Athens, Gree ce, May 26-30, 2013. Proceedings 32. Springer Berlin Heidelberg, 2013: 336-352.

[21] PARNO B, RAYKOVA M, VAIKUNTANATHAN V. How to delegate and verify in public: Verifiable computation from attribute-based encryption[C]. Proceedings of the Theory of Cryptography: 9th Theory of Cryptography Conference, TCC 2012, Taormina, Sicily, Italy, March 19-21, 2012. Proceedings 9. Springer Berlin Heidelberg, 2012: 422-439.

[22] BELLARE M, GOLDWASSER S, LUND C, et al. Efficient probabilistically checkable proofs and applications to approximations[C]. Proceedings of the twenty-fifth annual ACM symposium on Theory of computing. 1993: 294-304.

[23] BRAKERSKI Z, GENTRY C, VAIKUNTANATHAN V. (Leveled) fully homomorphic encryption without bootstrapping[J]. ACM Transactions on Computation Theory, 2014, 6(3): 1-36.

[24] GENNARO R, WICHS D. Fully homomorphic message authenticators[C]. Proceedings of the Advances in Cryptology-ASIACRYPT 2013: 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part II 19. Springer Berlin Heidelberg, 2013: 301-320.

[25] SHAN Z, REN K, BLANTON M, et al. Practical secure computation outsourcing: A survey[J]. ACM Computing Surveys, 2018, 51(2): 1-40.

[26] YU X, YAN Z, VASILAKOS A V. A survey of verifiable computation[J]. Mobile Networks and Applications, 2017, 22: 438-453.

[27] MATSUMOTO T, KATO K, IMAI H. Speeding up secret computations with insecure auxiliary devices[C]. Proceedings of the Advances in Cryptology—CRYPTO'88: Proceedings 8. Springer New York, 1990: 497-506.

[28] KILIAN J. Improved Efficient Arguments: Preliminary version[C]. Proceedings of the Advances in Cryptology—CRYPT0'95: 15th Annual International Cryptology Conference Santa Barbara, California, USA, August 27-31, 1995 Proceedings 15. Springer Berlin Heidelberg, 1995: 311-324.

[29] MICALI S. CS proofs[C]. Proceedings 35th Annual Symposium on Foundations of Computer Science. IEEE, 1994: 436-453.

[30] BARBOSA M, FARSHIM P. Delegatable Homomorphic Encryption with Applications to Secure Outsourcing of Computation[C]. Proceedings of the Cryptographers' Track at RSA Conference. 2012, 7178: 296-312.

[31] BENABBAS S, GENNARO R, VAHLIS Y. Verifiable delegation of computation over large datasets[C]. Proceedings of the Advances in Cryptology-CRYPTO 2011: 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings 31. Springer Berlin Heidelberg, 2011: 111-131.

[32] BACKES M, FIORE D, REISCHUK R M. Verifiable delegation of computation on outsourced data[C]. Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. 2013: 863-874.

[33] SONG W, WANG B, WANG Q, et al. Publicly verifiable computation of polynomials over outsourced data with multiple sources[J]. IEEE Transactions on Information Forensics and Security, 2017, 12(10): 2334-2347.

[34] GENNARO R, GENTRY C, PARNO B. Non-interactive verifiable computing: O utsourcing computation to untrusted workers[C]. Proceedings of the Advances in Cryptology -CRYPTO 2010: 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings 30. Springer Berlin Heidelberg, 2010: 465-482.

[35] JUVEKAR C, VAIKUNTANATHAN V, CHANDRAKASAN A. {GAZELLE}: A low latency framework for secure neural network inference[C]. Proceedings of the 27th {USENIX} Security Symposium. 2018: 1651-1669.

[36] SHOKRI R, SHMATIKOV V. Privacy-preserving deep learning[C]. Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. 2015: 1310-1321.

[37] AONO Y, HAYASHI T, WANG L, et al. Privacy-preserving deep learning via additively homomorphic encryption[J]. IEEE Transactions on Information Forensics and Security, 2017, 13(5):

1333-1345.

[38] ZHANG Q, YANG L T, CHEN Z. Privacy preserving deep computation model on cloud for big data feature learning[J]. IEEE Transactions on Computers, 2015, 65(5): 1351-1362.

[39] NANDAKUMAR K, RATHA N, PANKANTI S, et al. Towards deep neural network training on encrypted data[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2019: 40-48.

[40] LOU Q, FENG B, CHARLES FOX G, et al. Glyph: Fast and accurately training deep neural networks on encrypted data[J]. Advances in Neural Information Processing Systems, 2020, 33: 9193-9202.

[41] HAMM J, CAO Y, BELKIN M. Learning privately from multiparty data[C]. Proceedings of the International Conference on Machine Learning. 2016: 555-563.

[42] OHRIMENKO O, SCHUSTER F, FOURNET C, et al. Oblivious multi-party machine learning on trusted processors[C]. Proceedings of the USENIX Security Symposium. 2016, 16: 10-12.

[43] PAPERNOT N, ABADI M, ERLINGSSON U, et al. Semi-supervised knowledge transfer for deep learning from private training data[J]. arXiv preprint arXiv:1610.05755, 2016.

[44] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning[C]. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 1175-1191.

[45] LI P, LI J, HUANG Z, et al. Multi-key privacy-preserving deep learning in cloud computing[J]. Future Generation Computer Systems, 2017, 74: 76-85.

[46] MA X, ZHANG F, CHEN X, et al. Privacy preserving multi-party computation delegation for deep learning in cloud computing[J]. Information Sciences, 2018, 459: 103-116.

[47] HAO M, LI H, XU G, et al. Towards efficient and privacy-preserving federated deep learning[C].Proceedings of the IEEE international conference on communications. IEEE, 2019: 1-6.

[48] MA X, JI C, ZHANG X, et al. Secure multiparty learning from the aggregation of locally trained models[J]. Journal of Network and Computer Applications, 2020, 167: 102754.

[49] AGRAWAL N, SHAHIN SHAMSABADI A, KUSNER M J, et al. QUOTIENT: two-party secure neural network training and prediction[C]. Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019: 1231-1247.

[50] MOHASSEL P, ZHANG Y. Secureml: A system for scalable privacy-preserving machine learning[C]. Proceedings of the 2017 IEEE symposium on security and privacy. IEEE, 2017: 19-38.

[51] MOHASSEL P, RINDAL P. ABY3: A mixed protocol framework for machine

learning[C].Proceedings of the 2018 ACM SIGSAC conference on computer and communications security. 2018: 35-52.

[52] CHASE M, GILAD-BACHRACH R, LAINE K, et al. Private collaborative neural network learning[J]. Cryptology ePrint Archive, 2017.

[53] JIANG X, KIM M, LAUTER K, et al. Secure outsourced matrix computation and application to neural networks[C]. Proceedings of the 2018 ACM SIGSAC conference on computer and communications security. 2018: 1209-1222.

[54] CHEN H, DAI W, KIM M, et al. Efficient multi-key homomorphic encryption with packed ciphertexts with application to oblivious neural network inference[C]. Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019: 395-412.

[55] MA X, CHEN X, ZHANG X. Non-interactive privacy-preserving neural network prediction[J]. Information Sciences, 2019, 481: 507-519.

[56] ZHANG X, JIANG T, LI K C, et al. New publicly verifiable computation for batch matrix multiplication[J]. Information Sciences, 2019, 479: 664-678.

[57] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: Strategies for improving communication efficiency[J]. arXiv preprint arXiv:1610.05492, 2016.

[58] FREDRIKSON M, LANTZ E, JHA S, et al. Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing[C]. Proceedings of the 23rd {USENIX} Security Symposium. 2014: 17-32.

[59] FREDRIKSON M, JHA S, RISTENPART T. Model inversion attacks that exploit confidence information and basic countermeasures[C]. Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. 2015: 1322-1333.

[60] YUAN J, YU S. Privacy preserving back-propagation neural network learning made practical with cloud computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 25(1): 212-221.

[61] ABADI M, CHU A, Goodfellow I, et al. Deep learning with differential privacy[C]. Proceedings of the 2016 ACM SIGSAC conference on computer and communications security. 2016: 308-318.

[62] BARNI M, ORLANDI C, PIVA A. A privacy-preserving protocol for neural-network-based computation[C]. Proceedings of the 8th workshop on Multimedia and security. 2006: 146-151.

[63] ELGAMAL T. A public key cryptosystem and a signature scheme based on discrete logarithms[J]. IEEE transactions on information theory, 1985, 31(4): 469-472.

[64] PAILLIER P. Public-key cryptosystems based on composite degree residuosity

classes[C]. Proceedings of the Advances in Cryptology—EUROCRYPT'99: International Conference on the Theory and Application of Cryptographic Techniques Prague, Czech Republic, May 2-6, 1999 Proceedings 18. Springer Berlin Heidelberg, 1999: 223-238.

[65] DAMGARD I, GEISLER M, KROIGARD M. Homomorphic encryption and secure comparison[J]. International Journal of Applied Cryptography, 2008, 1(1): 22-31.

[66] SHAMIR A. How to share a secret[J]. Communications of the ACM, 1979,22(11): 612-613.

[67] GILBOA N. Two party RSA key generation[C]. Proceedings of the Advances in Cryptology—CRYPTO'99: 19th Annual International Cryptology Conference Santa Barbara, California, USA, August 15-19, 1999 Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999: 116-129.

[68] YAO A C C. How to generate and exchange secrets[C]. Proceedings of the 27th annual symposium on foundations of computer science. IEEE, 1986: 162-167.

[69] Blaze M, Bleumer G, Strauss M. Divertible protocols and atomic proxy cryptography[C]. Proceedings of the Advances in Cryptology—EUROCRYPT'98: International Conference on the Theory and Application of Cryptographic Techniques Espoo, Finland, May 31-June 4, 1998 Proceedings 17. Springer Berlin Heidelberg, 1998: 127-144.

[70] BONEH D, GENTRY C, LYNN B, et al. Aggregate and verifiably encrypted signatures from bilinear maps[C]. Proceedings of the Advances in Cryptology—EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4-8, 2003 Proceedings 22. Springer Berlin Heidelberg, 2003: 416-432.

[71] DWORK C. Differential privacy: A survey of results[C]. Proceedings of the Theory and Applications of Models of Computation: 5th International Conference, TAMC 2008, Xi'an, China, April 25-29, 2008. Proceedings 5. Springer Berlin Heidelberg, 2008: 1-19.

[72] DWORK C, KENTHAPADI K, MCSHERRY F, et al. Our data, ourselves: Privacy via distributed noise generation[C]. Proceedings of the Advances in Cryptology-EUROCRYPT 2006: 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28-June 1, 2006. Proceedings 25. Springer Berlin Heidelberg, 2006: 486-503.

[73] BRICKELL J, PORTER D E, SHMATIKOV V, et al. Privacy-preserving remote diagnostics[C]. Proceedings of the 14th ACM conference on Computer and communications security. 2007: 498-507.

[74] CHILLOTTI I, GAMA N, GEORGIEVA M, et al. TFHE: fast fully homomorphic

encryption over the torus[J]. Journal of Cryptology, 2020, 33(1): 34-91.

[75] BOURA C, GAMA N, GEORGIEVA M, et al. Chimera: Combining ring-lwe-based fully homomorphic encryption schemes[J]. Journal of Mathematical Cryptology, 2020, 14(1): 316-338.

[76] MCKEEN F, ALEXANDROVICH I, BERENZON A, et al. Innovative instructions and software model for isolated execution[J]. Hasp@ isca, 2013, 10(1).

基于机器学习的侧信道攻击方法研究进展报告

郁昱[1]，王伟嘉[2]

1.上海交通大学，上海，200030

2.山东大学，网络空间安全学院，青岛，266237

通讯作者：王伟嘉，E-mail：wjwang@sdu.edu.cn

摘　要：侧信道攻击（Side-Channel Attack）是一类额外利用如功耗、时间等物理侧信道信息的密码分析方法。它可以绕开密码在数学上的安全性，直接针对实现密码算法的硬件设备进行攻击，从而获得其内部的密钥。由于侧信道攻击对现实密码设备的强大攻击能力，已经给网络空间安全带来了严重威胁。国内外也推出了多个标准来规范密码实现的侧信道安全性。对侧信道攻击的研究可以更全面地发现密码实现的潜在安全问题，理解侧信道泄露的本质，从而更好地进行安全防护。机器学习（Machine Learning）是人工智能的一部分，专门分析和解释数据的模式及结构。研究发现，机器学习的应用可以有效地解决侧信道攻击中的诸多问题，近 15 年来，这一直是研究重点。本文把侧信道攻击分为泄露预处理、泄露建模、密钥恢复 3 个过程，梳理机器学习相关技术在这 3 个过程中的各类应用。

关键词：侧信道攻击；机器学习；统计学习；深度学习

Review of Machine learning-based Side-channel Attack Approaches

YU Yu[1]，WANG Weijia[2]

1. Shanghai Jiao Tong University, Shanghai, 200030, China

2.School of Cyber Science and Technology, Shandong University, Qingdao, 266237, China

Corresponding author: WANG Weijia，E-mail：wjwang@sdu.edu.cn

Abstract: Side-channel attack is a type of cryptographic analysis method that additionally utilizes side-channel information such as power consumption and timing. It can bypass the theoretical security of the cryptographic algorithm, and target the cryptographic hardware to achieve

its internal secret key. Side-channel attack has been posed an important threat to cyberspace security. Several standards have been enacted to enhance the side-channel security of cryptographic implementations. The research on side-channel attacks can help to discover the potential security issues of cryptographic implementations, understand the nature of side-channel leakage, and thus come up with better security protections. Machine learning is seen as a part of artificial intelligence. It focuses on analyzing and explaining the features and models of data. It has been found that the application of machine learning is able to solve many difficulties in the methods of side-channel attack, attracting much attention in the past 15 years. This article divides the side-channel attack into three phases, namely pre-processing, profiling and key recovering, and summarizes the application of machine learning to the side-channel attack based on the three phases.

Keywords: Side-Channel Attack, Machine Learning, Statistical Learning, Deep Learning

1 引言

1.1 侧信道攻击简介

密码是保障网络空间安全的核心技术和基础支撑，直接关系到国家政治安全、经济安全和国防安全。为了确保可靠性，当前学术界和工业界普遍采用经过严格数学论证的密码算法。在实际应用中，这些密码算法总会被运行在具体的软硬件设备上，如智能卡、通用处理器、专用电路等。因此，如图 1 所示，在运行过程中，设备往往会泄露许多额外的侧信道信息（如功耗、时间、电磁辐射等），攻击者就可利用此类侧信道泄露来恢复秘密信息。这类针对密码实现的攻击方法称为侧信道攻击（Side-Channel Attack），又称边旁路攻击、边信道攻击等。自 20 世纪 90 年代末起，侧信道攻击给网络空间安全带来了严重威胁，并引起国内外密码、计算机和微电子等领域学者的极大关注，也已成为密码学领域发展最为迅速的方向之一。

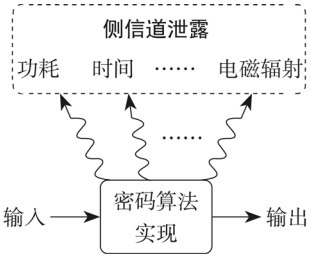

图 1　密码侧信道的安全模型

侧信道攻击主要包括泄露预处理、泄露建模、密钥恢复 3 个过程。下面进行简单的介绍。

1．泄露预处理

在侧信道攻击中，攻击者往往利用各种手段采集到如功耗、电磁等侧信道信息。以每次运行密码算法的时间内采集到的信息为单位，这些信息往往是一系列随着时间而变化的值（除了时间侧信道），可以看成一条横坐标是时间、纵坐标是泄露物理量（如电压、电流或磁场强度等）的曲线，称为泄露曲线轨迹。显然，一条泄露曲线对应密码算法的一次运行。多条泄露曲线组成的集合称为泄露曲线集。

在采集到泄露曲线之后，攻击者就可以开始侧信道攻击过程。而曲线预处理是侧信道攻击过程的第一步。直接采集到的泄露曲线一般存在噪声大、（无用）点数过多、多条曲线之间未对齐等问题，往往很难直接用于密码分析。攻击者/检测人员需要先对曲线进行一系列的预处理操作，如曲线降噪、选择兴趣点（Points Of Interest，POI）、曲线对齐等。

2．泄露建模

侧信道攻击本质上是一种根据侧信道泄露来推测密码算法运行中的秘密信息的过程。这里“运行中的秘密信息”往往指密码算法中的某个中间值。以 AES 加密为例，攻击者往往选择推测第一轮 S 盒输出，或者与它相关的信息（如汉明重量等），这个中间值由明文和密钥直接决定，它的泄露可以直接用于对密钥的恢复。值得一提的是，很多攻击不直接选择密钥本身的泄露，主要是因为密码算法多次运行时，密钥往往是不变的，此时直接通过侧信道泄露来推测密钥的攻击效果往往不够理想。所以，攻击者往往通过推测与密钥相关的一个变化的值从而间接推断密钥值。

侧信道泄露往往是如电压、电流、时间等物理量，而密码算法运行中的秘密信息则是离散的值，如何在侧信道泄露和密码算法运行中的秘密信息之间建立一个联系，是侧信道攻击能否成功的关键因素。泄露模型则是对以上的泄露信息和秘密信息联系的一种刻画，它可以是一个确定或随机函数，它的输入为密码算法运行中的秘密信息（离散值），输出为侧信道泄露的估计。

建立泄露模型的方法主要有两类。一类是利用密码算法电路实现上的一些已知性质，如微处理器总线中功耗往往和总线中传输数据的汉明重量成正比。这种方法需要攻击者/检测人员对密码算法的实现和硬件有所了解，然而以这种方式建立的泄露模型往往也是不准确的。另一类就是利用一个已知密钥的密码设备来建立泄露模型，如果这个已知密钥的设备和被攻击的设备泄露模型相近，这个方法就可以得到一个相对更加准确的泄露模型。

3．密钥恢复

侧信道攻击的最后一步是利用预处理后的曲线、建立好的泄露模型，以及密码算法的输入输出等已知信息来试图恢复密钥。前面提到攻击者往往通过推测与密钥相关的一个变化的值（如 AES 第一轮 S 盒输出）从而间接推断密钥值，这里的间接推断密钥值就是密钥恢复。

这里以典型的差分功耗分析为例来介绍。差分功耗分析利用密码算法的功耗（电磁辐射

往往也是功耗信息的一种形式）信息作为侧信道泄露，相应的泄露曲线也称为功耗曲线，根据在线分析所采用功耗曲线数量的不同，可以把功耗分析分为简单功耗分析和类差分功耗分析。

简单功耗分析只采用一条或几条功耗曲线就可以获得所需要的秘密信息（如密钥等），它对功耗模型的准确程度和功耗曲线的信噪比有比较高的要求。

类差分功耗分析由 Paul Kocher 等人在 1999 年提出的差分功耗分析[1]发展而来，由于对功耗模型的准确程度和功耗曲线的信噪比要求较为宽松，它们的攻击能力往往更为强大。类差分功耗分析需要较多的曲线，通过计算泄露的估计值和真实功耗曲线之间的相关程度来推测秘密信息的值，其中“相关程度”作为一个统计量，它的计算不可避免地需要多条功耗曲线。泄露的估计值由秘密信息（如密钥）、输入（如每条功耗曲线对应的明文或密文）和功耗模型来共同决定，其中输入和功耗模型是攻击者在实施在线攻击之前已经固定的；而秘密信息的值是攻击者所做的一个猜测，攻击者根据相关程度的大小，来判断对应秘密信息的猜测是否正确。

1.2 机器学习简介

机器学习（Machine Learning）是人工智能的一部分，属于计算科学领域，专门分析和解释数据的模式及结构，实现无须人工交互即可完成学习、推理和决策等行为的目的。简单来说，机器学习即支持用户向计算机算法馈送大量数据，然后让计算机分析这些数据，并仅根据输入数据给出建议和决策。机器学习可以分为监督学习、无监督学习以及强化学习。

监督学习（Supervised Learning）又称为有监督学习、监督式学习，是机器学习的一种方法，可以由训练资料中学到或创建一个模式（函数/Learning Model），并依此模式推测新的实例。训练资料由输入对象（通常是很多组数据）和预期输出组成。函数的输出可以是一个连续的值（称为回归分析），或者是预测的一个分类标签（称为分类）。监督学习可以分为以回归分析等统计方法为代表的统计学习和以多层神经网络为基础的深度学习。

与监督学习相对应的是无监督学习（Unsupervised Learning），它没有给定事先标记过的训练实例，自动对输入的资料进行分类或分群。无监督学习包括以 K 均值为代表的聚类方法、以主成分分析为代表的降维方法，以及以 Apriori 算法为代表的关系规则学习（Association Rule）。在人工神经网络中，生成对抗网络（Generative Adversarial Nets）、自组织映射（Self-Organizing Map）和适应性共振理论（Adaptive Resonance Theory）则是最常用的无监督学习。

强化学习是智能体（Agent）以“试错”的方式进行学习，通过与环境进行交互获得的奖赏指导行为，目标是使智能体获得最大的奖赏。强化学习不同于连接主义学习中的监督学习，主要表现在强化信号上。在强化学习中，由环境提供的强化信号是对产生动作的好坏做一种评价（通常为标量信号），而不是告诉强化学习系统如何产生正确的动作。

如图 2 所示，根据密码侧信道攻击的 3 个重要组成部分，本文在第 2~4 节分别阐述机器学习在泄露建模、泄露预处理、在线攻击上的应用。最后在第 5 节介绍一些无法直接分类的研究成果。另外，强化学习目前与侧信道攻击领域的结合主要包括基于强化学习的超参数寻找和基于隐马尔可夫链的攻击方法及其衍生，分别在 2.2.2 节和 4.3 节对它们进行阐述。

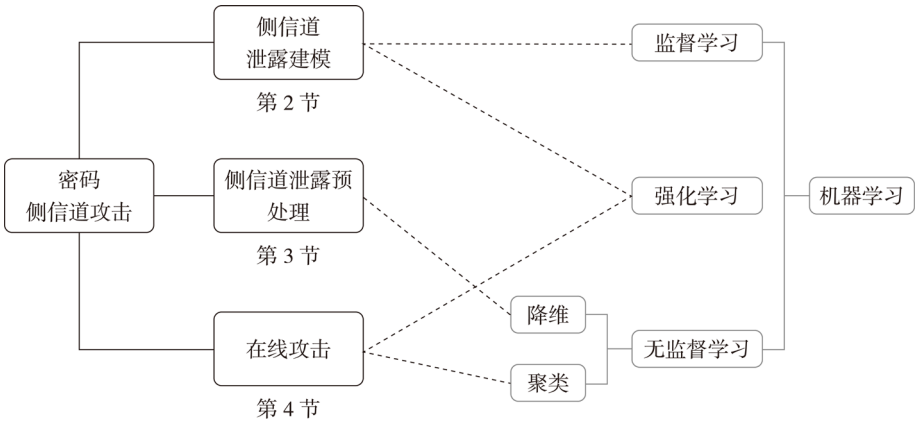

图 2　侧信道与机器学习之间的联系

2　侧信道泄露建模

2.1　基于统计学习的侧信道建模方法

统计学习在侧信道建模中的应用比较早。因为经典侧信道攻击方法就是以统计学为基础的，而统计学习方法是统计学的一个延伸。基于统计学习的侧信道建模和传统的侧信道建模方法区分并不明显，很多方法（如基于线性回归的侧信道建模[2-4]）也逐渐成为经典的建模方法。这类方法的主要特点是结构简单、参数可控、解释性较强、理论完善，它们的主要目的是解决以下两种情况的高效建模问题。

情况 1：建模曲线数量不足。传统的高斯建模方法[5]往往需要多个建模曲线才能得到一个可用的模型，如果建模曲线不够多，在线攻击就会很难成功。另外，随着电子芯片工艺的提升，泄露函数变得越来越复杂，噪声也随之增大，此时为了得到一个有效的功耗模型，传统的建模方法往往需要大量的建模曲线。采用统计学习的方法，即使建模曲线数量比较少，也可以得到一个不太准确的泄露模型，攻击者只要在在线攻击阶段相应地增加攻击曲线数量，就可以成功地进行攻击。也就是说，采用统计学习的方法可以在建模曲线和攻击曲线数量之间达到一个平衡。

情况 2：建模设备和待测设备泄露不匹配。在侧信道攻击中，往往假设建模设备和待测

设备的泄露函数相同，如它们是同一批次的密码设备。然而，随着电子芯片工艺的提升，芯片中的寄生效应越来越明显，导致即便是同一批次的密码设备，其泄露函数也存在较大的差异。传统的建模方法很难在上述情景下发挥有效作用，而统计学习方法则可以更好地适应这种情况，具备更好的鲁棒性。

基于统计学习的侧信道建模方法起始于 2005 年左右，在 2017 年发展达到了高峰，出现了大量的相关理论和方法。下面来回顾这类建模方法的代表性技术和理论。

1．基于线性回归的建模方法

基于线性回归的建模方法是较为成熟的一类方法。Schindler 等人[2]在 2005 年首次提出了基于线性回归的侧信道攻击方法。Whitnall 等人[3]在 2015 年分析了这种方法在曲线数量和建模曲线之间的平衡。Wang 等人[6-7]把岭回归引入了侧信道攻击中，证明了岭回归作为线性回归的一种改良方法，在侧信道攻击中具有更好的效果，并可以很好地处理建模设备和待测设备泄露存在不一致的情况。同时，Wang 等人基于线性回归的可解释性，完善了岭回归、线性回归的适用范围和参数选择方法，并给出了高效的计算方法。Dabosville 等人[8]给出了基于线性回归的二阶攻击方法，可以用于对掩码防护的密码实现进行侧信道攻击。

2．基于分类的建模方法

Whitnall 等人[9]首次提出了基于分类的建模方法，他们使用 k-聚类和层次聚类对不同中间值对应的泄露曲线进行分类，从而可以对不同的中间值进行分类，两个中间值在同一类则说明它们的泄露相同。例如，对于汉明重量泄露的情况下，相同汉明重量的中间值就是一类。这种分类就是一种不那么精确但是具备很好鲁棒性的模型。值得注意的是，这种功耗模型无法使用常见的基于相关性的在线攻击，因为后者需要功耗模型的输出与真实功耗泄露成正比。适用于这种功耗模型的常见在线攻击包括互信息攻击[10]、聚类攻击[11]等。

3．其他统计学习建模方法

Hospodar 等人[12]首次提出了基于支持向量机（Support Vector Machine，SVM）的侧信道攻击方法，同传统的模板攻击进行了比较，他们发现当泄露函数接近汉明重量的情况下，这种方法的效果和使用的参数有比较大的关系。Lerman 等人[13]给出了基于随机森林（Random Forests，RF）的建模方法，并与基于支持向量机的方法以及传统的建模方法进行了比较，发现该方法有较好的效果。

4．建模的方差-偏差平衡理论

基于统计学习的侧信道攻击的可解释性主要来自统计学习的理论方差-偏差的平衡（Variance-Bias Trade-Off）理论。Wang 等人[6-7]和 Lerman 等人[13]分别利用这种理论对相应的建模方法进行了分析。这里做一个简单的介绍。偏差-方差平衡理论是一种来自统计学习领域的理论工具，它主要是为了衡量模型存在的两种泛化误差：偏差和方差。其中偏差是指真实泄露模型和功耗模型之间的偏差；而方差是指功耗模型输出值的方差。建模的目的是同时减

少模型的偏差和方差，但是根据偏差-方差理论（在曲线不足的情况下），偏差和方差是不可能同时达到最小的。例如，利用高方差的建模方法（如线性回归建模）可能会得到一个比较准确的模型，但是相应地，模型的偏差可能会由于过拟合变得较大，导致攻击效果不理想。基于以上理论，既然偏差和方差不可能同时达到最小，那么一个比较好的策略是在模型中人为地加入一些偏差，以达到更好的效果。

2.2　基于神经网络的侧信道建模方法

在侧信道攻击的实践中，由于泄露曲线存在大量噪声，以及密码实现往往加入了侧信道防护措施，使得曲线预处理方法变得极为多样化，攻击者/检测人员不得不花费大量的时间来找到一组合适的曲线预处理和泄露模型建立方法的组合，大大增加了攻击/分析的成本。为了有效地降低侧信道攻击的成本，研究者逐渐思考更加“智能化”、人工干预更加少的分析方法。例如，曲线预处理本质是对曲线中的数据进行操作，从而简化曲线中的点和中间值之间的对应关系。对于未处理的曲线，点和中间值之间的对应关系比较复杂，此时就需要综合能力更强大的深度学习方法。

1．适用于侧信道建模的新型神经网络结构

Cagli 等人[14]把深度学习引入需要预处理过程的侧信道模板攻击。他们把不需要预处理过程的侧信道攻击称为端到端的攻击，提出了一个基于卷积神经网络的建模方法，同时结合在深度学习中常用的数据增强技术，极大地提高了当侧信道泄露曲线存在未对齐情况下的建模效率。值得注意的是，Cagli 等人通过曲线的随机移动等方式对曲线进行微调，以模拟时钟抖动效应（Clock Jitter Effect），从而扩大建模曲线集。而在使用传统的建模攻击方法时，攻击者只有重新对曲线进行对齐的预处理操作，才能建立有效的泄露模型。卷积神经网络可以有效地实施端到端攻击的主要原因在于卷积层的操作对泄露曲线进行适当的组合，使得其可以很好地适应错位的曲线。但是，该工作实际上只考虑到曲线在一定程度上未对齐的情况，而其他诸如随机时钟频率等现实需要预处理的问题没有考虑。

Lu 等人[15]对端到端攻击情况下基于深度学习的建模方法进行了进一步的挖掘。考虑到现实环境中泄露曲线往往是比较长的（含几十万个点），他们使用了称为局部连接（Local Connection）层的新型网络结构，这种结构可以在小的区域内更好地提取有用的信息，并克服卷积层不善于减少整体维度的弊端。同时，作者在实验中发现，使用他们的网络甚至可以获得比使用降维预处理情况下的攻击更好的结果。

Cao 等人[16]考虑在跨设备的场景下，提出了一种基于深度学习的建模方法。他们主要考虑的是建模设备和待测设备的泄露存在一定程度上不一致的情况。为了达到跨设备建模的目的，他们使用了迁移学习领域的一些方法，在建立的预训练模型之后引入了一个额外的微调阶段来调整预训练网络。具体地说，通过对无标签的攻击轨迹数据进行训练，来调整训练所

得模型，使其适应不同设备间的差异性。

Hoang 等人[17]在存在高阶掩码等侧信道防护的情况下，给出了一个新的深度学习架构。他们提出了一种新的多种卷积核组合的模型，而神经网络的结构也被设计成深度较深，但宽度较窄的形式，使得这个架构可以很好地应用于存在多种侧信道防护（如随机延迟、高阶掩码等）的情况。同时，他们发现，在现有深度学习模型上，将明文作为一个特征是有优势的，如果去掉这一特征会大大降低模型在攻击对抗措施中的有效性，他们建议尽可能用明文或密文作为额外的输入特征来提高攻击效率。

一些深度学习应用比较广泛的相关领域中的经典方法也可以直接被用于侧信道攻击的建模中。Kim 等人[18]基于音频领域的模型给出了一个新的卷积网络，它是 VGG 架构的一维形式。这表明，为其他领域开发的卷积神经网络架构（但输入数据与侧信道攻击领域有共同之处）可以成功地应用于侧信道攻击领域。

Yang 等人[19]认为，大多数攻击集中在分析原始曲线上，是在时域进行的，而频域信息则被忽略。一维时域曲线中含有的特征数量不足，同时抖动和错位会带来噪声，导致攻击效率较低。考虑到卷积神经网络能更好地处理二维图像，他们利用短时傅里叶变换，将原始曲线扩展成时间-频率图像，并构建了有效的 2D 卷积网络来有效地学习时频特征。

Masure 等人[20]研究掩码的先验知识对攻击效率的影响。在白盒模型中，攻击者可以获得密码运算过程中的所有信息，包括掩码的值。在黑盒模型中，攻击者只能获得密码运算的输入和输出数据，无法获得掩码的信息，甚至无法了解密码的运算机制。黑盒模型是最常见的攻击场景，但是让神经网络自动学习掩码方案，难度较高。作者提出了一种新的威胁模型——灰盒模型，在这种模型中，攻击者可以获知密码的源代码，因此可以获知所使用的掩码方案，但是在运算过程中，仍然无法获得掩码，攻击者可以通过分析代码在曲线中找出掩码运算的位置。在训练时，针对不同的掩码构建不同的网络，将各个网络输出的概率分布通过离散卷积与掩码运算结合起来，形成最终的概率分布。在这种场景下，攻击速度大大超过了黑盒模型。

大多数场景都是假设攻击者在建模阶段可以获取大量的曲线，而攻击阶段只能获取少量曲线。Picek 等人[21]提出了相反的场景，即在建模阶段受到限制，而攻击阶段的限制较少。他们在各种机器学习模型下，使用建模阶段的少量标记数据和攻击阶段的无标记数据，结合半监督学习领域中的自训练和标签传播算法来进行攻击。但是我们注意到，在有防护的情况下，这种半监督攻击的效果有限。

2．网络的超参数选择

深度网络在训练之前首先需要确定一些网络的超参数，如层数、每层的节点数等，这些参数直接决定了建模的效果。下面简单介绍一些经典的超参数选择的研究工作。

Zaid 等人[22]利用可视化技术，来评估深度学习中模型超参数对建模效果的影响。该工作

的评估方法包括权重可视化和热图两个技术。其中，权重可视化的策略是，如果一个神经元对于分类有较大的影响，就会给这些神经元赋予大的权重；而热图技术的策略是从卷积处的可视化入手来分析卷积层的特征。如果一个位置对分类有较大的影响，就会给这个位置分配较大的权重。同时，根据可视化技术的分析结果，文中提出了一种构建适用于侧信道分析的卷积神经网络架构的方法。

Wouters 等人[23]研究了深度学习应用在建模阶段参数选择的问题。该文章发现 Zaid 等人[22]工作中对参数选择的结论存在一定的问题，主要的原因来自 Zaid 等人的工作中使用的可视化技术。文中指出卷积核尺寸与泄露曲线中的错位量没有严格的对应关系，而增加卷积核的尺寸和卷积核的数量实际上是可以有效提高网络性能的。同时，文中测试了如果先使用传统的预处理方法，再使用深度学习技术进行建模，就能得到不错的结果。这个结论从一定程度上说明了基于深度学习的建模技术虽然可以自动进行一定的预处理（达到端到端的效果），一定程度上减少了测试人员的工作量，但是预处理的效果很难做到比传统的预处理方法好。

其他一些对神经网络参数先验性的研究包括：Li 等人[24]考察了不同权重初始化方法对攻击的影响。在不同数据集下，不同的初始化方法各有优劣，其背后的数学解释仍待探索。同时，权重初始化器和激活函数的不同组合也会对结果有很大影响。Kerkhof 等人[25]对传统损失函数和侧信道领域开发的新型损失函数的性能进行了详细的评估和系统的比较，结果表明，交叉熵比（Cross-Entropy Ratio）的表现在大多数场景下优于其他损失函数，深度学习中的常见损失函数分类交叉熵同样表现良好。Wu 等人[26]探讨了池化层的不同类型和池化步长如何影响卷积神经网络的攻击性能，发现用较大的池化层，可以在保持攻击性能的情况下，训练出较小的网络结构。Perin 等人[27]探讨了不同优化器的优劣，他们发现，Adam 和 RMSprop 容易过拟合，往往需要配合正则化方法；带动量的 SGD 优化器效果较好；Adagrad 可用于大型模型的攻击场景。

利用深度学习神经网络来进行侧信道攻击是近期的研究热点，而神经网络中的超参数需要攻击者/检测人员反复多次进行随机搜索才能确定。这样的搜索时间较长，代价较高。Rijsdijk 等人[28]提出了利用强化学习方法来寻找超参数，形成一个称为 Q-Learning 的超参数寻找算法，并设计了强化学习的奖励函数。作者在多个数据集上进行了实验验证，他们发现，同之前论文中的一些最优模型相比，新的方法参数数量少，且攻击效果更好。值得注意的是，该工作只考虑了卷积神经网络架构下参数的搜索，可以尝试在神经网络（如 Multilayer Perception）下用强化学习进行参数搜索。Knezevic 等人[29]使用进化算法和遗传编程，尝试演化出适合侧信道场景下的激活函数。他们发现，进化出的激活函数一定程度上改善了攻击效果。在未来可以使用进化算法演化其他结构，如损失函数，甚至去进化整个神经网络结构。

在之前的许多工作，大多使用传统的随机搜索或网格搜索来寻找超参数，Wu 等人[30]借鉴了神经架构搜索（Neural Architecture Search，NAS）领域的贝叶斯优化算法，自动进行参

数的调整。贝叶斯优化算法可以工作得很好，但是他们发现随机搜索同样能找到一大批相似性能的模型。由于目前所使用的模型较为轻量，数据集也容易攻破。在哪种侧信道场景下，需要使用 NAS 算法来替代随机搜索或网格搜索，是一个有待研究的问题。

3 侧信道泄露的预处理

3.1 基于统计学习的降维方法

Archambeau 等人[31]首次提出了基于主成分分析的降维方法，并与传统的模板攻击相结合，取得了不错的攻击效果。之后基于主成分分析的降维方法在各类侧信道攻击中得到了极为广泛的应用，逐渐成为侧信道分析领域内事实上的标准。

另外一个比较常用的降维方法是线性判别分析（Linear Discriminant Analysis），Bruneau 等人[32]研究了线性判别分析在侧信道攻击中的应用，他们首先建立了一种所谓最优（但是很难用于实际攻击）的降维方法，这种降维方法保障降维之后的信噪比（Signal-Noise Ratio）是最高的。然后，作者证明了线性判别分析的效果在渐进意义上等价于上述最优的降维方法。即便线性判别分析的效果好于主成分分析，但是由于其效果在噪声较大情况下不太稳定，以及运行时间较长，目前使用最广的还是主成分分析方法。

3.2 数据增强技术

很多利用深度学习的侧信道攻击方法都在使用卷积神经网络的基础上额外增加了数据增强技术，这暗示了数据增强技术可以作为一个独立的方法用于侧信道攻击中。Pu 等人[33]把数据增强技术与基于统计学习的侧信道建模相结合，针对未对齐曲线的情况也得到了很好的效果。这种方法从一定程度上体现了数据增强技术本身的重要性可能超过了卷积神经网络。

Kim 等人[18]也发现在泄露曲线中加入一定量特定的噪声信息，往往可以显著提升建模的效果，同时这项研究指出加入噪声可以看作是一种使用噪声训练的正则化技术。这种增加噪声的方法也可以看成一种数据增强技术，但是与之前工作略有不同的是，这种方法使用了加噪声的方式来随机化曲线，并且在建模阶段不再使用增加噪声前的数据。

Picek 等人[34]考虑不平衡数据对建模过程的影响，提出了新的数据采样技术来提高在出现不平衡数据情况下的建模效果，本质上这种新的数据采样技术也可以看成是数据增强的一种，只是他们更多地考虑如何消除（或减轻）多数带来的偏差。同时，还讨论了实际攻击效果和侧信道常用的判别标准（如成功率或猜测熵）之间可能会存在不一致的情况。文献[34]指出，模型的低准确率可能并不表明在在线攻击过程中使用更多的功耗曲线一定能达到 90%的成功率；反之，也不一定成立。

3.3　基于神经网络的预处理方法

Wu 等人[35]把自动编码器的技术应用在侧信道攻击中，提出使用去噪自动编码器来“智能地”过滤掉一些侧信道防护，证明卷积去噪自动编码器在处理高斯噪声、均匀噪声、去同步化、时钟抖动和洗牌（Shuffling）等侧信道防护的有效性。该方法的前提是攻击者完全控制了一个设备（设备 A），并可以启用/禁用其中的侧信道防护；该方法的攻击目标是启用了侧信道防护的真实设备（设备 B）。该方法如下：攻击者首先从设备 A 获取启用防护和关闭防护的侧信道泄露来建立训练集；然后攻击者使用这些侧信道泄露来训练去噪自动编码器，一旦训练过程结束，经过训练的模型就可以过滤掉设备 B 对应泄露的防护部分。文献也指出，在某些情况下，攻击者也可以不用完全控制设备 A，只要通过已有的预处理方法（如对齐等）成功去除侧信道防护对泄露曲线的影响，从而得到有防护和无防护的泄露曲线即可。但是，我们注意到，这种方法无法有效地对掩码防护进行过滤操作。

Won 等人[36]将多种预处理技术集成到一个网络中。他们提出了一个名为多尺度卷积网络的架构，该网络有多个分支，每个分支接收不同的预处理曲线数据，该网络能够学习每个分支中的不同特征，在公开数据集上取得了较好的效果。

4　密钥恢复攻击方法

4.1　基于分类的在线攻击方法

Batina 等人[11]首次正式提出了基于分类的在线攻击方法，它的主要思想是根据泄露模型和猜测密钥对曲线进行分类，即相同的估计泄露分为一类，其中估计泄露是由猜测密钥计算中间值再使用泄露模型计算得到的。例如，泄露模型是汉明重量，那么只要把中间值的汉明重量相同的曲线分为一类即可。然后，攻击者计算曲线的类内距和类间距，如果猜测密钥是正确的，那么类内距会比较小，而类间距会比较大。在实践中，人们可以使用类内距/类间距作为猜测密钥判断的依据。

值得一提的是，Whitnall 等人[37]发现，如果把上面的方法改为只根据中间值进行分类（而不是根据中间值的泄露估计），就可以得到一种不依赖于任何泄露模型的在线攻击方法。但是，作者也发现，如果（在固定输入的情况下）中间值和假设密钥是一一对应的关系，那么这种攻击一定不会成功。他们进一步也证明了，如果（在固定输入的情况下）目标中间值和假设密钥是一一对应的关系，那么任何不依赖于泄露模型的在线攻击方法都无法有效地进行攻击。当目标中间值和假设密钥是一一对应的关系时，一个泄露模型是必要的。

对分块密码的常规攻击手段是按照分而治之的策略进行的。以 AES 为例，密钥的 8 个比特为一个块。首先训练模型攻击密钥的第一块；然后重新训练模型攻击密钥的第二块，重复

相同的过程，直至恢复所有密钥。Maghrebi 等人[38]使用多标签分类的技术，一次攻击两个密钥块，即 16 比特，而所需的学习时间和使用传统方式攻击一个密钥块所需的时间相当，这提高了攻击效率。同时，Maghrebi 等人尝试用多标签方法一次攻击多个泄露位置，在这种情况下，攻击速度快于传统的方法。Zhang 等人[39]利用多标签分类，将一个字节的密钥拆分成 8 个比特，每份数据使用 8 个标签进行训练，降低了模型的复杂性，在一些场景下好于传统的多类模型。

4.2 基于拟合优度的在线攻击方法

在很多情况下，基于建模的侧信道攻击和非建模的侧信道攻击存在一定的内在联系。具体来说，对建模方法稍加改造，往往可以得到一个非建模的侧信道攻击方法。

这类改造方法基于一个监督式学习的事实：在噪声相同的情况下，学习一个代数结构复杂的函数往往比代数结构简单的函数需要更多的数据。假设目标中间值是$z = f(x,k)$，其中x是输入，k是部分的密钥。令$h(z)$是中间值的泄露，监督学习的目标就是估计函数h。对于一个错误的密钥$k' \neq k$，计算得到的错误中间值为$f(x,k') = f(f^{-1}(z,k),k')$，对应错误的泄露估计为$h(f(f^{-1}(z,k),k'))$，显然这是一个非常复杂的函数。那么，攻击者就可以针对每个假设的密钥k^*，采用监督学习来估计$h(f(f^{-1}(z,k),k^*))$。我们把学习所需的数据量称为拟合优度，那么，假设的密钥k^*为正确密钥时，则拟合优度最低。该思路形成于在文献[37]和文献[4]等给出的从模板到非模板攻击的通用转化方法，已经被广泛用于基于统计学习的侧信道攻击中。

这类方法在基于深度学习的建模方法中也有应用。Timon 等人[40]首次给出了基于深度学习的非建模的侧信道攻击方法，结果显示，在某些情况下效果好于经典的非建模攻击，如相关功率分析等。该方法的主要思路为：每次用一个密钥假设来进行标注，标注完成之后进行训练；若密钥不对，则训练会出问题，如各种训练指标会不正常；若密钥猜对了，训练过程则会比较“顺利”，作者引入了基于拟合优度分析的指标来衡量这种训练“顺利”的程度。另外，作者也给出了针对掩码实现的高阶攻击方法。

4.3 基于强化学习的在线攻击方法

强化学习的目标是给定一个马尔可夫链，寻找最优策略，策略就是状态到动作的映射，使得最终的累计回报最大。Karlof 等人[41]早在 2003 年就把隐马尔可夫模型引入到侧信道攻击中。隐马尔可夫模型是马尔可夫链的一种，它的状态虽然不能直接观察到，但能通过观测向量序列以一定概率观察到，每个观测向量都是通过某些概率密度分布表现为各种状态。他们的主要思想是把密码算法建模成一个有限状态机，然后转化为一个隐马尔可夫链，最终形成一个用于恢复密钥的有效算法。Brumley 等人[42]将隐马尔可夫链和向量量化理论运用到了时间攻击中，实现了针对 ECC 的攻击，对于 160 比特的密钥，破解的平均复杂度由2^{160}降低

到2^{48}。

虽然现在此攻击方法在现实中比较少见，但我们认为很多后续的侧信道攻击方法，包括代数侧信道攻击[43]、柔性侧信道攻击[44]等，都借鉴了以上基于隐马尔可夫链的方法。值得一提的是，在以上攻击方法中，目前最为实用的是柔性侧信道攻击[44]。它的主要思想是把密码算法表示成一个有向图，图中每个节点对应每个中间变量或变量之间的运算。其中，中间变量的节点还附带了这个中间变量所有取值的概率。节点和节点之间由于运算的存在而相互影响。柔性侧信道攻击利用编码学中的信心传递算法（Belief-Propagation Algorithm）把某个变量节点上的侧信道泄露“传递”到图中的密钥变量上，从而得到密钥变量每个取值的概率。与代数侧信道攻击等相似的方法相比，该方法的好处是可以和差分功耗分析一样利用多条功耗曲线进行密钥恢复。值得一提的是，柔性侧信道攻击方法也和编码技术相结合，也被用于侧信道泄露的理论评估中[45]。

5　其他相关研究

5.1　对公钥密码算法的攻击方法

与对称密码算法不同，公钥密码算法运行时间较长，产生的侧信道信息也比较丰富。在这种情况下，攻击者往往可以从一条或几条曲线中较为准确地推测如计算分支的选择、中间比特等信息。针对公钥密码算法的特殊情况，基于深度学习的泄露建模往往可以达到更好的建模效果。下面简单回顾一下这方面比较重要的研究工作。

Perin 等人[46]基于深度学习的方法提出了一种新的方法来改进单次跟踪攻击。在通常情况下，公钥密码算法如 ECC 或 RSA 往往采用“标量盲法”的防护方法来抵抗侧信道攻击，相应地，也存在一类针对这种防护的攻击方法，称为横向攻击。但是当侧信道噪声过大时，由于时间和计算的限制，攻击很难有效地实施。作者采用基于无监督学习的深度学习技术来解决上述问题，实验结果显示，即使恢复的部分私钥包含高达 48%的错误位，神经网络也能返回具有 90%以上正确位的私钥。

Carbone 等人[47]基于 RSA 算法的深度学习模板攻击，他们对带有消息盲化、指数盲化、模数盲化防护的 RSA 实现进行攻击，利用电磁辐射泄露，并把横向攻击方法和深度学习相结合，有效地恢复了一些秘密信息。此外，该工作还研究了当建模设备和攻击设备不一致时这种攻击的效果，并发现设备的不一致不会对攻击的有效性产生较大影响。

5.2　对机器学习方法的改进

侧信道攻击的场景往往和很多机器学习的应用场景（如图像识别、声音识别等）是大相

径庭的，直接把机器学习的方法（尤其是基于深度网络的机器学习方法）用于侧信道攻击往往是不合理的。这就反向推动了对机器学习方法的改进。下面简单介绍一些基于侧信道攻击场景下对机器学习方法的改进。

Zhang 等人[48]发现在评价训练/学习到的模型的优劣程度时，深度学习和侧信道领域的指标存在很大差别，尤其是对于不平衡数据的情况，这大大降低了机器学习的效率。文献提出了一个新的指标称为交叉熵比，来评估侧信道攻击的深度学习模型的性能，并证明了交叉熵比与传统的侧信道指标猜测熵和成功率密切相关。同时，该工作将交叉熵比指标调整为一种损失函数，解决了评价深度学习和侧信道领域的指标不一致的问题，显著提高了建模的效率。

Perin 等人[49]分析了神经网络输出类概率及其与成功密钥恢复的关系。输出类概率是机器学习模型泛化能力的评估标准。文献探讨了神经网络输出层分类概率所包含的信息，以此作为验证侧信道攻击训练的神经网络性能的有效信息。同时，文中提出用组合方式来提高泛化能力。这种方法对几个训练好的神经网络模型进行组合，而不是简单地从搜索举例超参数中选择最佳模型。由于在集合中，几个模型被结合在一起，如果少数模型表现不佳，这些模型引入的波动将通过拥有泛化的模型而被消除，并且组合可以明显获得更好的攻击性能。

Masure 等人[50]证明了在训练深度神经网络时负对数似然损失（Negative Log Likelihood Loss，NLL Loss）实际上渐进地等同于最大化感知信息（Perceived Information，PI），而 PI 可以作为泄露和目标密钥之间互信息的下限。这个结论证明了利用深度学习建立模型在理论上的有效性和高效性，论证了深度学习用于侧信道攻击的合理性，也说明了利用深度神经网络可以很好地对 PI 进行估计，从而在理论上评估某类实现的泄露程度。同时，该工作也通过仿真和实验说明即便是在有高阶掩码等侧信道防护的情况下，通过 NLL 来估计 PI 也是非常有效的。

Zaid 等人[51]采用机器学习中集成模型的建模方式，提出了一种新的损失函数融合损失（Ensembling Loss，EI），该损失函数提高了集成模型中成员函数之间的多样性，从而帮助模型提升了准确性。另外，该工作结合丰富的实验和可视化结果，从实验方面证明了 EI 确实能够提高成员函数间的多样性。相比于其他基于深度学习的侧信道攻击方案该工作有更好的性能表现，可以生成有效的集成模型应用于侧信道攻击。

如果训练时间太短，网络就无法获得其全部能力；而如果训练时间太长，网络就会过拟合。找到训练停止的正确时机对于侧信道分析来说比较困难，因为深度学习的指标和侧信道指标之间没有明确的联系。深度学习指标无法作为侧信道中模型的监测指标。对模型进行完整的评估需要在不同的训练阶段进行一次完整的攻击，得到猜测熵。这会带来较大的运算量。Robissout 等人[52]考虑使用成功率，在训练的各个阶段计算训练集和验证集上成功率的差异，以此作为性能指标。但是计算成功率并不稳定，且计算成本同样较高，达到模型最佳状态的时间窗口太短。Perin 等人[53]借鉴了信息瓶颈理论，使用互信息作为早期停止策略的监控指

标。他们的具体做法是将深度学习模型的每层视为随机变量，计算各个隐藏层与输入输出标签的互信息。在训练阶段，输入层和隐藏层的互信息快速增加。而在过拟合阶段，神经网络开始压缩输入数据的信息，其互信息则会减少。他们的方法的缺点是精确计算互信息需要较高的成本。

5.3　基于深度学习的泄露检测

Moos 等人[54]首次探讨了深度学习是否可以应用在泄露检测方法中，并给出了第一个基于深度学习的全面泄露检测方法。泄露检测是指一类跳过复杂的侧信道攻击流程，直接分析密码实现是否存在侧信道泄露的技术。典型的（非机器学习方法的）泄露检测方法首先采集两组不同输入的侧信道功耗曲线（同一组曲线的输入相同），然后利用 T-Test 等统计学方法[55-56]检测两组曲线是否在统计上同分布，若存在泄露，两组曲线则有明显区别。在区分两组功耗曲线的基础上，Moos 等人引入了监督学习和敏感性分析，形成了一种基于深度学习的泄露检测方法：如果存在泄露，那么利用训练集训练神经网络，可以有效地区分评估验证集中的两组功耗曲线。通过实验，与传统方法相比，采用了多层感知器和卷积神经网络两种模型（在不同数据集上兼容性较好）能从更小的数据集中检测到泄露，而且在数据集相同时达到更高的检测精度。

5.4　用于研究的数据集

许多文献中的攻击是作者用自己采集的数据进行评估的，缺少统一的标准。Prouff 等人[57]发布了名为 ASCAD 的公共数据集。它包含 3 组轨迹和相关的元数据，每组轨迹都有不同程度的抖动，均为一阶掩码 AES 的软件实现。该数据集质量较高，攻击难度适中，目前已被侧信道攻击研究人员广泛使用，成为评估攻击效果的基准之一。

参考文献

[1] KOCHER R, JAFFE J, JUN B. Differential power analysis[C/OL]. In: Annual international cryptology conference: 1999. Springer: 388-397. https://doi.org/10.1007/3-540-48405-1_25.

[2] SCHINDLER W, LEMKE K, PAAR C. A stochastic model for differential side channel cryptanalysis[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2005. Springer: 30-46. https://doi.org/10.1007/11545262_3.

[3] WHITNALL C, OSWALD E. Profiling DPA: Efficacy and Efficiency Trade-Offs [C/OL]. In: International Conference on Cryptographic Hardware and Embedded Systems: 2013. Springer: 37-54. https://doi.org/10.1007/978-3-642-40349-1_3.

[4] WANG W, YU Y, LIU J, et al. Evaluation and improvement of generic-emulating DPA attacks[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2015. Springer: 416-432. https://doi.org/10.1007/978-3-662-48324-4_21.

[5] CHARI S, RAO JR, ROHATGI P. Template attacks[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2002. Springer: 13-28. https://doi.org/10.1007/3-540-36400-5_3.

[6] WANG W, YU Y, STANDAERT F-X, et al. Ridge-Based profiled differential power analysis[C/OL]. In: Cryptographers' Track at the RSA Conference: 2017. Springer: 347-362. https://doi.org/10.1007/978-3-319-52153-4_20.

[7] WANG W, YU Y, STANDAERT F-X, et al. Ridge-Based DPA: Improvement of Differential Power Analysis for Nanoscale Chips[J/OL]. IEEE Trans Inf Forensics Secure 2018, 13(5):1301-1316. https://doi.org/10.1109/TIFS.2017.2787985.

[8] DABOSVILLE G, DOGET J, PROUFF E. A new second-order side channel attack based on linear regression[J/OL].IEEE Transactions on Computers 2012,62(8):1629-1640. https://doi.org/10.1109/ TC.2012.112.

[9] WHITNALL C, OSWALD E. Robust profiling for DPA-style attacks[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2015. Springer: 3-21. https://doi.org/10.1007/978-3-662-48324-4_1.

[10] GIERLICHS B, BATINA L, TUYLS P, et al. Mutual information analysis[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2008. Springer: 426-442. https://doi.org/10.1007/978-3-540-85053-3_27.

[11] BATINA L, GIERLICHS B, LEMKE-RUST K. Differential cluster analysis[C/OL]. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2009. Springer: 112-127. https://doi.org/10.1007/978-3-642-04138-9_9.

[12] HOSPODAR G, GIERLICHS B, DE MULDER E, et al. Machine learning in side-channel analysis: a first study[J/OL]. Journal of Cryptographic Engineering 2011, 1(4):293-302. https://doi.org/ 10.1007/s13389-011-0023-x.

[13] LERMAN L, POUSSIER R, MARKOWITCH O, et al. Template attacks versus machine learning revisited and the curse of dimensionality in side-channel analysis[J/OL]: extended version. Journal of Cryptographic Engineering 2018, 8(4):301-313. http://dx.doi.org/10.1007/s13389-017-0162-9.

[14] CAGLI E, DUMAS C, PROUFF E. Convolutional neural networks with data augmentation against jitter-based countermeasures[C/OL]. In: International Conference on Crypt

ographic Hardware and Embedded Systems: 2017. Springer:45-68. http://dx.doi.org/10.1007/978-3-319-66787-4_3.

[15] LU X, ZHANG C, CAO P, et al. Pay attention to raw traces: A deep learning architecture for end-to-end profiling attacks[J/OL]. IACR Transactions on Cryptographic Hardware Embedded Systems 2021: 235-274.https://tches.iacr.org/index.php/TCHES/article/view/8974.http://dx.doi.org/10.46586/tches.v2021.i3.235-274.

[16] CAO P, ZHANG C, LU X, et al. Cross-Device Profiled Side-Channel Attack with Unsupervised Domain Adaptation[J/OL]. IACR Transactions on Cryptographic Hardware Embedded Systems 2021:27-56. http://dx.doi.org/10.46586/tches.v2021.i4.27-56.

[17] HOANG A-T, HANLEY N, O'NEIL M. Plaintext: A missing feature for enhancing the power of deep learning in side-channel analysis? Breaking multiple layers of side-channel countermeasures[J/OL]. IACR Transaction on Cryptographic Hardware Embedded Systems 2020:49-85. http://dx.doi.org/10.46586/tches.v2020.i4.49-85.

[18] KIM J, PIECK S, HEUSER A, et al. Make some noise. Unleashing the power of convolutional neural networks for profiles side-channel analysis[J/OL]. IACR Transactions on Cryptographic Hardware 2019:148-179. http://dx.doi.org/10.46586/tches.v2019.i3.148-179.

[19] YANG G, LI H, MING J, et al. Convolutional neural network based side-channel attacks in time-frequency representations[C/OL]. In: International Conference on Smart Card Research and Advanced Applications: 2018. Springer:1-17. http://dx.doi.org/10.1007/978-3-030-15462-2_1.

[20] MASURE L, CRISTIANI V, LECOMTE M, et al. Don't Learn What You Already Know: Scheme-Aw are Modeling for Profiling Side-Channel Analysis against Masking[A/OL]. Cryptology ePrint Archive: 2022. http://dx.doi.org/10.46586/tches.v2023.i1.32-59.

[21] PICEK S, HEUSER A, JOVIC A, et al. Improving side-channel analysis through semi-supervised learning[C/OL]. In: International Conference on Smart Card Research and Advanced Applications: 2018. Springer:35-50. http://dx.doi.org/10.1007/978-3-030-15462-2_3.

[22] ZAID G, BOSSUET L, HABRARD A, et al. Methodology for efficient CNN architectures in profiling attacks[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020(1):1-36. http://dx.doi.org/10.46586/tches.v2020.i1.1-36.

[23] WOUTERS L, ARRIBAS V, GIERLICHS B, et al. Revisiting a methodology for efficient CNN architectures in profiling attacks[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems 2020:147-168. http://dx.doi.org/10.46586/tches.v2020.i3.147-168.

[24] LI H, KRČEK M, PERIN G. A comparison of weight initializers in deep learning-bases side-channel analysis[C/OL]. In: International Conference on Applied Cryptography and Network

Security: 2020. Springer:126-143. https://doi.org/10.1007/978-3-030-61638-0_8.

[25] KERKHOF M, WU L, PERIN G, et al. No (Good) Loss no Gain: Systematic Evaluation of Loss functions in Deep Learning-based Side-channel Analysis[A]. Cryptology ePrint Archive, 2021.

[26] WU L, PERIN G. On the importance of pooling layer tuning for profiling side-channel analysis[C/OL]. In: International Conference on Applied Cryptography and Network Security:2021. Springer:114-132. https://doi.org/10.1007/978-3-030-81645-2_8.

[27] PERIN G, PICEK S. On the Influence of Optimizers in Deep Learning-based Side-channel Analysis[C/OL]. In: International Conference on Selected Areas inCryptography: 2020. Springer:615-636. https://doi.org/10.1007/978-3-030-81652-0_24.

[28] RIJSDIJK J, WU L, PERIN G, et al. Reinforcement Learning for Hyperparamete r Tuning in Deep Learning-based Side-channel Analysis. IACR Transactions on Cryptograph ic Hardware and Embedded Systems (2021)[J/OL]:677-707. https://doi.org/10.46586/ tches.v2 021.i3.677-707.

[29] KNEZEVIC K, JURAJ F, JAKOBOVIC D, et al. NeuroSCA: Evolving Activation Functions for Side-Channel Analysis. IEEE Access (2022)[J/OL].https://doi.org/10.1109/ACCESS. 2022.3232064.

[30] WU L, PERIN G, PICEK S. I Choose You: Automated Hyperparameter Tuning for Deep Learning-based Side-channel Analysis. IEEE Transactions on Emerging Topics in Computing[J/OL]. https://doi.org/10.1109/TETC.2022.3218372.

[31] ARCHAMBEAU C, PEETERS E, STANDAERT F-X, et al. Template Attacks in Principal Subspaces. In: International Workshop on Cryptographic Hardware and Embedded Systems: 2006. Springer: 1-14[J/OL]. https://doi.org/10.1007/11894063_1.

[32] BRUNEAU N, GUILLEY S, HEUSER A, et al. Less is More[C/OL]. In: Interna tional Workshop on Cryptographic Hardware and Embedded Systems: 2015. Springer: 22-41. https://doi.org/10.1007/ 978-3-662-48324-4_2.

[33] PU S, YU Y, WANG W. Trace augmentation: What can be done even before preprocessing in a profiled sca? [C/OL] In: International Conference on Smart Card Research and Advanced Applications: 2017.Springer:232-247. https://doi.org/10.1007/978-3-319-75208-2_14.

[34] PICEK S, HEUSER A, JOVIC A, et al. The Curse of Class Imbalance and Conflicting Metrics with Machine Learning for Side-channel Evaluations[J/OL]. IACR Transactions on Cryptographic Hardware 2019,2019(1):1-29. https://doi.org/10.13154/tches.v2019.i1.209-237.

[35] WU L, PICEK S. Remove some noise: On pre-processing of side-channel measurements

with autoencoders[J/OL]. IACR Transactions on Cryptographic Hardware Embedded Systems 2020:389-415. https://doi.org/10.46586/TCHES.V2020.I4.389-415.

[36] WON Y-S, HOU X, JAP D, et al. Back to the Basics: Seamless Integration of Side-Channel Pre-Processing in Deep Neural Networks. IEEE Transactions on Information Forensics and Security 16 (2021):3215-3227. https://doi.org/10.1109/TIFS.2021.3076928.

[37] WHITNALL C, OSWALD E, STANDAERT FX. The Myth of Generic DPA…and the Magic of Learning: The Cryptographer's Track at the RSA Conference 2014 , San Francisco, CA, USA, February 25-28, 2014 [C/OL]. 2014: 183-205. https://doi.org/10.1007/978-3-319-04852-9_10.

[38] MAGHREBI H. Deep Learning based Side-Channel Attack: a New Profiling Methodology based on Multi-Label Classification. Cryptology ePrint Archive[A/OL]. 2020. https://eprint.iacr.org/2020/436.

[39] ZHANG L B, XING X, FAN J F, et al. Multilabel Deep Learning-Based Side-Channel Attack[J/OL]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2021: 1207-1216. 10.1109/TCAD.2020.3033495.

[40] TIMON B. Non-Profiled Deep Learning-based Side-Channel attacks with Sensitivity Analysis[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems. 2019: 107-131. https://doi.org/10.13154/tches.v2019.i2.107-131.

[41] KARLOF C, WAGNER D. Hidden Markov Model Cryptanalysis: Cryptographic Hardware and Embedded Systems, Cologne, Germany, September 8-10, 2003[C/OL]. 2003: 17-34. https://doi.org/ 10.1007/978-3-540-45238-6_3.

[42] BRUMLEY D, BONEH D. Remote timing attacks are practical [J/OL]. Computer Networks, 2005,48(5): 701-716.

[43] RENAULD M, STANDAERT FX. Algebraic Side-Channel Attacks: Information Se curity and Cryptology, Beijing, China, December 12-15, 2009[C/OL]. 2009: 393-410. https://doi.org/10.1007/ 978-3-642-16342-5_29.

[44] VEYRAT-CHARVILLON N, GÉRARD B, STANDAERT FX. Soft Analytical Side-Channel Attacks: 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, China, December 7-11, 2014[C/OL]. https://doi.org/10.1007/ 978-3-662-45611-8_15.

[45]GUO Q, GROSSO V, STANDAERT F-X, et al. Modeling soft analytical side-channel attacks from a coding theory viewpoint[J]. IACR Trans. Cryptogr. Hardw. Embed. Syst., 2020(4): 209-238.

[46]PERIN G, CHMIELEWSKI Ł. BATINA L. Keep it unsupervised: Horizontal attacks meet

deep learning[C]. IACR Transactions on Cryptographic Hardware Embedded Systems 2021:343-372.

[47]CARBONE M, CONIN V, CORNELIE MA, et al. Deep learning to evaluate secure RSA implementations[J]. IACR Trans. Cryptogr. Hardw. Embed. Syst., 2019(2):132-161.

[48]ZHANG J, ZHENG M, NAN J, et al. A Novel Evaluation Metric for Deep Learning-Based Side Channel Analysis and Its Extended Application to Imbalanced Data[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020:73-96. https://doi.org/10.13154/tches.v2020.i3.73-96.

[49]PERIN G, CHMIELEWSKI Ł, PICEK S. Strength in numbers: Improving generalization with ensembles in machine learning-based profiled side-channel analysis[J]. IACR Transactions on Cryptographic Hardware Embedded Systems 2020:337-364.

[50]MASURE L, DUMAS C, PROUFF E. A Comprehensive Study of Deep Learning for Side-Channel Analysis[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020:348-375. https://doi.org/10.13154/tches.v2020.i1.348-375.

[51]ZAID G, BOSSUET L, HABRARD A, et al. Efficiency through Diversity in Ensemble Models applied to Side-Channel Attacks: A Case Study on Public-Key Algorithms[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2021:60-96. https://doi.org/10.46586/tches.v2021.i3.60-96.

[52]ROBISSOUT D, ZAID G, COLOMBIER B, et al. Online Performance Evaluation of Deep Learning Networks for Side-Channel Analysis[C/OL]. In: International workshop on Constructive Side-Channel Analysis and Secure Design: 2020, Springer:200-218. https://doi.org/10.1007/978-3-030-68773-1_10.

[53]PERIN G, BUHAN I, PICEK S. Learning when to stop: a mutual information approach to fight overfitting in profiled side-channel analysis[C/OL]. In: International workshop on Constructive Side-Channel Analysis and Secure Design: 2021, Springer:53-81. https://doi.org/10.1007/978-3-030-89915-8_3.

[54]MOOS T, WEGENER F, MORADI A. DL-LA: Deep Learning Leakage Assessment: A modern roadmap for SCA evaluations[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2021:552-598. https://doi.org/10.46586/tches.v2021.i3.552-598.

[55]WHITNALL C, OSWALD E. A Critical Analysis of ISO 17825(Testing Methods for the Mitigation of Non-invasive Attack Classes Against Cryptographic Modules)[J/OL]. In: International Conference on the Theory and Application of Cryptology and Information Security:2019, Springer:256-284. https://doi.org/10.1007/978-3-030-34618-8_9.

[56]MORADI A, RICHTER B, SCHNEIDER T, et al. Leakage Detection with the x2-

Test[J/OL]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2018:209-237. https://doi.org/ 10.13154/tches.v2018.i1.209-237.

[57]Ryad B, Prouff E, Strullu R, et al. Deep learning for side-channel analysis and introduction to ASCAD database. Journal of Cryptographic Engineering 10, 2020:163-188.https://doi.org/ 10.1007/s13389-019-00220-8.

基于密码技术的横向联邦学习隐私保护协议

田海博[1] 李茂楠[1] 邵云峰[2] 李秉帅[2]

1.中山大学 计算机学院，广州，510275

2.华为技术有限公司 诺亚方舟实验室，北京，100085

通讯作者：田海博，E-mail：tianhb@mail.sysu.edu.cn

摘 要：McMahan 等人提出横向联邦平均学习算法以来，人们对联邦学习隐私保护技术的研究不断深入。在一开始，鉴于联邦平均算法中数据不出本地的特性，人们认为该算法本身就具有良好的隐私保护特性。然而，很快就出现了针对联邦学习中梯度信息的攻击算法，可以高概率地还原联邦学习算法中输入的数据。Bonawitz 等人把密码技术引入到联邦学习中，采用简单的密钥交换、秘密分享、混合加密、伪随机函数等密码学工具，设计了安全的聚合协议，可以完成对百万级模型参数的保护。随后，针对该协议的通信和计算代价，研究者进一步给出了采用随机图的安全聚合协议和采用同态伪随机函数的安全聚合协议。本文对联邦平均算法、安全聚合协议及其改进协议进行介绍，探讨协议设计背后的原理。

关键词：横向联邦学习；隐私保护；安全聚合协议

Protocols for Privacy Protection of Horizontal Federated Learning Based on Cryptographic Techniques

TIAN Haibo[1] LI Maonan[1], SHAO Yunfeng[2], LI Bingshuai[2]

1. School of Computer Science and Engineering, Sun Yat-Sen Univeristy, Guanzhou, 510275

2. Huawei Noah’s Ark Lab, Beijing, 100085

Corresponding author: TIAN Haibo, E-mail：tianhb@mail.sysu.edu.cn

Abstract: McMahan et al. propose a horizontal federated average learning algorithm. People deepen the research of privacy protection technologies for their algorithm. At the beginning, in view of the fact that the data in the federal average algorithm are local, people think that the algorithm

itself has good privacy protection characteristics. However, attack algorithms against gradient information in federated learning soon appear, which can restore the input data of the federated learning algorithm with high probability. Bonawitz et al. introduce cryptographic technologies into federated learning, and design secure aggregation protocols using simple cryptographic tools including key exchange, secret sharing, hybrid encryption, pseudo-random functions, which can protect millions of model parameters. Then, to reduce the communication and computing costs of the protocol, researchers further propose secure aggregation protocols using random graphs and secure aggregation protocols using homomorphic pseudo-random functions. This paper introduces the federated averaging algorithm, security aggregation protocols and their improved protocols, and discusses the principles behind their design.

Keywords: Horizontal Federated Learning; Privacy Protection; Secure Aggregation Protocols

1　引言

近几年来，机器学习[1]在诸如智能推荐、图像识别、智能驾驶等许多应用场景中得到广泛应用，数据隐私的问题随之受到关注。首先，大批量优质隐私数据可以训练出在医疗、金融等敏感领域非常有用的机器学习模型，提高人们的健康水平、保护资金安全。其次，《中华人民共和国个人信息保护法》明确规定了信息处理者对信息负有保护义务，因而企事业单位涉及个人隐私的数据受到法律保护，不能任意分享使用。最后，大批量数据可能分布在不同的企事业单位中，敏感数据对于企事业单位而言是其自身的重要资产，用于机器学习自然特别慎重。总之，数据隐私的问题使得机器学习获得大批量优质敏感数据是困难的，阻碍了机器学习技术在敏感领域的广泛应用。

联邦学习[2-3]是为了解决上述问题提出的一种机器学习方法。按照杨强等人[3]的分类，联邦学习可以根据训练数据的分布特征分为横向、纵向和迁移联邦学习。这些联邦学习的共同特点之一是在本地进行模型训练，以在数据不离开数据拥有方设备的基础上，实现跨主体、跨设备的机器学习，避免数据集中的训练模式下明显的隐私问题。利用联邦学习，Google 公司进行了大规模的移动终端键盘输入软件（Gboard APP）的训练[4]，在安卓设备定位为北美和加拿大的用户输入数据上，训练了用户输入的预测模型，以提高用户的输入体验。Lee 等人[5]实现了在不暴露患者信息的前提下，在数据不互通的多家医疗机构间快速找寻相似病情的患者。郭睿等人[6]联合多方关于旅客的消费记录以及出行信息，设计了航空出行的推荐方法。

然而，在本地进行数据训练并不意味着本地数据的安全。以横向联邦学习为例，参与训练的各方需要与服务器交换模型参数，以更新待训练的模型。而模型参数是可能泄露训练数

据的。Zhu 等人[7]发现当横向联邦学习泄露单个参与方模型的梯度信息时，可以构建网络模型拟合泄露的梯度信息，重建训练数据。对于图像数据而言，Zhu 等人恢复的数据是像素级别的。Zhao 等人[8]进一步展示了如何提取真实标签。Jonas 等人[9]展示了在图像上进行几次迭代后得到的平均梯度依旧泄露原始训练数据。Yin 等人[10]展示了如何从平均梯度中恢复批次图像，他们的算法中一个批次含有 8~48 幅图像，采用的是 ResNets 网络和 ImageNet 数据库。这些观察意味着联邦学习本身依旧有数据隐私的问题，需要慎重对待。

2 基于密码的横向联邦学习隐私保护技术

针对横向联邦学习的数据隐私问题，Abadi 等人[11]在 2016 年提出了采用差分隐私（DP）的思路，在参与方提交给服务器模型参数之前，先增加一个特定方差的高斯噪声，以在不特别损害训练模型精度的情况下，保护单个参与方的隐私。Fan 等人[12]在 2021 年提出了采用可信计算环境（TEE）的思路，参与方在对服务器的可信环境认证后，向服务器的可信环境提交模型参数，由服务器在可信环境下完成聚合。

DP 和 TEE 对密码技术的依赖性较小，在效率方面具有优势，然而 DP 方法需要在精度和隐私保护之间平衡，TEE 方法需要信赖硬件安全环境，都有各自难以克服的缺点。基于密码技术的横向联邦学习隐私保护技术在这两个方面没有明显缺陷，在效率方面则需要精巧的设计来提升。

对于横向联邦学习而言，最显而易见适用的密码技术是同态加密。如果同态加密的解密密钥只有客户端知道，且客户端是诚实的，那么横向联邦学习的隐私保护可以简单到只需要一轮通信，客户端加密模型参数给服务器，服务器同态聚合给客户端，客户端本地解密后继续训练。早期的一些横向联邦学习隐私保护方案或一些特定的应用场景中采用了这种密钥假设[13-18]，这些方案在密钥的生成方式和采用的同态加密方案上有所不同。如果同态加密的解密密钥是一个可信第三方持有的，那么横向联邦学习的隐私保护也可以简单到只需要一轮通信，服务器只需要在聚合后给可信第三方，由可信第三方解密一次即完成一轮聚合。在一些特定的应用场景中，确实也有学者提出了这样一些方案[19-20]。最后如果同态加密的解密密钥是由客户端分布持有的，且允许攻击者腐化服务器和部分客户端，横向联邦学习的隐私保护就要复杂一些了。Truex 等人[21]给出了采用门限 Paillier 同态加密[22]的联邦学习隐私保护方法。其中，门限 Paillier 同态加密[22]中建议的密钥生成方式是需要一个可信的密钥分发中心的。文献[23-26]给出的横向联邦学习隐私保护技术也是基于该方案完成的。Jiang 等人[27]给出的横向联邦学习隐私保护方案中采用了 Bresson 等人的方案[28]，本质上也是需要一个可信密钥生成中心的。文献[29-30]分别采用了门限的 ElGamal 加密和门限的格基加密方案来完成横向联邦学习的隐私保护。最后这两个方案虽然不需要可信密钥分发中心，但是通信和计算的代价

依旧比较高。

安全多方计算技术作为一种普适性的隐私保护技术当然也可以用在横向联邦学习隐私保护的场景中。文献[31-32]采用了对模型参数做秘密分享的思路，每个参与方将本地训练的模型参数按照份额分享给其他参与方，各自做聚合。为了降低秘密分享方案的通信量，Kadhe等人[33]采用了多秘密分享的技术。文献[34-35]提议首先在参与者中选举一个委员会，然后在委员会成员间进行安全多方计算。文献[36-37]直接采用了两个或多个服务器完成横向联邦学习的安全多方计算任务。

安全聚合协议是一种以隐私保持的求和为主要目标的协议。早期该协议用来对时间序列数据进行加和[38]，采用了同态加密和秘密分享技术。Bonawitz 等人[39]将安全聚合这一概念用于横向联邦学习隐私保护，主要采用了伪随机函数和 Diffie-Hellman 密钥协商协议的特性，用于高维数据的隐私保持聚合。该协议没有涉及复杂的密码操作，巧妙地组合了系列的基本密码技术，得到了一个可以有效处理上百万维数据聚合的协议。该协议自 2017 年提出后，受到了广泛关注，人们提出了各种思路来继续提升协议的效率。

第一个思路就是压缩模型参数的大小和规模，以减小安全聚合协议的输入规模。这里主要的优化体现为对模型参数的量化方法和实际上传的模型参数的裁剪[40-42]。第二个思路体现为缩小安全聚合协议参与方的规模，通过分组的方式提高效率。文献[43-44]采用了预分组的方式，文献[45-46]采用了随机分组的方式。第三个思路体现为降低协议中组成要素的复杂度，如文献[47-48]通过引入可信第三方来提高生成伪随机函数种子的效率，Liu 等人[49]采用同态伪随机函数的方法去掉了 Diffie-Hellman 密钥协商协议，仅依赖秘密分享完成了安全聚合，效率提升明显。

此外，人们还对联邦学习中安全聚合协议在多轮情况下的安全性[50]以及对全局模型参数的保护上[51-52]进行了深入研究，取得了一些研究成果。

下面主要介绍文献[2,39,46,49]的研究内容，包括横向联邦平均算法[2]、基本和增强的安全聚合协议[39]、基于同态伪随机函数的安全聚合协议[49]、基本和增强的随机图安全聚合协议[46]，探讨这些协议设计的特点和背后的原理。

3　联邦平均算法

横向联邦学习的基础算法之一是联邦平均算法[2]，本文中的安全聚合协议都是基于该算法进行的工作。表 1 给出了该算法的详细描述。

表 1　联邦平均算法

服务器执行：
初始化ω_0
对第$t = 1,2,\cdots$轮，执行
设置$S_t =$ (随机选择的 $\max(C \cdot K, 1)$ 个客户端)
让每个客户端$k \in S_t$并行执行
$\omega_t^k \leftarrow \text{ClientUpdate}(k, \omega_{t-1})$
根据客户端返回值计算 $\omega_t \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} \omega_t^k$
ClientUpdate(k, ω)：//在客户端k上执行的算法
对一个世代$i \in [1, E]$，执行
将本地数据P_k分为大小为B的批量数据
对批量数据中的每批数据b，执行
$\omega \leftarrow \omega - \eta \nabla \ell(\omega; b)$
返回ω给服务器

在表 1 中，ω表示联邦学习模型的参数，其下标为 0 时，表示初始化参数，其他情况表示相应轮数时模型的参数；t表示训练的轮数，该算法没有明确写出训练终止的条件，因此只是表示训练轮数渐次增长，实际训练时可以通过采用固定轮数或其他基于测试数据的方式来终止训练；C 表示每轮初始参与训练的客户端的比例；K表示总的客户端的数量；S_t集合表示第t轮初始参与训练的客户端。

每个客户端并行执行ClientUpdate函数，该函数的输入包括客户端标识k和上一轮的模型参数ω_{t-1}；当客户端返回客户端本地更新的模型参数ω_t^k后，服务器计算加权平均，其中n表示所有客户端的训练数据总量，n_k表示客户端k所拥有的训练数据的量，$\frac{n_k}{n}$为权重，因为实际参与训练的客户端是所有客户端的子集，因此当k从 1 到K求和时，有些不参与训练的客户端其参数默认为 0。

客户端运行ClientUpdate函数，每次运行会用所有的客户端数据P_k进行训练，训练的次数为E，每次训练称为一个世代；对于每个世代的训练，都是把本地数据P_k分批训练，批量数据的规模为 B，本地模型参数的更新是以批次数据b为单位更新的，即在输入批次b的数据后，形成模型参数的梯度值$\nabla \ell(\omega; b)$，乘上学习率η之后，得到模型参数应该修正的值。

图 1 给出了联邦平均算法的通信模型。其中服务器执行部分用云图来表示，客户端执行部分用数据集和模型两个模块来表示。数据集分为K份，表示为$P_1, \cdots, P_K$，对应可以训练K个本地模型，表示为$\omega^1, \cdots, \omega^K$。每个本地模型训练前的输入都是全局模型参数。本地模型通过

执行E个世代的训练完成更新，更新后的本地模型参数返回给服务器。服务器在收到本轮参与训练的客户端返回的本地模型参数后，计算其加权平均，更新全局模型参数，然后进行下一轮的训练。

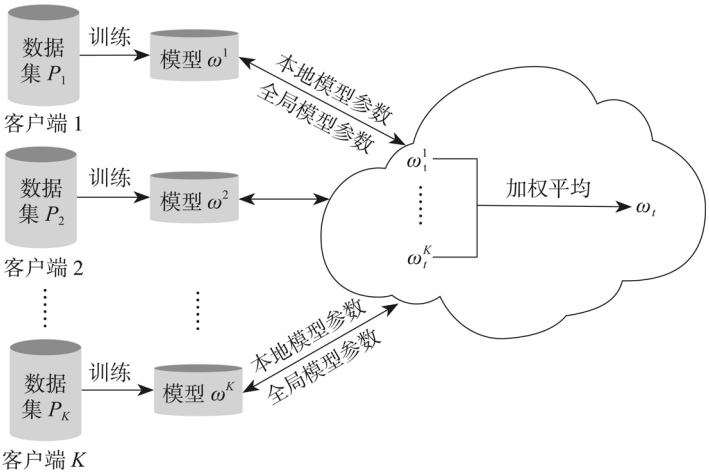

图 1　联邦平均算法的通信模型

联邦平均算法适用于损失函数，可以采用求和方式来表达任意人工智能算法。一般地，定义人工智能算法的优化目标为$\min_{\omega\in\mathbb{R}^d} f(\omega)$。

$$f(\omega) \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} f_i(\omega) \tag{1}$$

式中，n为训练数据的数量。那么当训练数据分散到K个客户端后，客户端$k \in K$拥有数据P_k，规模为$|P_k| = n_k$，则式（1）可重写为

$$f(\omega) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(\omega) \tag{2}$$

其中，

$$F_k(\omega) = \frac{1}{n_k}\sum_{i\in P_k} f_i(\omega)$$

此时，$f(\omega)$的计算代表了服务器的加权平均，而$F_k(\omega)$的计算代表了客户端的本地训练。如果客户端的数据是独立同分布的，那么上述两种计算方式本质上是相同的。当客户端数据不是独立同分布的时，实验表明，联邦平均算法具有很好的表现[2]。特别地，联邦平均算法适用于具有以下特点的数据。

- 非独立同分布：单个客户端数据的分布不能代表全局数据的分布。
- 非平衡：不同客户端的数据量不均匀。
- 大规模分布式：参与的客户端数量远大于每个客户端的平均数据量。

通过对联邦平均算法的探讨，可以看到该算法的计算是带权重的平均。如果权重参数可以公开，那么该计算可以进一步简化为求和运算。这样，安全聚合协议就可以用于横向联邦学习隐私保护了。

4 安全聚合协议

Bonawitz 等人[39]提出了一种安全聚合协议，可以隐私保持地计算若干输入的和。表 2 给出了基本的安全聚合协议的详细描述，图 2 给出了该协议的通信模型。

表 2 基本的安全聚合协议

输入：客户端k的d维模型参数ω^k；服务器的初始用户集合U_0^S；安全参数κ；Diffie-Hellman 密钥协商的群参数(g,q)，即循环群$G =< g >$的生成元g和阶q；Shamir 秘密分享的有限域F及门限值th；对称加密算法AE.enc和相应的解密算法AE.dec；伪随机函数PRF
输出：$\sum_{k\in U_3^S}\omega^k$，其中$U_3^S$集合含有不少于th个客户端
1. **广播密钥**：客户端$k\in U_0^S$生成两对 Diffie-Hellman 密钥协商协议临时密钥对(g^{c^k},c^k)和(g^{s^k},s^k)，其中$c^k,s^k\in\mathbb{Z}_q$，发送$\text{msg}_1^k=\left(k,g^{c^k},c^k\right)$给服务器$S$，进入下一轮； 服务器$S$收集至少th个$\text{msg}_1$，否则超时退出；服务器$S$构造集合$U_1^S$，包含所接收消息的发送者身份，规模为$\lvert U_1^S\rvert=n_1^S$，构造并向$U_1^S$中的客户端发送消息$\text{msg}_1^S=\{\text{msg}_1^k\}$，$k\in U_1^S$
2. **分享秘密**：客户端k确认收到的消息msg_1^S中有至少th对不同的公钥，否则退出；客户端k构造集合U_1^k，包含msg_1^S中所有客户端的身份，规模为$\lvert U_1^k\rvert=n_1^k$；客户端k选择一个随机数b^k，计算两次 Shmair 秘密分享，得到n_1^k份随机数b^k的秘密份额$\{b_u^k\in F\}_{u\in U_1^k}$和临时密钥中$s^k$的$n_1^k$份秘密份额$\{s_u^k\in F\}_{u\in U_1^k}$；客户端$k$计算与客户端$u$，$u\in U_1^k\backslash\{k\}$的共享密钥$\text{key}^{u,k}=\left(g^{c^u}\right)^{c^k}$，用该密钥加密秘密份额对$(b_u^k,s_u^k)$得到密文$\text{AE.enc}_{key^{u,k}}(b_u^k,s_u^k)$；最后客户端$k$构造消息$\text{msg}_2^k=\left(k,\left\{\left(u,\text{AE.enc}_{\text{key}^{u,k}}(b_u^k,s_u^k)\right)\right\}_{u\in U_1^k\backslash\{k\}}\right)$，发送消息$\text{msg}_2^k$给服务器$S$，进入下一轮； 服务器$S$收集至少th个$\text{msg}_2$，否则超时退出；服务器$S$构造集合$U_2^S$，包含所接收消息的发送者身份，规模为$\lvert U_2^S\rvert=n_2^S$，对每个客户端$u\in U_2^S$，构造消息$\text{msg}_{2_u}^S=(u,\left\{(k,\text{AE.enc}_{\text{key}^{u,k}}(b_u^k,s_u^k))\right\}_{k\in U_2^S\backslash\{u\}})$，并发送消息$\text{msg}_{2_u}^S$给该客户端

续表

3. **收集密文**：客户端k根据收到的消息中密文的发送者构造集合U_2^k，设定U_2^k=$U_2^k \cup \{k\}$，$\left|U_2^k\right| = n_2^k$，检查$n_2^k \geqslant$th，否则退出；客户端$k$存储消息$\text{msg}_{2_k}^S$；对每个客户端$u \in U_2^k \backslash \{k\}$，客户端$k$计算与$u$的共享密钥$sk^{u,k} = \left(g^{s^u}\right)^{s^k}$，然后计算$d$维向量$zs^{u,k} = \delta_{u,k} \cdot \text{PRF}(sk^{u,k})$，其中当$k > u$时，$\delta_{u,k} = 1$，否则$\delta_{u,k} = -1$，进而计算$\text{mask}_1^k = \sum_{u \in U_2^k \backslash \{k\}} zs^{u,k}$；客户端$k$计算$\text{mask}_2^k = \text{PRF}(b^k)$；客户端$k$计算$\gamma^k = \omega^k + \text{mask}_1^k + \text{mask}_2^k$；客户端$k$发送$\text{msg}_3^k = (k, \gamma^k)$给服务器$S$，并进入下一轮； 服务器$S$接收至少th个$\text{msg}_3$，否则超时退出；服务器$S$构造集合$U_3^S$，包含所接收消息的发送者身份，规模为$\left|U_3^S\right| = n_3^S$，构造消息$\text{msg}_3^S = U_3^S$，发送给$U_3^S$集合中的客户端
4. **聚合解密**：客户端k确认收到的msg_3^S中包含不同身份的数量至少为th，形成U_3^k集合，否则退出；客户端k解密消息$\text{msg}_{2_k}^S$中的密文，恢复U_2^k中的客户端给k的秘密份额；客户端k构造并发送消息$\text{msg}_4^k = \{(u, bs^u)\}_{u \in U_2^k}$，其中如果$u \notin U_3^k$，$bs^u = s_k^u$，否则$bs^u = b_k^u$；之后退出协议； 服务器$S$接收至少th个$\text{msg}_4^k$，形成$U_4^S$集合，否则超时退出；对于$k \in U_2^S \backslash U_3^S$的客户端，服务器$S$使用 Shamir 秘密恢复得到临时密钥$s^k$，进而对于$u \in U_3^S$，计算$\text{mask}_3^u = \sum_{k \in U_2^S \backslash U_3^S} zs^{u,k}$；另外对于$u \in U_3^S$，服务器$S$使用 Shamir 秘密恢复得到随机数$b^u$，进而能够重新计算$\text{mask}_2$；最后聚合解密得到$\omega = \sum_{k \in U_3^S} \omega^k = \sum_{k \in U_3} (\gamma^k - \text{mask}_2 - \text{mask}_3^u)$

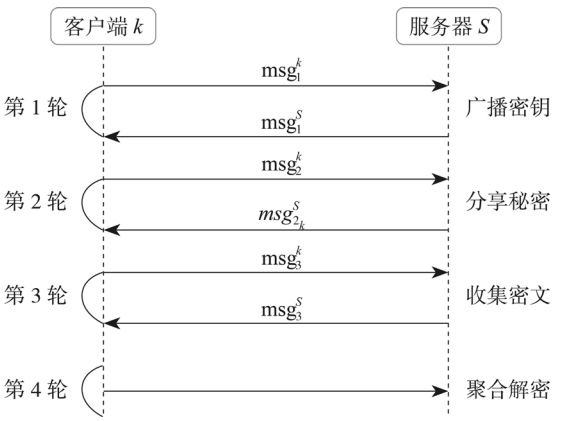

图 2　基本的安全聚合协议的通信模型

4.1　协议设计的原理

Bonawitz 等人[39]给出的是一个针对实际场景的协议。实际场景的主要问题是模型参数规

模过大，难以采用同态加密等技术处理。模型参数的维度很容易超过十万，稍微复杂一些的神经网络模型参数的规模就会超过百万，甚至千万。这样的规模给同态加密带来了严重挑战。例如，采用 Paillier 同态加密[53]，假设明文空间为 2048 比特，每个模型参数为 8 比特，客户端数量为 128，则即便把多个参数放在一个明文中，采用所谓的单指令多数据（SIMD）技术，十万维度也意味着约 782 个密文，百万维度还要乘以 10，计算和通信代价还是比较高的。因此，该协议采用了伪随机函数（PRF）来处理高维数据，降低计算和通信的代价。

在协议设计上，该协议采用了成熟的 Diffie-Hellman 密钥协商协议，利用了共享密钥的对称性。该协议设计时并没有寻找具有同态能力的 PRF，而是通过构造的方法形成了可以对消的掩码。表 3 给出了一个 5 个客户端的例子。假设这 5 个客户端都属于U_2^S，且$U_2^S \backslash U_3^S = \emptyset$，那么所有用户的$\text{mask}_1$之和恰好为 0。

表 3　可对消掩码示例

	1	2	3	4	5
1		$-\text{PRF}(sk^{2,1})$	$-\text{PRF}(sk^{3,1})$	$-\text{PRF}(sk^{4,1})$	$-\text{PRF}(sk^{5,1})$
2	$\text{PRF}(sk^{1,2})$		$-\text{PRF}(sk^{3,2})$	$-\text{PRF}(sk^{4,2})$	$-\text{PRF}(sk^{5,2})$
3	$\text{PRF}(sk^{1,3})$	$\text{PRF}(sk^{2,3})$		$-\text{PRF}(sk^{4,3})$	$-\text{PRF}(sk^{5,3})$
4	$\text{PRF}(sk^{1,4})$	$\text{PRF}(sk^{2,4})$	$\text{PRF}(sk^{3,4})$		$-\text{PRF}(sk^{5,4})$
5	$\text{PRF}(sk^{1,5})$	$\text{PRF}(sk^{2,5})$	$\text{PRF}(sk^{3,5})$	$\text{PRF}(sk^{4,5})$	

进一步，该协议采用了双掩码，防止网络延迟带来的用户模型参数泄露。和 Bonawitz 等人[39]在论文中讨论的一样，可以首先考虑仅有单个掩码mask_1的情况。假设服务器S在第 3 轮收集密文时因为网络延迟没有收到某个客户端k^*的消息$\text{msg}_3^{k^*}$，发送了消息msg_3^U。之后，客户端k^*的消息$\text{msg}_3^{k^*}$又被服务器S收到了。此时按照协议，客户端会将k^*关于s^{k^*}的份额提交给服务器S。如果只有掩码mask_1，服务器S就可以恢复$\text{msg}_3^{k^*}$中的模型参数了。但是如果有两个掩码mask_1和mask_2，而客户端又只提供其中一个掩码的份额，服务器S就无法从延迟消息中恢复客户端的模型参数。

最后，该协议采用了 Shamir 秘密分享来对抗客户端掉线的情况。从表 1 中的联邦平均算法可以看到，当每轮训练时，服务器都会随机选择$C \cdot K$个客户端。这些客户端在表 2 的协议中对应U_0^S。在预想的场景中，移动端随时可能会退出训练，造成客户端掉线。因此，在执行表 2 协议的过程中，客户端的数量可能会越来越少，也就是$U_3^k \subseteq U_2^k \subseteq U_1^k$和$U_4^S \subseteq U_3^S \subseteq U_2^S \subseteq U_1^S \subseteq U_0^S$，且$U_i^k \subseteq U_i^S$，$1 \leqslant i \leqslant 3$。那么当$U_2^S \backslash U_3^S \neq \emptyset$时，表 3 的对消并不能实现。因此，需要一种方法来恢复不能对消的共享秘密，顺利完成聚合解密，这种方法即第 2 轮的 Shamir 秘密分享。事实上，如果不考虑客户端掉线的情况，表 2 的协议可以简化很多，也不用 4 轮通信这么多。

4.2　协议的安全性

安全聚合协议假设了半诚实的攻击者，即攻击者可以腐化服务器或客户端，但是需要诚实地执行协议。攻击者的目标是诚实用户的模型参数。首先我们注意到有一些安全聚合协议是假设所有的客户端有一个共同的密钥，而服务器不知道这个密钥，然后进行的隐私保护计算。这个假设在上述半诚实攻击者下是不成立的，因为攻击者可以腐化客户端，所以知道客户端的密钥，进而腐化服务器，通过服务器来获取客户端的输入，并解密获得模型参数。然后我们还需要看到这个假设依旧很弱，如在该假设下，即使攻击者腐化了服务器，也只能按照服务器的协议执行，而不能任意发送消息。

在半诚实的攻击者假设下，Bonawitz 等人[39]证明了协议的安全性，我们抄录其定理 6.3，在本文中编号为定理 4-1。

定理 4-1　存在一个概率多项式时间的仿真者SIM，对于任意给定的安全参数κ，门限值th，用户集合U_0^S，用户集合的输入ω_U，训练过程中服务器形成的用户标识集合U_1^S、U_2^S、U_3^S和U_4^S，腐化实体的集合C，如果满足$C \subseteq U_0^S \cup \{S\}$，$|C\backslash\{S\}| < \text{th}$，$U_4^S \subseteq U_3^S \subseteq U_2^S \subseteq U_1^S \subseteq U_0^S$，那么SIM的输出与真实交互$\text{REAL}_C^{U_0^S,\text{th},\kappa}$的输出在计算上不可区分，即

$$\text{REAL}_C^{U_0^S,\text{th},\kappa}(\omega_U, U_1^S, U_2^S, U_3^S, U_4^S) \approx_c \text{SIM}_C^{U_0^S,\text{th},\kappa}(\omega_C, z, U_1^S, U_2^S, U_3^S, U_4^S) \qquad (3)$$

其中，当$|U_3^S| \geqslant$th时，$z = \sum_{u \in U_3^S \backslash C} \omega_u$，否则无定义。

定理 4-1 表明，只要攻击者腐化的客户端数量少于 Shamir 秘密分享的门限值，存在仿真者，在仅有腐化用户的输入和真实环境下输出的情况下，可以给出真实情况下的全部输出脚本，而该脚本在攻击者看来是与真实情况下的交互不可区分的。也就是说，攻击者从真实情况下的脚本中难以获得诚实用户的隐私输入。

4.3　协议的增强版本

当考虑腐化的服务器可以发送任意消息时，表 2 的基本协议是不安全的。特别地，腐化的服务器可以通过向不同的客户端发送不同的消息，使得不同客户端的U_1^k、U_2^k、U_3^k有所不同，从而能够获得某个客户端k^*的隐私输入。一个简单的巫师攻击是腐化的服务器S自身生成许多的公私钥对，然后在U_1^S中放入客户端k^*的公钥和自己生成的公钥，发送给客户端k^*；最后腐化的服务器S冒充不同的客户端分享秘密份额给客户端k^*，同时获得客户端k^*的秘密份额；最后在客户端k^*发送密文后，腐化的服务器S就可以用恢复的客户端k^*分享的秘密来解密，获得k^*的模型参数。

为此，Bonawitz 等人[39]给出了一个对抗主动攻击者的增强的安全聚合协议。在具体陈述协议构造之前，我们先来明确论文中主动攻击者的含义。主动攻击者是指在半诚实攻击者能力的基础上，腐化的实体可以进一步背离协议，发送错误的或任意选择的消息给诚实的客户端，可以终止或不响应消息。显然，我们刚刚给出的巫师攻击场景中的攻击者就是一个主动攻击者，通过构造服务器响应消息完成了攻击。

在协议设计上，增强的安全聚合协议增加了公钥基础设施（PKI）、数字签名和一轮称之为一致性检查的消息交互。表 4 给出了该协议。相应地，图 3 是该协议的通信模型。

表 4 中的协议，第 1 轮的数字签名，意图在于防止腐化的服务器的巫师攻击，比较好理解。一致性检查是为了防止腐化的服务器通过给不同的客户端不同的U_3^S视图，以获得更多的份额。表 5 给出了一个较为具体的例子。

表 4　增强的安全聚合协议

输入：在表 2 协议的基础上，增加证书权威 CA 或假设服务器在第 1 轮是半诚实的。客户端k有签名私钥σk_k及验证公钥vk_k；当有 CA 时，客户端有包含vk_k的数字证书cert_k；数字签名算法Sig. sign和验证算法Sig. ver
输出：与表 2 协议相同
1. **广播密钥**：客户端k额外生成数字签名$\sigma_0^k = \mathrm{Sig.sign}(\sigma k_k, g^{c^k}, g^{s^k})$；然后消息$\mathrm{msg}_1^k = \left(k, g^{c^k}, c^k, \sigma_0^k\right)$；其他不变。服务器$S$不变
2. **分享秘密**：客户端k确认收到的消息msg_1^S中其他客户端的数字签名都是对的，如果有证书，那么需要验证数字证书都是好的，然后执行其他操作。服务器 S 不变
3. **收集密文**：客户端k不变。服务器S不变
4. **一致性检查**：客户端k确认收到的msg_3^S中包含不同身份的数量至少为th，形成U_3^k集合，否则退出；客户端k生成数字签名$\sigma_1^k = \mathrm{Sig.sign}(\sigma k_k, U_3^k)$，形成并发送消息$\mathrm{msg}_4^k = (k, \sigma_1^k)$给服务器$S$，然后进入下一轮；服务器$S$接收至少$th$个$\mathrm{msg}_4^k$，形成$U_4^S$集合，否则超时退出；形成并发送$\mathrm{msg}_4^S = \{\mathrm{msg}_4^k\}_{k \in U_4^S}$给在$U_4^S$集合中的客户端
5. **聚合解密**：客户端k确认收到的msg_4^S中的数字签名都是正确的，形成U_4^k集合，然后进行后续操作。服务器S不变

表 5　腐化的服务器给不同客户端不同的$U_2^S \backslash U_3^S$视图

1	2	3	4	5	6	7	8	9
		5		3	3	3	3	3
		6		6	5	5	5	4
						6	6	6

在表 5 中，一共有 9 个客户端，其中客户端 1 和 2 是腐化的客户端，客户端 4 是攻击者的攻击目标。秘密分享体制的门限$\mathrm{th}=6$。攻击者通过给不同的客户端不同的视图来进行攻击。具体来看，腐化的服务器在第 2 轮设置$U_{2_4}^S=\{1,2,\cdots,6\}$，设置$U_{2_i}^S=\{1,2,\cdots,9\}$，$i\neq 4, i\in\{1,2,\cdots,9\}$。然后在第 3 轮给客户端 3 发送$U_{3_3}^S=\{1,2,3,4,7,8,9\}$，给客户端 5 发送$U_{3_5}^S=\{1,2,4,5,7,8,9\}$，给客户端 6 发送$U_{3_6}^S=\{1,2,4,6,7,8,9\}$，给客户端 7、8、9 发送$U_{3_{7,8,9}}^S=\{1,2,4,7,8,9\}$。这样对每个诚实客户端而言，都可以进入第 4 轮，向服务器发送份额。因为用户 1 和 2 是腐化的，所以攻击者从 1、2 可以得到每个其他客户端每个秘密的 2 个份额，进而可以获得客户端 3、5、6 的 Diffie-Hellman 临时私钥，同时可以获得客户端 4 的b^4相关份额。那么，因为客户端 4 的$U_{2_4}^S=\{1,2,\cdots,6\}$，其掩码$\mathrm{mask}_1^4$可以通过其他客户端的临时私钥去掉，其掩码$\mathrm{mask}_2^4$可以通过$b^4$去掉，所以攻击者就获得了客户端 4 的模型参数。显然，该攻击可以通过检查服务器U_2^S或U_3^S的一致性来防范。一致性检查完成的正是对U_3^S的一致性检查。

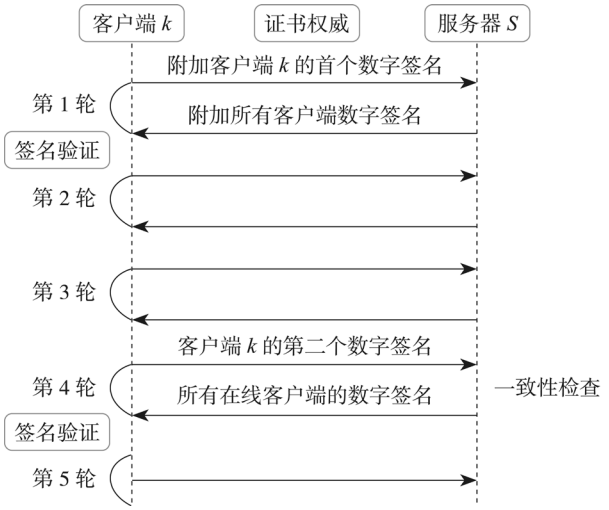

图 3　增强的安全聚合协议的通信模型及主要内容

Bonawitz 等人[39]给出了表 4 增强的安全聚合协议的一个定理，我们抄录其定理 6.5，编号为定理 4-2。

定理 4-2　存在一个概率多项式时间的仿真者（SIM），对于任意给定的安全参数κ，门限值th，用户集合U_0^S，腐化实体的集合$C\subseteq U_0^S\cup\{S\}$，诚实用户集合的输入$\omega_{U\backslash C}$，设$|U_0^S|=n$，$|C\backslash\{S\}|=n_C$，如果$2\mathrm{th}>n+n_C$，那么SIM的输出与真实交互$\mathrm{REAL}_C^{U_0^S,\mathrm{th},\kappa}$的输出在计算上不可区分，即

$$\mathrm{REAL}_{C}^{U_0^S,\mathrm{th},\kappa}\left(\omega_{U\setminus C},M_C\right)\approx_c \mathrm{SIM}_{C}^{U_0^S,\mathrm{th},\kappa,\mathrm{Ideal}_{\{\omega_u\}_{u\in U_0^S\setminus C}}^{\delta}}(M_C) \tag{4}$$

其中，$\delta=\mathrm{th}-n_C$，M_C表示概率多项式时间的攻击者，Ideal表示诚实用户谕言机。

定理 4-2 表明，只要攻击者腐化的客户端数量与初始客户端的总量之和小于 2 倍的门限值，对于表 4 的协议，诚实用户的输入隐私可以得到保护。

需要注意的是，上述定理证明中所用的条件并没有在协议设计上得到保证。下面给出两个观察。

（1）定理中的参数n在协议中并没有得到客户端的确认。事实上，客户端关于n的设定是从服务器来的，而服务器可能已经腐化。一个简单的攻击场景是腐化的服务器设置U_0^S为平时的 2 倍大，并划分为规模类似的$U_{0_1}^S$和$U_{0_2}^S$，并设置$\mathrm{th}<|U_0^S|/2$。在第 1 轮广播公钥之后，腐化的服务器S向客户端k^*输入$U_1^S\subseteq U_0^S$，然后向其他客户端输入$U_{1_1}^S\subseteq U_{0_1}^S$或$U_{1_2}^S\subseteq U_{0_2}^S$；在收到客户端$\mathrm{msg}_2^{k^*}$之后，腐化的服务器$S$向该客户端提供足够的密文，让该客户端继续发送$\mathrm{msg}_3^{k^*}$；同时设置$k^*\in U_{3_1}^S$且$k^*\neq U_{3_2}^S$，以从$U_{3_1}^S$的诚实客户端和$U_{3_2}^S$的诚实客户端分别获得客户端$k^*$的th份秘密份额，恢复客户端$k^*$的$s^{k^*}$和$b^{k^*}$，进而可以去掉$\mathrm{msg}_3^{k^*}$的掩码，获得该客户端的模型参数。这个简单的攻击当然很容易防范，对于客户端k^*来说，其门限值th显然过小，但是这些检查并没有在表 4 的协议中体现。

（2）定理中的参数n_C在协议中也没有让客户端确认。事实上，客户端通过在第 1 轮检查有效数字签名的数量来确认不同的其他客户端的数量进而推断n，并假设被腐化的客户端只是少量达到间接认可n_C的效果。然而第 1 轮有效的数字签名并没有新鲜因子，因此是可以重放的，这直接导致有效的数字签名的数量不能与参与一轮训练的用户数量对等起来。假设训练轮次是 1000 轮，每轮选择大约 100 个用户，那么腐化的服务器只要在前 99 轮每轮腐化一个用户，就可以从第 100 轮开始，用腐化的 99 个用户的临时密钥及其数字签名来进行类巫师攻击，获取任意诚实用户的模型参数。这个简单的攻击可以通过在数字签名的内容中增加时间戳等新鲜因子来防范。

5 安全聚合协议的发展

本节主要介绍基于同态伪随机函数的安全聚合协议[49]、基本和增强的随机图安全聚合协议[46]这 3 个协议。

5.1 同态伪随机函数方法

针对 Bonawitz 等人[39]的工作，南洋理工大学的 Liu 等人[49]给出了基于同态 PRF 的改进方案。前面已经指出，对于大规模模型参数的保护，PRF 是一种较为经济有效的方法。Bonawitz

等人[39]利用了 Diffie-Hellman 共享密钥的对称性和 PRF 来设计协议。Liu 等人[49]则是结合了一种同态 PRF[54]来设计协议。

表 6 给出了 Liu 等人[49]基础的安全聚合协议。我们调整了一些步骤，使得该协议的通信模型与 Bonawitz 等人[39]的通信模型一致，即客户端间不直接通信。图 4 给出了该协议的通信模型。

表 6　基于同态 PRF 的基础安全聚合协议

输入：客户端k的d维模型参数ω^k；服务器的初始用户集合U_0^S；安全参数κ；同态 PRF 的群参数(g,q)，即循环群$G =< g >$的生成元g和阶q；Shamir 秘密分享的有限域F及门限值th；客户端k的长期加密密钥对$(\mathrm{cpk}^k,\mathrm{csk}^k)$；服务器的长期加密密钥对$(\mathrm{cpk}^S,\mathrm{csk}^S)$；公钥加密算法 Enc 及解密算法 Dec；安全的哈希函数 H
输出：$\sum_{k\in U_3^S}\omega^k$，其中$U_3^S$集合含有不少于th个客户端
1. **分享秘密**：客户端k选择一个随机数$s^k \in \mathbb{Z}_q$，计算 Shmair 秘密分享，得到$\lvert U_0^S\rvert = n$份随机数s^k的秘密份额$\{s_u^k \in F\}_{u\in U_0^S}$；客户端$k$根据$U_0^S$获得其他客户端的加密公钥；对每个客户端$u$，客户端$k$加密秘密份额$s_u^k$得到密文$\mathrm{Enc}_{\mathrm{cpk}^u}(s_u^k)$；最后客户端$k$构造消息$\mathrm{msg}_1^k = \left(k,\{(u,\mathrm{Enc}_{\mathrm{cpk}^u}(s_u^k))\}_{u\in U_0^S\setminus\{k\}}\right)$，发送消息$\mathrm{msg}_1^k$给服务器$S$，进入下一轮； 服务器$S$收集至少th个$\mathrm{msg}_1$，否则超时退出；服务器$S$构造集合$U_1^S$，包含所接收消息的发送者身份，规模为$\lvert U_1^S\rvert = n_1^S$，对每个客户端$u \in U_1^S$，构造消息$\mathrm{msg}_{1_u}^S = (u,\{(k,\mathrm{Enc}_{\mathrm{cpk}^u}(s_u^k))\}_{k\in U_1^S\setminus\{u\}})$，并发送消息$\mathrm{msg}_{1_u}^S$给该客户端
2. **收集密文**：客户端k根据收到的消息$\mathrm{msg}_{1_k}^S$中密文的发送者构造集合U_1^k，设定U_1^k=$U_1^k \cup \{k\}$，$\lvert U_1^k\rvert = n_1^k$，检查$n_1^k \geqslant$th，否则退出；客户端$k$解密$\mathrm{msg}_{1_k}^S$中的密文，得到份额集合$\{s_k^u\}_{u\in U_1^k}$；计算$\gamma^k = \{\gamma_i^k = g^{\omega_i^k}H(i)^{s^k}\}_{i\in\{1,\cdots,d\}}$。客户端$k$发送$\mathrm{msg}_2^k = (k,\gamma^k)$给服务器$S$，并进入下一轮； 服务器$S$接收至少th个$\mathrm{msg}_2$，否则超时退出；服务器$S$构造集合$U_2^S$，包含所接收消息的发送者身份，规模为$\lvert U_2^S\rvert = n_2^S$，构造消息$\mathrm{msg}_2^S = U_2^S$，发送给$U_2^S$集合中的客户端
3. **聚合解密**：客户端k确认收到的msg_2^S中包含不同身份的数量至少为th，形成U_2^k集合，否则退出；客户端k计算$b^k = \sum_{u\in U_2^k} s_k^u$；客户端$k$构造消息$\mathrm{msg}_3^k = (k,b^k)$；客户端$k$给服务器$S$发送消息$\mathrm{msg}_3^k$，之后退出协议；服务器$S$接收至少th个$\mathrm{msg}_3^k$，形成$U_3^S$集合，否则超时退出；服务器$S$使用 Shamir 秘密恢复得到临时密钥$b = \sum_{u\in U_2^S} s^u$，进而计算$g^{\omega_i} = g^{\sum_{u\in U_2^S}\omega_i^u} = \prod_{u\in U_2^S}\gamma_i^u\,/H(i)^b$；之后计算离散对数得到$\omega = \{\omega_i\}_{i\in\{1,\cdots,d\}}$

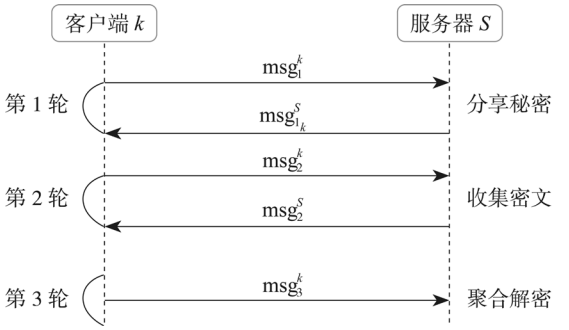

图 4　基于同态 PRF 的基础安全聚合协议通信模型

1．协议的设计原理

表 6 的基础安全聚合协议通信轮数较低，主要是通过将加密密钥长期化，省去了广播密钥的步骤。在实际操作中，每个用户在向服务器注册参与训练时，就可以把长期密钥传递给服务器，之后服务器在每轮选择$C \cdot K$个客户端时，就可以将这些客户端所需要的加密密钥传输给这些客户端。当然也可以由每个被选中的客户端根据自身缓存其他客户端加密密钥的情况，向服务器拉取自己所需的公钥。

该协议设计上暗含了一个分布式 Shamir 秘密分享的过程。表 7 给出了一个包含 3 个客户端的例子来展示这一过程。设3 个客户端{1，2，3}分别有秘密s^1、s^2、s^3要分享，按照 Shamir 秘密分享的方法，每个客户端选一个随机多项式计算给其他客户端的份额。表 7 表明每个客户端对份额求和之后，得到的 3 个份额恰好对应 3 个秘密之和的份额。

表 7　分布式 Shamir 秘密分享

	秘密	多项式	1	2	3
1	s^1	$s^1 + xf^1(x)$	s_1^1	s_2^1	s_3^1
2	s^2	$s^2 + xf^2(x)$	s_1^2	s_2^2	s_3^2
3	s^3	$s^3 + xf^3(x)$	s_1^3	s_2^3	s_3^3
求和	$b = \sum_{u=1}^{3} s^u$	$\sum_{u=1}^{3} (s^u + xf^u(x))$	$b^1 = \sum_{u=1}^{3} s_1^u$	$b^2 = \sum_{u=1}^{3} s_2^u$	$b^3 = \sum_{u=1}^{3} s_3^u$

2．协议的性能和安全性

在性能上，表 6 中所示协议每个用户都减少了一个待分享的秘密，降低了一半与秘密份额相关的通信量和计算量。然而，在每个用户加密时，即使通过预计算先把$H(i)^{s^k}$算出来，也需要多次乘法才能完成单个维度的加密，当维度达几百万时，相比于 Bonawitz 等人[39]的工作，加密时间显然要长一些。另外，服务器需要进行离散对数的计算，当模型参数和用户数

都较小时，可以通过查表的方式快速完成，否则需要通过如大步小步法等算法来完成，需要相对较长的时间。

在安全性方面，Liu 等人[49]提供了 Bonawitz 等人[39]增强协议的版本，采用的是类似的技术。针对巫师攻击的问题，通过数字签名和数字证书的方式来防止；针对不同客户端U_2^S可能不同的问题，采用了增加一致性检查的方式来避免。另外，Liu 等人[49]的协议设计上不需要考虑客户端消息延迟的问题。因为每个诚实客户端提交的都是共享份额的和，所以服务器并不能恢复单个用户的秘密。因此，服务器无法通过延迟消息获利。

5.2　随机图方法

如前所述，针对 Bonawitz 等人[39]的工作，Choi 等人[45]和 Bell 等人[46]都给出了采用随机图的提议，以降低通信量与计算量。在 Bonawitz 等人[39]的工作中，每个客户端向其他所有客户端发送秘密份额，形成客户端之间的全连接关系。采用随机图的基本思想是客户端向随机的一些客户端发送秘密份额，与随机的一些客户端建立 Diffie-Hellman 共享秘密。

Choi 等人[45]使用了 Erdos-Renyi（ER）随机图，并假定 ER 随机图中顶点的连接关系是在可信环境下生成并得到所有参与者认可的。ER 随机图中的顶点即客户端，具有连接关系的两个客户端彼此交换秘密份额，并建立共享秘密。Bell 等人[46]采用了随机化的哈拉里图，分别设计了两种哈拉里图的生成协议，用于对抗半诚实和主动攻击者。下面介绍 Bell 等人[46]的工作。表 8 给出了基本的随机图安全聚合协议[46]。

表 8　基本的随机图安全聚合协议

输入：客户端k的d维模型参数ω^k；服务器的初始用户集合U_0^S；安全参数κ；Diffie-Hellman 密钥协商的群参数(g,q)，即循环群$G =< g >$的生成元g和阶q；Shamir 秘密分享的有限域F及门限值th；对称加密算法AE.enc和相应的解密算法AE.dec；伪随机函数PRF；用户最大掉线率δ
输出：$\sum_{k\in U_3^S}\omega^k$，其中$U_3^S$集合含有不少于th个客户端
1. **广播密钥**：客户端k的操作与表 2 协议相同； 服务器S收集至少$(1-\delta)\lvert U_0^S\rvert$个$\mathrm{msg}_1$，否则超时退出；服务器$S$构造集合$U_1^S$，包含所接收消息的发送者身份，规模为$\lvert U_1^S\rvert = n_1^S$；服务器$S$生成顶点数为$n_1^S$，边数为$n_e^S$的哈拉里图，之后对图的顶点进行随机置换，得到随机的哈拉里图；对于图中的一个客户端u，用$N_G(u)$表示该客户端的所有邻居客户端；对每个客户端u，服务器S构造消息$\mathrm{msg}_{1_u}^S = \{\mathrm{msg}_1^k\}_{k\in N_G(u)}$，然后向其发送消息$\mathrm{msg}_{1_u}^S$

续表

2. **分享秘密**：客户端k的操作与表 2 相同，唯一的差别在于处理的是消息$\mathrm{msg}_{1_k}^S$； 服务器S收集至少$(1-\delta)n_1^S$个msg_2，否则超时退出；服务器S构造集合U_2^S，包含所接收消息的发送者身份，规模为$\lvert U_2^S\rvert = n_2^S$，对每个客户端$u \in U_2^S$，构造消息$\mathrm{msg}_{2_u}^S = (u, \{(k, \mathrm{AE.enc}_{\mathrm{key}^{u,k}}(b_u^k, s_u^k))\}_{k\in U_2^S \cap D_G(u)})$，并发送消息$\mathrm{msg}_{2_u}^S$给该客户端
3. **收集密文**：客户端k的操作与表 2 相同； 服务器S收集至少$(1-\delta)n_1^S$个msg_3，否则超时退出；服务器S构造集合U_3^S，包含所接收消息的发送者身份，规模为$\lvert U_3^S\rvert = n_3^S$。对每个客户端$u \in U_3^S$，构造消息$\mathrm{msg}_{3_u}^S = U_3^S \cap D_G(u)$，发送给该客户端
4. **聚合解密**：客户端k的操作与表 2 相同，唯一的差别在于处理的是消息 $\mathrm{msg}_{3_k}^S$； 服务器S收集至少$(1-\delta)n_1^S$个msg_3，否则超时退出；服务器S构造集合U_4^S，包含所接收消息的发送者身份，规模为$\lvert U_4^S\rvert = n_4^S$。对于$k \in U_2^S \backslash U_3^S$的客户端，服务器$S$使用 Shamir 秘密恢复得到临时密钥$s^k$，进而对于$u \in U_3^S$，计算$\mathrm{mask}_3^u = \sum_{k\in U_2^S\backslash U_3^S \cap D_G(u)} zs^{u,k}$。另外，对于$u \in U_3^S$，服务器$S$使用 Shamir 秘密恢复得到随机数$b^u$，进而能够重新计算$\mathrm{mask}_2^u$；最后聚合解密得到$\omega = \sum_{k\in U_3^S} \omega^k = \sum_{k\in U_3}(\gamma^k - \mathrm{mask}_2^u - \mathrm{mask}_3^u)$

1．协议的设计原理

对比表 2 和表 8 可知，基本的安全聚合协议和基本的随机图安全聚合协议有以下 2 个不同点。

（1）表 8 的协议由服务器生成一个随机图，并根据该图确定了每个客户端的邻居关系，之后每个客户端的操作仅与邻居节点相关。

（2）表 8 的协议中服务器不能再用门限来判断每轮接收消息的数量是否足够，因此采用了一个掉线率参数δ来确保协议可以进行下去，最终第 4 轮能否恢复临时密钥和随机数，取决于掉线率δ和随机图的n_e^S连通度。

随机图的属性对于表 8 协议的正确性有影响。Bell 等人[46]给出了随机图应该具有的 3 个属性，通过定义 5-1~定义 5-3 的描述。

定义 5-1（有限腐化邻居事件） 令C为被腐化的客户端集合，$C \subseteq U_1^S$，$\lvert U_1^S\rvert = n$，$\lvert C\rvert \leqslant \gamma n$。事件$E_1(C, G, \mathrm{th}) = 1$定义为对于所有的客户端$u \in U_1^S \backslash C$，都有$\lvert N_G(u) \cap C\rvert < \mathrm{th}$。

定义 5-2（客户掉线连通事件） 令C为被腐化的客户端的集合，D为掉线客户端的集合，$D, C \subseteq U_1^S$，$\lvert U_1^S\rvert = n$，$\lvert C\rvert \leqslant \gamma n$，$\lvert D\rvert \leqslant \delta n$。事件$E_2(C, D, G) = 1$定义为$G(U_0^S \backslash C \backslash D)$是连通的。

定义 5-3（客户掉线重构事件） 令D为掉线客户端的集合，$D \subseteq U_0^S$，$\lvert U_0^S\rvert = n$，$\lvert D\rvert \leqslant \delta n$。

事件$E_3(D,G,\text{th})=1$定义为对于所有的客户端$u\in U_0^S$，都有$|N_G(u)\cap(U_0^S\backslash D)|\geqslant$th。

显然，如果服务器生成的随机图使得事件E_3以极大概率发生，那么表 8 第 4 轮解密失败的概率可以忽略。如果事件E_1以极大概率发生，就满足了表 2 协议的安全性要求，可以对抗半诚实的攻击者。事件E_2是在随机图情况下出现的一个独有事件。如果$E_2(C,D,G)=0$出现，意味着掉线的用户导致图 G 成为了至少 2 个连通子图，从表 8 的协议可以明确看到，对于每个连通子图，服务器都可以恢复该子图的模型参数之和，因此服务器至少可以得到两个部分和，这与协议的安全目标，让服务器只得到U_3^S中客户端模型参数之和相违背。

如果一个随机图满足定义 5-1~定义 5-3，就称该随机图为一个好随机图[46]。

定义 5-4（好随机图）　令$\mathcal{D}$是数据对(G,th)上的一个分布。称$\mathcal{D}$分布是(σ,η)好的，如果对所有掉线客户端的集合和腐化客户端集合，$D,C\subseteq U_0^S$，$|U_0^S|=n$，$|D|\leqslant\delta n$，$|C|\leqslant\gamma n$，满足：

（1）$\Pr[E_1(C,G',\text{th}')\wedge E_2(C,D,G',\text{th}')=1|(G',\text{th}')\leftarrow\mathcal{D}]>1-2^{-\sigma}$；

（2）$\Pr[E_3(D,G',\text{th}')=1|(G',\text{th}')\leftarrow\mathcal{D}]>1-2^{-\eta}$。

表 8 中服务器用随机置换和哈拉里图生成了一个随机图，该随机图可以被证明其极大概率是一个好随机图。Bell 等人[46]给出了定理 3.10 来表明这一事实，本文编号为定理 5-1。

定理 5-1　设$\gamma,\delta\in[0,1]$满足$\frac{\gamma n}{n-1}+\delta<1$。表 8 中生成随机图的算法可以形成关于对$(G,\text{th})$上的一个分布$\mathcal{D}$，该分布是$(\sigma,\eta)$好的，只要$\beta=\text{th}/n_e^S$满足$\frac{\gamma n}{n-1}<\beta<1-\delta$，且

$$n_e^S\geqslant\max\left(\frac{(\sigma+1)\ln 2+\ln n}{c}+1,\frac{\eta\ln 2+\ln n}{2\left(\frac{n(1-\delta)}{n-1}-\beta\right)^2}\right)\tag{5}$$

其中，$c=\min(2(\beta-2\gamma)^2,-\ln(\gamma+\delta))$。

可以看到，定理 5-1 较为复杂，需要对多个参数进行调整。Bell 等人[46]给出了一个具体的例子，其中$\gamma=\delta=0.2$，$\beta=0.5$，$n=10^6$，$\sigma=40$，$\eta=30$，代入定理 5-1 的表达式可以得到$n_e^S\geqslant 2113$。通信代价和计算代价仅约为表 2 基本协议的0.2%。

为了更为准确地理解随机图方法，下面摘录 Bell 等人[46]关于定理 5-1 的推导过程。首先是一个教科书中可以看到的超几何分布。

定义 5-5（超几何分布）　超几何分布是一种离散概率分布，该分布的随机变量表示为$X\sim H(n,m,n_e)$，描述的是在n个物件中，有m个特定类型的物件，从中不放回地取出n_e个物件，成功抽取出特定类型的物件X次。该分布的概率质量函数为

$$P(X=t)=\frac{C_m^t\cdot C_{n-m}^{n_e-t}}{C_n^{n_e}}\tag{6}$$

令$\gamma=\frac{m}{n}$，随机变量X的期望为$E(X)=n_e\gamma$。该分布有以下两个边缘概率分布。

（1）对于所有的$\gamma \geqslant d \geqslant 0$，有$\Pr[X \leqslant (\gamma - d)n_e] \leqslant \mathrm{e}^{-2d^2 n_e}$。

（2）对于所有的$1-\gamma \geqslant d \geqslant 0$，有$\Pr[X \geqslant (\gamma + d)n_e] \leqslant \mathrm{e}^{-2d^2 n_e}$。

该分布与随机图方法的关系在于，当$n = |U_0^S|$，$m = |C|$，$n_e = n_e^S$时，超几何分布描述了随机图中任意一个客户端，其邻居节点中腐化客户端的均值为$n_e\gamma$。因此，当客户端总体中有1/3是腐化时，从期望的角度，随机图中一个节点的邻居应该也是1/3是腐化的。而两个边缘概率分布，则描述了实际采样的随机图中，一个节点的邻居中腐化节点的数量比期望值少dn_e，或者比期望值多dn_e，其概率都是指数下降的。在假设掉线用户为均匀随机分布时，令$m = |D|$，超几何分布也可以描述一个节点的邻居节点掉线的概率，图5给出了当$n = 10^4$，$n_e = 200$，$\gamma = 0.2$，$\delta = 0.1$时，随机图中一个节点的邻居节点中腐化节点数量取不同值的概率X_i和不掉线节点的数量取不同值的概率Y_i。当取门限值$\mathrm{th} = 100$时，$\Pr[X_i \geqslant \mathrm{th}] \leqslant \mathrm{e}^{-36}$且$\Pr[Y_i \leqslant \mathrm{th}] \leqslant \mathrm{e}^{-64}$。

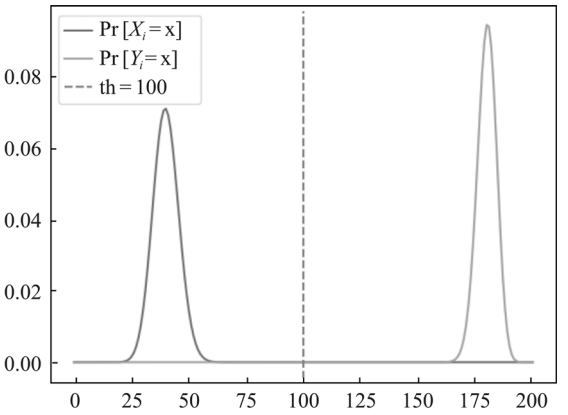

图5　$X_i \sim H(n-1, \gamma n, n_e)$和$Y_i \sim H(n-1, (1-\delta)n, n_e)$的概率质量函数

在明确了超几何分布与随机图的关系后，事件$E_1(C, G, \mathrm{th}) = 1$发生的概率可以描述为

$$\Pr(E_1(C, G, \mathrm{th}) = 1) = 1 - \Pr[E_1(C, G, \mathrm{th}) = 0] \geqslant 1 - n\Pr[X_i \geqslant \mathrm{th}] \quad (7)$$

类似地，事件$E_3(D, G, \mathrm{th}) = 1$发生的概率可以描述为

$$\Pr(E_3(D, G, \mathrm{th}) = 1) = 1 - \Pr(E_3(D, G, \mathrm{th}) = 0) \geqslant 1 - n\Pr[Y_i \leqslant \mathrm{th} - 1] \quad (8)$$

事件$E_2(C, D, G) = 1$与哈拉里图的性质相关。在哈拉里图中，每个顶点都与n_e^S个其他顶点相连，即图的连通度为n_e^S。Bell等人[46]分析了一种简单的情况。令n_e^S为偶数，然后哈拉里图在构造时，首先将所有顶点排成一个圆，然后每个节点连接其左右相邻的$n_e^S/2$个邻居。现在假设这样构造的一个哈拉里图通过删除顶点分成了两个子图，其中一个子图中有节点1，另一个子图中的顶点与节点1没有连接，设这些顶点中编号最小的节点为m，则m顶点与1没有直接连接，且$m-1, \cdots, m-n_e^S/2$顶点均与1没有直接连接，否则顶点m就可以通过这些顶

点与 1 连通；然而与顶点 1 没有连接的编号最小的节点为m，因此$m-1, \cdots, m-n_e^S/2$顶点应该与顶点 1 连通。这种矛盾说明在分成的两个子图中，$m-1,\cdots,m-n_e^S/2$顶点被删除了。注意到在对顶点随机置换时，并不改变这种顶点间的连接关系。这种观察可以得出结论，当事件$E_2(C,D,G)=0$发生时，至少有$n_e^S/2$个顶点属于腐化和掉线客户端集合，并且由于这些顶点在置换之前具有特定的关系，并不是随机选择的，因此对于任意一个顶点，该事件发生的概率为$(\delta+\gamma)^{\frac{n_e^S}{2}}$。进而，事件$E_2(C,D,G)=1$发生的概率可以描述为

$$\Pr(E_2(C,D,G)=1)=1-\Pr(E_2(C,D,G)=0)\geqslant 1-n(\delta+\gamma)^{\frac{k}{2}} \tag{9}$$

根据“好随机图”的定义，条件 1 可以等价地描述为

$$\Pr[E_1(C,G,\mathrm{th})=0]+\Pr(E_2(C,D,G)=0)=n(\Pr[X_i\geqslant \mathrm{th}]+(\delta+\gamma)^{\frac{n_e^S}{2}})\leqslant 2^{-\sigma} \tag{10}$$

条件 2 可以等价地描述为

$$n\Pr[Y_i\leqslant \mathrm{th}-1]\leqslant 2^{-\eta} \tag{11}$$

代入X_i和Y_i的边缘概率分布之后，结合哈拉里图的简单构造方法，可以得到定理 5-1。例如，该定理中max函数的第二项对应的是事件$E_3(D,G,\mathrm{th})=1$时n_e^S参数的取值。

2．协议的安全性及增强版本

在表 8 所示的随机图安全聚合协议中，随机图由服务器单独生成，这在考虑主动攻击者时是行不通的。腐化的服务器只需要设置攻击目标的邻居大部分为腐化客户端，即可获取攻击目标的模型参数。因此，在安全性增强的安全聚合协议中，首先要解决的问题就是安全地生成随机图。为了达到该目标，Bell 等人[46]给出了表 9 所示的有向随机图生成方法。

表 9 中的有向随机图协议采用了不常用的攻击者假设，即协议第 1 轮假设攻击者是半诚实的，第 2 轮开始服务器才是主动攻击者。这样设计的主要原因是避免采用 PKI 基础设施，简化协议设计。另外，有向随机图G中的一个顶点u具有顶点集合out^u和in^u，其中out^u表示所有以 u 为起点的边的终点集合，in^u表示所有以u为终点的边的集合。最后，在验证步骤对公钥数量小于$4n_e^S$的检查源于随机图的特性。按照表 9 协议，每个客户端u选择n_e^S个邻居，因而自身被选为其他节点邻居的期望也是n_e^S，实际采样的随机图中$\left|\mathrm{in}_{1_u}^S\right|>3n_e^S$的概率可以忽略，因此将服务器返回公钥数量为$3n_e^S+n_e^S$的情况视为异常。

需要指出的是，表 9 协议的描述与 Bell 等人[46]的描述并不完全相同，主要是细化了关于掉线客户端、公钥列表的传输及数量检查等。然而，这些修改并不影响有向随机图的生成过程。

表 9　有向随机图生成方法

输入：服务器S的初始用户集合U_0^S；随机图的连通度n_e^S
输出：服务器S输出有向随机图，客户端输出随机图中邻居节点的临时公钥
1. **公钥承诺**：客户端k的操作与表 2 协议相同； 假设服务器S在此步骤是半诚实的。服务器S首先按照表 8 协议构建U_1^S客户端集合，然后计算$pk_1^S = \{g^{c^u}\}_{u\in U_1^S}$和$pk_2^S = \{g^{s^u}\}_{u\in U_1^S}$的默克尔树根，并将其作为承诺值与$U_1^S$一起发送给$U_1^S$中的客户端
2. **承诺打开**：客户端k从U_1^S中随机选择n_e^S个邻居节点，形成集合out_1^k，发送给服务器S； 假设服务器S从本步骤开始是主动攻击者。服务器S首先收集至少$(1-\delta)n_1^S$个msg_2，否则超时退出；服务器S构造集合U_2^S，包含所接收消息的发送者身份，规模为$\lvert U_2^S\rvert = n_2^S$；服务器$S$根据接收的消息$\{\text{out}_1^u\}_{u\in U_2^S}$构造一个有向图 G，图中顶点在U_1^S中，在边(u,v)中，$v\in \text{out}_1^u$；对每个客户端$u\in U_2^S$，服务器S图G中导出顶点集合$\text{in}_{1_u}^S$，对任意$v\in \text{in}_{1_u}^S$，存在边(v,u)；构造消息 $$\text{msg}_{2_u}^S = (u, \{(k, g^{c^k}, g^{s^k}, \text{aux}^k)\}_{k\in U_2^S\cap \text{in}_{1_u}^S}, \{(k, g^{c^k}, g^{s^k}, \text{aux}^k)\}_{k\in U_2^S\cap \text{out}_1^u})$$ 其中，aux^k包含给客户端k的默克尔树验证信息；服务器S发送消息$\text{msg}_{2_u}^S$给客户端u；服务器本地输出随机图G
3. **验证**：客户端k检查服务器返回的公钥数量小于$4n_e^S$，否则退出协议；之后验证这些公钥在步骤 1 承诺的默克尔树上；全部验证通过，客户端k根据消息$\text{msg}_{2_u}^S$的第二部分确定in_2^k集合；根据消息$\text{msg}_{2_u}^S$的第三部分确定$\text{out}_2^k \subseteq \text{out}_1^k$集合，检查$\lvert\text{out}_2^k\rvert \geqslant \text{th}$；并本地输出$\text{in}_2^k$和$\text{out}_2^k$集合公钥列表

与随机图的安全定义类似，好的有向随机图由定义 5-6~定义 5-8 描述[46]。

定义 5-6（有向图中的有限腐化邻居事件）　令C为被腐化的客户端集合，$C\subseteq U_1^S$，$\lvert U_1^S\rvert = n$，$\lvert C\rvert \leqslant \gamma n$；令$n_e^S$为有向图中每个顶点的出度，th为门限。事件$E_4(C,G,n_e^S,\text{th}) = 1$定义为：对于所有的客户端$u\in U_1^S\backslash C$，都有$\lvert\text{out}_u\cap C\rvert < 2\text{th} - n_e^S$。

定义 5-7（无小连通图事件）　令C为被腐化的客户端的集合，O为诚实客户端的集合的任意子集。事件$E_5(C,G,\text{th},\alpha) = 1$定义为存在$\lvert O\rvert < \alpha n,\ u\in O$，且$\lvert\text{out}_u\cap(C\cup O)\rvert < \text{th}$。

定义 5-8（有向图中的客户掉线重构事件）　令D为掉线客户端的集合，$D\subseteq U_0^S$，$\lvert U_0^S\rvert = n$，$\lvert D\rvert \leqslant \delta n$。事件$E_6(D,G,\text{th}) = 1$定义为对于所有的客户端$u\in U_0^S$，都有$\lvert\text{out}_u\cap(U_0^S\backslash D)\rvert \geqslant \text{th}$。

可以看到，事件E_6是事件E_3的自然平推。事件E_4是事件E_1在主动攻击下的自然要求。在主动攻击下，假设$\lvert\text{out}_u\cap C\rvert = n_c$，则服务器可获得关于$u$的秘密份额的总数为$(n_e^S - n_c) + 2n_c$，该值当然应该小于$2\text{th}$，否则客户端$u$的两个掩码就都被去掉了。事件$E_5$主要是考虑主动攻击者生成的随机图中存在小连通子图，使得主动攻击者能够恢复每个子图的模型参数之和的问

题。当事件$E_5(C,D,\text{th},\alpha)=1$时，腐化的服务器计算的模型参数之和至少包含$\alpha n$个客户端的输入。

好有向随机图定义[46]如下。

定义 5-9（好有向随机图） 令$\mathcal{D}$是数据对(G,th)上的一个分布，$\alpha\in[0,1]$。称$\mathcal{D}$分布是(σ,η,C,α)好的，如果对所有掉线客户端的集合和腐化客户端集合，$D,C\subseteq U_1^S$，$|U_1^S|=n$，$|D|\leqslant\delta n$，$|C|\leqslant\gamma n$，满足：

（1）$\Pr[E_4(C,G',\text{th}')\wedge E_5(C,G',\text{th}',\alpha)=1|(G',\text{th}')\leftarrow\mathcal{D}]>1-2^{-\sigma}$；

（2）$\Pr[E_6(D,G',\text{th}')=1|(G',\text{th}')\leftarrow\mathcal{D}]>1-2^{-\eta-1}$。

Bell 等人[46]证明表 9 给出的有向随机图生成算法可以生成好的有向随机图，我们摘录其引理 4.7，编号为引理 5-1。

引理 5-1 设$\gamma,\delta\geqslant 0$，$\gamma+2\delta<1$，c为常数，$c(1+\ln(n)+\eta+\sigma)\leqslant n_e^S<\frac{n-1}{4}$，$\text{th}=\lceil\frac{(3+\gamma-2\delta)n_e^S}{4}\rceil$，$\alpha=\frac{1-\gamma-2\delta}{12}$，对于足够大的$n$，表 9 是有向随机图生成算法，可以形成关于对$(G,\text{th})$上的一个分布$\mathcal{D}$，该分布是$(\sigma,\eta,C,\alpha)$好的。

类似定理 5-1 的分析，引理 5-1 也是根据好有向随机图的定义和 3 个事件的定义结合超几何分布的边缘概率推导出来的。

在表 9 有向随机图生成算法的基础上，表 10 给出了增强的随机图安全聚合协议。图 6 是该协议的通信模型。

表 10 协议的设计思路从图 6 可以清楚地看到，表 10 的协议基本上是两个协议拼合而成的，一个是表 9 中的随机图生成算法，一个是安全聚合协议。其中，安全聚合协议在第 4 轮和第 5 轮嵌入了两个数字签名。第 4 轮的数字签名用于客户端u向其选择的邻居节点$v\in\text{out}_u$陈述共享密钥$sk^{u,v}$用于客户端u的掩码计算中；言外之意是客户端u告诉客户端v，u将v作为自己的邻居之一，放在了掩码计算中，从而v可以像表 2 的协议一样，向服务器提供u的秘密份额。第 5 轮的数字签名用于客户端v向选择其作为邻居的客户端u返回一个确认；该确认的言外之意是客户端v告诉客户端u，v知道u在U_4^S中，从而不会把u关于s^u的秘密份额提供出去。

Bell 等人[46]描述了一种简单攻击。在表 8 协议的第 3 轮收集密文时，腐化的服务器针对特定客户端u制定以下攻击策略。

（1）设u的邻居节点$v\in N_G(u)$，v的邻居节点$k\in N_G(v)$，服务器设置$\text{msg}_{3_k}^S=(U_3^S\backslash\{v\})$，从而获得$u$的邻居节点的秘密$s^v$，进而可以计算$u$与其邻居的共享密钥。

（2）对于u的邻居节点$v\in N_G(u)$，服务器设置$\text{msg}_{3_k}^S=(u\in U_3^S)$，从而获得$u$的秘密$b^u$。

表 10　增强的随机图安全聚合协议

输入：表 2 输入，表 4 输入，表 9 输入
输出：表 2 输出
1. **公钥承诺**：客户端k及服务器S按照表 9 协议操作
2. **承诺打开**：客户端k及服务器S按照表 9 协议操作，客户端确定out_1^k集合，服务器确定随机图G
3. **验证及分享**：客户端k按照表 9 协议验证承诺，确定in_2^k集合和out_2^k集合。之后按照表 2 分享秘密，不同之处在于仅给out_2^k集合的客户端计算和传输秘密份额； 服务器S按照表 7 协议操作，将属于某个客户端u的秘密份额通过消息$\text{msg}_{3_u}^S$打包发送，同时在消息$\text{msg}_{3_u}^S$中包含客户端u当前在线的out_1^u客户端集合列表$\text{out}_{3_u}^S = U_3^S \cap \text{out}_1^u$，其中$U_3^S$为此轮服务器收到的客户端身份标识
4. **收集密文**：客户端k根据收到的消息中密文的发送者构造集合in_3^k，确认$\text{in}_3^k \subseteq \text{in}_2^k$，提取消息的$\text{out}_{3_k}^S$，构造集合$\text{out}_3^k$，确认$\text{out}_3^k \subseteq \text{out}_2^k$，设置$U_3^k = \text{in}_3^k \cup \text{out}_3^k$，检查$\lvert\text{out}_3^k\rvert \geqslant t\text{h}$，否则退出；客户端$k$解密消息$\text{msg}_{3_k}^S$中的密文，若解密失败则退出；按照表 2 协议计算$\gamma^k$；对每个邻居$u \in \text{out}_3^k$，客户端$k$计算数字签名$\sigma_{0_u}^k = \text{Sig.sign}(\sigma k_k, k, u, \text{"include"})$，发送$\text{msg}_4^k = (k, \gamma^k, \{\sigma_{0_u}^k\}_{u \in \text{out}_3^k})$给服务器$S$，并进入下一轮； 服务器$S$收集至少$(1-\delta)n_1^S$个$\text{msg}_4$，否则超时退出；服务器$S$构造集合$U_4^S$，包含所接收消息的发送者身份，规模为$\lvert U_4^S\rvert = n_4^S$。对每个客户端$u \in U_4^S$，构造消息$\text{msg}_{4_u}^S = (U_4^S \cap (\text{out}_1^u \cup \text{in}_{1_u}^S), \{k, \sigma_{0_u}^k\}_{k \in U_4^S \cap \text{in}_{1_k}^S})$，发送给该客户端
5. **一致性检查**：客户端k将收到的$\text{msg}_{4_k}^S$中包含的不同身份的客户端设置为U_4^k集合，分别设置$\text{in}_4^k = U_4^k \cap \text{in}_3^k$和$\text{out}_4^k = U_4^k \cap \text{out}_3^k$，确认$\text{out}_4^k$集合成员数量至少为$t\text{h}$，否则退出；客户端$k$验证$\text{msg}_{4_k}^S$中的数字签名$\{\sigma_{0_k}^u\}_{u \in \text{in}_4^k}$，若验证失败则退出；对每个邻居$u \in \text{in}_4^k$，客户端$k$生成数字签名$\sigma_{1_u}^k = \text{Sig.sign}(\sigma k_k, k, u, \text{"ack"})$；形成并发送消息$\text{msg}_5^k = (k, \{\sigma_{1_u}^k\}_{u \in \text{in}_4^k})$给服务器$S$，然后进入下一轮； 服务器$S$接收至少$(1-\delta)n_1^S$个$\text{msg}_5$，形成$U_5^S$集合，否则超时退出；对每个客户端$u \in U_5^S$，构造消息$\text{msg}_{5_u}^S = (U_5^S \cap (\text{out}_1^u \cup \text{in}_{1_u}^S), \{k, \sigma_{1_u}^k\}_{k \in U_5^S \cap \text{out}_1^k})$，发送给该客户端
6. **聚合解密**：客户端k将收到的$\text{msg}_{5_k}^S$中包含的不同身份的客户端设置为U_5^k集合，分别设置$\text{in}_5^k = U_5^k \cap \text{in}_4^k$和$\text{out}_5^k = U_5^k \cap \text{out}_4^k$，确认$\text{out}_5^k$集合成员数量至少为$t\text{h}$，否则退出；客户端$k$验证$\text{msg}_{5_k}^S$中的数字签名$\{\sigma_{1_k}^u\}_{u \in \text{out}_5^k}$，若验证失败或成功验证的数字签名的数量不足$n_\sigma$则退出；否则客户端$k$构造消息$\text{msg}_6^k = \{(u, bs^u)\}_{u \in \text{out}_3^k}$，其中如果$u \notin \text{out}_4^k$，$bs^u = s_k^u$，否则$bs^u = b_k^u$；客户端$k$给服务器$S$发送消息$\text{msg}_6^k$，之后退出协议； 服务器$S$收集至少$(1-\delta)n_1^S$个$\text{msg}_6$，否则超时退出；服务器$S$构造集合$U_6^S$，包含所接收消息的发送者身份，规模为$\lvert U_6^S\rvert = n_6^S$。类似表 4，聚合解密得到$\omega = \sum_{k \in U_4^S} \omega^k = \sum_{k \in U_4^S} (\gamma^k - \text{mask}_2^u - \text{mask}_3^u)$，注意，此时掉线用户集合为$U_3^S \backslash U_4^S$

Bell 等人[46]的简单攻击对于采用表 9 方法生成的随机图同样适用。事实上，第二个攻击策略属于正常操作；而第一个攻击策略与随机图的生成方式无关。从协议设计上看，表 10 中第 5 轮的数字签名是抵御 Bell 等人[46]描述的简单攻击的一种方法，即诚实客户端v在分享秘密份额时需要先接收到足够多（n_σ个）邻居节点的确认签名，上述攻击策略（1）的情况不会发生。

第 4 轮的数字签名主要是防止腐化的服务器给客户端u特定的$\mathrm{out}_{3_u}^S$集合来获利。例如，设置$\mathrm{out}_{3_u}^S=\emptyset$，同时假设客户端$u$此时依旧会参与协议，仅使用$b^u$来做一次掩码保护（尽管$U_3^u$集合包含选择 u 的客户端集合，但是从信任关系上看，u所信任的是自己随机选择的邻居客户端，因此可以简化地看成仅做一次掩码保护）；则服务器S可以通过其实依旧在线的$\mathrm{out}_{3_u}^S$获得b^u的秘密份额，进而获得u的模型参数。第 4 轮数字签名可以阻止在线的$\mathrm{out}_{3_u}^S$提供b^u的秘密份额。

此外，表 10 的协议修正了 Bell 等人[46]的协议，在第 4 轮中客户端应该与所有的邻居节点建立共享密钥，进而计算掩码，否则会破坏表 3 的对称性。另外，在第 4 轮中检查$|\mathrm{out}_3^k|\geqslant t\mathrm{h}$，也可以防止腐化的服务器设置$\mathrm{out}_{3_u}^S=\emptyset$的情况。

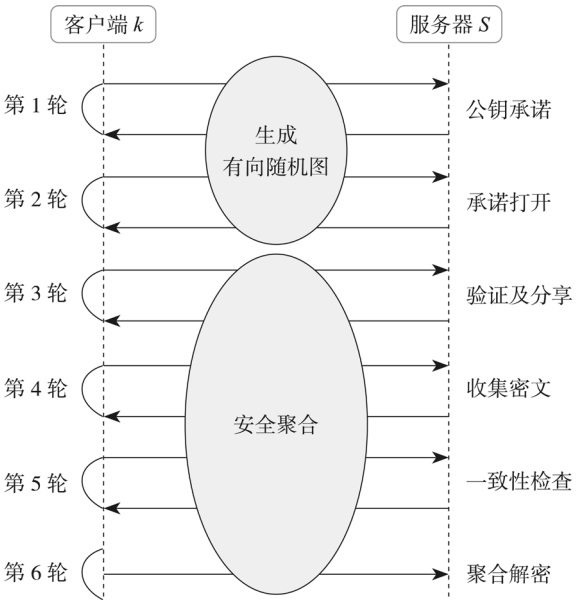

图 6　增强的随机图安全聚合协议通信模型

Bell 等人[46]给出了定理 4.9 来陈述协议的安全性，本文编号为定理 5-2。

定理 5-2　沿用引理 5-1 的符号，如果

$$n_\sigma = n_e^S - \left(\mathrm{th} - \frac{n_e^S \gamma n}{n-1} + \sqrt{\frac{n_e^S}{2}(\ln n + (\sigma+1)\ln\ 2)} \right) + 1 \qquad (12)$$

采用的对称加密算法 AE 是 IND-CPA 和 IND-CTXT 安全的，数字签名算法 Sig 是 EUF-CMA 安全的，密钥协商协议κ是安全的，那么存在一个 PPT 仿真者 Sim，对所有腐化客户端 $C \subseteq U_0^S$，$|C| \leqslant \gamma n$，对所有输入$\vec{\omega} = \{\omega^u\}_{u \in U_0^S \setminus C}$，对所有控制服务器和 C 集合客户端的主动攻击者$\mathcal{A}$，当主动攻击者$\mathcal{A}$在第 1 轮表现半诚实时，仿真者 Sim 的输出与服务器和 C 集合客户端的输出不可区分，即$\mathrm{Real}_C \approx_{\sigma,\kappa} \mathrm{Sim}^{F_{\vec{\omega},\alpha}}(C)$，其中$F_{\vec{\omega},\alpha}$是仿真者仅询问一次的理想函数，输入为诚实客户端形成的随机图 G 的一个子图顶点集合，当该集合规模大于$\alpha|U_0^S \setminus C|$时，返回这些客户端的模型参数之和，否则返回结束标志。

定理 5-2 主要是给出了邻居节点返回的确认数量的一个限制，其推导过程与随机图中每个节点的最大腐化邻居节点及掉线邻居节点的数量相关。理想函数的引入主要与随机图方法下能够推导的最大诚实用户子图规模相关。事实上，根据 Bell 等人[46]研究，可以得知当客户端规模为10^8、腐化比例为 0.2、掉线比例为 0.2 时、$\sigma = 40$、$\eta = 30$时，$\alpha = 0.32$，即此时理论上只能保证服务器输出的聚合结果至少包含 32%的诚实客户端输出。

6 展望与结论

联邦学习的隐私保护包含诸多内容，本文只是就横向联邦学习中安全聚合协议进行了探讨，设计了一些示例展示协议设计背后的原理。我们看到安全聚合协议设计精巧，效率较高，但依旧存在一些瑕疵，尤其是对抗主动攻击者的协议，还有一些可以改进的空间。

另外，当把安全聚合协议嵌入到联邦平均算法中时，就可以清楚地看到，为了实现隐私保护，需要付出较高的通信和计算代价，因此在当前工作的基础上，如何能够进一步提升效率，同时尽量减少损害模型的训练精度是值得研究的一个方向。

参考文献

[1] JORDAN M I, MITCHELL T M. Machine learning: Trends, perspectives, and prospects [J]. Science,2015, 349(6245):255-260.

[2] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-Efficient Learning of Deep Networks from Decentralized Data [DB]. arXiv preprint arXiv:1602.05629, 2016.

[3] YANG Q, LIU Y, CHEN T, et al. Federated machine learning: Concept and applications [J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2019, 10(2):1-19.

[4] YANG T, ANDREW G, EICHNER H, et al. Applied Federated Learning: Improving Google Keyboard Query Suggestions [DB]. arXiv preprint arXiv:1812.02903, 2018.

[5] LEE J, SUN J, WANG F, et al. Privacy-preserving patient similarity learning in a federated environment: development and analysis[J]. JMIR medical informatics, 2018, 6(2):e20. doi: 10.2196/medinform.7744.

[6] 郭睿, 陈涛, 刘志强. 基于航空旅客隐私数据保护的联邦学习算法应用模型研究 [J]. 信息网络安全,2020(S1):35-39.

[7] ZHU L, LIU Z, HAN S. Deep leakage from gradients [J]. Advances in Neural Information Processing Systems, 2019, 32.

[8] ZHAO B, MOPURI K R, BILEN H. iDLG: Improved Deep Leakage from Gradients[DB]. arXiv preprint arXiv:2001.02610, 2020.

[9] JONAS G, HARTMUT B, HANNAH D, et al. Inverting Gradients—How easy is it to break privacy in federated learning? [DB]. arXiv preprint arXiv:2003.14053, 2020.

[10] YIN H, MALLYA A, VAHDAT A, et al. See through Gradients: Image Batch Recovery via GradInversion [DB]. arXiv preprint arXiv:2104.07586v1,2021.

[11] ABADI M, CHU A, GOODFELLOW I, et al. Deep Learning with Differential Privacy[C]. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. October 2016:308-318.

[12] FAN M, HAMED H, KLEOMENIS K, et al. PPFL: Privacy-preserving Federated Learning with Trusted Execution Environments[C]. Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services. 2021:94-108.

[13] PHONG L T, AONO Y, HAYASHI T, et al. Privacy-preserving deep learning via additively homomorphic encryption [J]. IEEE Transactions on Information Forensics and Security. 2017, 13(5): 1333-1345.

[14] ZHOU C, FU A, YU S, et al, Privacy-preserving federated learning in fog computing [J]. IEEE Internet of Things Journal. 2020, 7(11):10782-10793.

[15] YANG W, LIU B, LU C, et al. Privacy preserving on updated parameters in federated learning [C]. Proceedings of the ACM Turing Celebration Conference-China.2020:27-31.

[16] XU D, YUAN S, WU X. Achieving differential privacy in vertically partitioned multiparty learning [C]. 2021 IEEE International Conference on Big Data. 2021:5474-5483.

[17] FANG C, GUO Y, HU Y, et al. Privacy preserving and communication-efficient federated learning in internet of things [J]. Computers & Security. 2021, 103:102199.

[18] ALEXANDRU A B, PAPPAS G J. Private weighted sum aggregation for distributed control

systems [J]. IFAC-PapersOnLine. 2020, 53(2):11081-11088.

[19] ZHOU C, FU A, YU S, et al. Privacy preserving federated learning in fog computing [J]. IEEE Internet of Things Journal. 2020, 7(11):10782-10793.

[20] ZHANG J, CHEN B, YU S, et al. PEFL: A privacy-enhanced federated learning scheme for big data analytics. 2019 IEEE Global Communications Conference. 2019:1-6.

[21] TRUEX S, BARACALDO N, ANWAR A, et al. A hybrid approach to privacy preserving federated learning [C]. Proceedings of the 12th ACM workshop on artificial intelligence and security. 2019:1-11.

[22] DAMGÅRD I, JURIK M. A GENERALISATION, a Simplification and Some Applications of Paillier's Probabilistic Public-Key System [C]. International Workshop on Public Key Cryptography PKC 2001. 2001, LNCS 1992:119-136.

[23] LIU Y, MA, LIU X, et al. Boosting privately: Federated extreme gradient boosting for mobile crowdsensing [C]. 2020 IEEE 40th International Conference on Distributed Computing Systems. 2020:1-11.

[24] LI Y, LI H, XU G, et al. Efficient privacy preserving federated learning with unreliable users [C]. IEEE Internet of Things Journal. 2021, 9(13):11590-11603.

[25] MA J, NAAS S A, SIGG S, et al. Privacy-preserving federated learning based on multi-key homomorphic encryption [J]. International Journal of Intelligent Systems. 2022, Early View.

[26] JIANG Z L, GUO H, PAN Y, et al. Secure neural network in federated learning with model aggregation under multiple keys [C]. 2021 8th IEEE International Conference on Cyber Security and Cloud Computing /2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom). 2021: 47-52.

[27] JIANG Z L, GUO H, PAN Y, et al. Secure neural network in federated learning with model aggregation under multiple keys [C]. in 2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom). 2021:47-52.

[28] BRESSON E, CATALANO D, POINTCHEVAL D. A Simple Public-Key Cryptosystem with a Double Trapdoor Decryption Mechanism and Its Applications[C]. International Conference on the Theory and Application of Cryptology and Information Security, ASIACRYPT 2003. 2003, LNCS 2894: 37-54.

[29] ZHU H, WANG R, JIN Y, et al. Distributed additive encryption and quantization for privacy preserving federated deep learning [J]. Neurocomputing. 2021, 463(6):309-327.

[30] TIAN H, ZHANG F, SHAO Y, et al. Secure linear aggregation using decentralized

threshold additive homomorphic encryption for federated learning [DB]. arXiv preprint arXiv:2111.10753, 2021.

[31] BOER D, KRAMER S. Secure sum outperforms homomorphic encryption in (current) collaborative deep learning [DB]. arXiv preprint arXiv:2006.02894, 2020.

[32] SOTTHIWAT E, ZHEN L, LI Z, et al. Partially encrypted multi-party computation for federated learning [C]. 2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing. 2021:828-835.

[33] KADHE S, RAJARAMAN N, KOYLUOGLU O O, et al. Fastsecagg: Scalable secure aggregation for privacy-preserving federated learning [DB]. arXiv preprint arXiv:2009.11248, 2020.

[34] KANAGAVELU R, LI Z, SAMSUDIN J, et al. Two-phase multi-party computation enabled privacy preserving federated learning [C]. 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing. 2020:410-419.

[35] ZHU H, GOH R S M, NG W K. Privacy-preserving weighted federated learning within the secret sharing framework [J]. IEEE Access. 2020, 8:198275-198284.

[36] XU Y, PENG C,TAN W, et al. Noninteractive verifiable privacy-preserving federated learning [J]. Future Generation Computer Systems. 2022, 128:365-380.

[37] CHEN H, LI H, XU G, et al. Achieving privacy-preserving federated learning with irrelevant updates over e-health applications [C]. 2020 IEEE International Conference on Communications. 2020: 1-6.

[38] SHI E, CHAN H, RIEFFEL E, et al. Privacy-preserving aggregation of time-series data [C]. In Annual Network & Distributed System Security Symposium. 2011.

[39] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning [C]. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017:1175-1191.

[40] BONAWITZ K, SALEHI F, KONECN J, et al. Federated learning with autotuned communication-efficient secure aggregation [C]. 53rd Asilomar Conference on Signals, Systems, and Computers. 2019:1222-1226.

[41] ELKORDY A R, AVESTIMEHR A S. Heterosag: Secure aggregation with heterogeneous quantization in federated learning [J]. IEEE Transactions on Communications. 2022, 70(4):2372-2386.

[42] ERGUN I, SAMI H U, GULER B. Sparsified secure aggregation for privacy-preserving federated learning [DB]. arXiv preprint arXiv:2112.12872, 2021.

[43] SO J, GÜLER B, AVESTIMEHR A S. Turbo-aggregate: Breaking the quadratic aggregation

barrier in secure federated learning [J]. IEEE Journal on Selected Areas in Information Theory. 2021, 2(1):479-489.

[44] JAHANI-NEZHAD T, MADDAH-ALI M A, LI S, et al. Swiftagg: Communication-efficient and dropout-resistant secure aggregation for federated learning with worst-case security guarantees [DB]. arXiv preprint arXiv:2202.04169, 2022.

[45] CHOI B, SOHN J, HAN D, et al. Communication computation efficient secure aggregation for federated learning [DB]. arXiv preprint arXiv:2012.05433, 2020.

[46] BELL J H, BONAWITZ K, GASCON A, et al. Secure single-server aggregation with (poly) logarithmic overhead [C]. Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security. 2020:1253-1269.

[47] MANDAL K, GONG G, LIU C. Nike-based fast privacy preserving high dimensional data aggregation for mobile devices [DB]. https://cacr.uwaterloo.ca/techreports/2018/cacr2018-10.pdf.

[48] JIANG Z, WANG W, LIU Y. Flashe: Additively symmetric homomorphic encryption for cross-silo federated learning [DB]. arXiv preprint arXiv:2109.00675, 2021.

[49] LIU Z, GUO J, LAM K Y, et al. Efficient dropout-resilient aggregation for privacy-preserving machine learning [J]. IEEE Transactions on Information Forensics and Security. 2022, Early Access.

[50] SO J, ALI R E, GULER B, et al. Securing secure aggregation: Mitigating multi-round privacy leakage in federated learning [DB]. arXiv preprint arXiv:2106.03328, 2021.

[51] FENG Y, YANG X, FANG W, et al. Practical and bilateral privacy-preserving federated learning [DB]. arXiv preprint arXiv:2002.09843v2, 2020.

[52] MANDAL K, GONG G. Privfl: Practical privacy-preserving federated regressions on high-dimensional data over mobile networks [C]. Proceedings of the 2019 ACM SIGSAC Conference on Cloud Computing Security Workshop. 2019:57-68.

[53] PAILLIER P. Public-key cryptosystems based on composite degree residuosity classes [C]. International conference on the theory and applications of cryptographic techniques, EUROCRYPT 1999. 1999, LNCS 1592: 223-238.

[54] BANERJEE A, FUCHSBAUER G, PEIKERT C, et al. Key-homomorphic constrained pseudorandom functions[C]. Theory of Cryptography Conference TCC 2015. 2015, LNCS 9051: 31-60.

密码分析与自动学习

孙玲[1,2]，王美琴[1,2]

1.山东大学密码技术与信息安全教育部重点实验室，济南，250100
2.山东大学网络空间安全学院，青岛，266237
通讯作者：王美琴，E-mail：mqwang@sdu.edu.cn

摘　要：经过30年的发展，对称密码算法分析理论已日渐成熟。多种多样的分析方法可对算法的安全性进行全面综合评估；便捷高效的自动化密码分析方法间接增强了密码科研人员的算法设计水平。对称密码算法设计技术和分析技术增强了学术界与工业界对于对称密码算法理论与实际安全性的信心。本文将从对称密码算法分析理论框架出发，首先介绍自动化密码分析方法研究进展，然后对近年来重回焦点的人工智能与密码分析交叉研究进行回顾与讨论。

关键词：密码分析；MILP；SAT/SMT；CP；机器学习

Cryptanalysis with Automatic Methods and Machine Learning

SUN Ling[1,2], WANG Meiqin[1,2]

1.Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education
Shandong University, Jinan, China
2.School of Cyber Science and Technology, Shandong University, Qingdao, China
Corresponding author: WANG Meiqin, E-mail: mqwang@sdu.edu.cn

Abstract: After 30 years of development, the cryptanalytic theory of symmetric-key primitives has gradually matured. A variety of cryptanalytic methods can evaluate the security of the cipher comprehensively; convenient and efficient automatic methods indirectly enhance the designing level of cryptographers. The mutually reinforcing design and cryptanalysis techniques improve the confidence of academia and industry in the theoretical and practical security of symmetric-key primitives. Following the theoretical framework of cryptanalysis, this paper first introduces the research progress of automatic methods. Then, it reviews and discusses the cross-research of artificial intelligence and cryptanalysis that has returned to the focus in recent years.

Keywords: Cryptanalysis; MILP; SAT/SMT; CP; Machine Learning

1 引言

密码作为保障网络与信息安全最有效、最可靠、最经济的关键核心技术，直接关系国家政治安全、经济安全、国防安全和信息安全。在信息化高度发展的今天，密码的应用已渗透到社会生产和生活的方方面面，从涉及政权安全的保密通信、军事指挥到波及国民经济的金融交易、防伪税控，再到触及公民权益的电子支付等，密码都在背后发挥着基础支撑作用。

密码算法包括对称密码算法和公钥密码算法。对称密码算法凭借计算量小、加密速度快、加密效率高等优势，应用更为广泛。对称密码算法的安全性通常建立在其对各种已有攻击方法抵抗性的基础上，因此设计安全高效的算法离不开丰富的密码分析经验。经过 30 多年的发展，对称密码算法的分析已渐成体系。在评估算法安全性时，首先考查算法对各种已有攻击方法的抵抗性，以勾勒安全性轮廓；其次具体算法具体分析，探索是否存在安全漏洞。对称密码算法是学术界和工业界普遍认可能经受住各种攻击方法考验的算法。

对称密码算法分析发展至今，种类繁多。密码算法只有通过各种攻击方法的检验后，才能证明它在现阶段是安全的。因此，密码研究人员需要从密码算法的自身特点出发，使用多用途编程语言（如 C++、Java、Python 等）为每种密码分析方法编写专门的程序。此外，为了提升程序的运行效率，需要结合密码算法自身特点加入定制化的优化条件，这对密码研究人员的密码分析功底和编程水平提出了极高的要求。即使是对专业的密码研究人员，完成一个算法的安全性检测可能也需要数月时间。由此可知，对称密码算法的安全性分析是一项复杂且耗时费力的工作。尤其在新算法设计阶段，大到算法结构，小到部件参数，都需要进行频繁的调整，保证高效地为调整后的算法提供可靠的安全性评估结果逐渐成为决定算法设计水平的关键。在密码分析和设计领域需求的双重驱动下，自动化密码分析方法应运而生，基于国内外研究团队的共同推动，其取得了长足进展。

自动化密码分析方法在解放密码研究人员劳动力和提升密码算法设计水平方面功不可没，但随着研究的深入，其在针对大分组、长轮数算法搜索任务中不尽如人意的表现似乎暗示了瓶颈期的到来。与此同时，人工智能在各行各业对传统计算方式的冲击使得密码研究人员开始思考，能否将人工智能中的算法用于密码算法非随机统计特征的识别与挖掘，进一步突破手动和自动化分析的局限性，从而研发更强的密码分析工具。

本文将从对称密码算法分析理论框架出发，首先介绍自动化密码分析方法研究进展，然后对近年来重回焦点的人工智能与密码分析交叉研究进行回顾和讨论。

2　自动化密码分析方法研究进展

对称密码的分析基于对密码算法中非随机统计特征的有效提取与应用。以 20 世纪 90 年代初 Eli Biham 和 Adi Shamir[1]在美密会提出差分分析方法为标志性事件，对称密码算法分析的研究渐入正轨。发展至今，已初步形成以差分分析、线性分析[2]、积分分析[9]等最具代表性的经典分析方法为基础，以在经典分析方法指导下进行延拓与综合形成的系列新型分析方法为重要组成部分的对称密码分析理论框架，如图 1 所示。

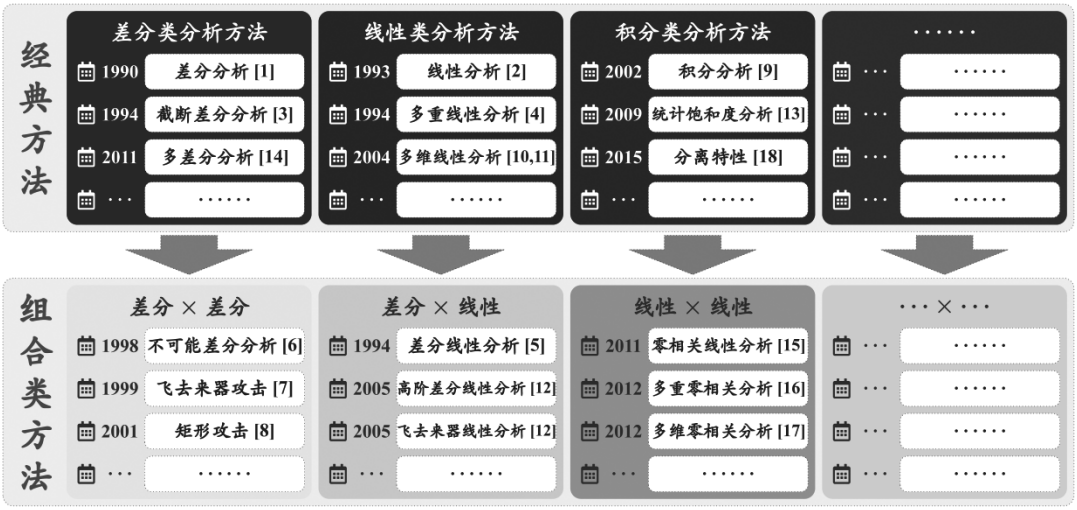

图 1　对称密码分析理论框架

对称密码算法的分析经过 30 多年发展，分析方法种类繁多，密码算法只有通过各种攻击方法的检验后，才能证明它在现阶段是安全的。为此，密码研究人员需要从密码算法的自身特点出发，使用多用途编程语言（如 C++、Java、Python 等）为每种密码分析方法编写专门的程序。此外，为了提升程序运行效率，需要结合密码算法自身特点加入定制化优化条件，这对密码研究人员的密码分析功底和编程水平提出了极高要求。即便是对专业的密码研究人员，完成一个算法的安全性检测可能也需要数月时间。由此可知，对称密码算法的安全性分析是一项复杂且耗时费力的工作。尤其在新算法设计阶段，大到算法结构，小到部件参数，都需要进行频繁调整，保证高效地为调整后算法提供可靠的安全性评估结果逐渐成为决定算法设计水平的关键。在密码分析和设计领域需求的双重驱动下，自动化密码分析方法应运而生。

自动化密码分析方法的原理（图 2）是将密码分析中烦琐复杂的问题转化为一系列可借助现成求解器求解的数学问题，求解所得相关数学问题的解又可等价变换回密码分析问题的

结果。自动化密码分析方法主要基于数学建模实现，按照所依赖的数学问题进行划分，主要有以下 3 类。

① 基于混合整数线性规划（Mixed Integer Linear Programming，MILP）的自动化密码分析方法。

② 基于布尔可满足性问题（Boolean Satisfiability Problem，SAT）或可满足性模理论（Satisfiability Modulo Theories，SMT）的自动化密码分析方法。

③ 基于约束规划（Constraint Programming，CP）的自动化密码分析方法。

自动化密码分析方法自 20 世纪初雏形初显，在国内外研究团队的共同推动下，取得了长足进展，也逐渐形成 3 条清晰的发展脉络。

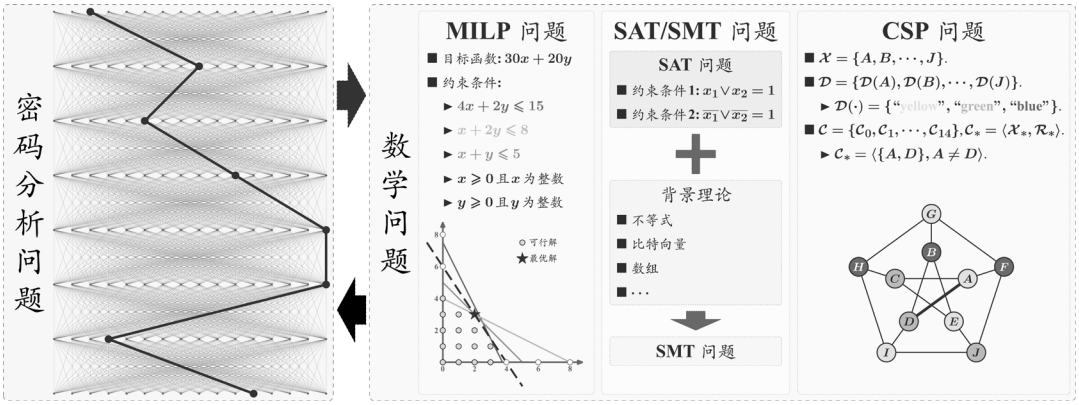

图 2　自动化密码分析方法的原理

（1）逐步实现算法密码特性更精细刻画。

首先，这一发展脉络体现在基于 MILP 的差分和线性类区分器搜索方面。如图 3 所示，基于 MILP 的自动化方法起初分别由 Nicky Mouha 等人[19]以及吴生宝和王明生[20]用于目标算法差分与线性活跃 S 盒下界的搜索，虽然该方法的搜索结果可为密码算法的安全性评估提供参考，但精度低且 MILP 问题的大量可行解无法解析为真实攻击路线。2014 年亚密会，孙思维等人[21]将计算几何学中凸集的 H-representation 理论和贪心算法引入 MILP 模型构建过程，完成了 S 盒差分性质的精简刻画。融合该模型后的 MILP 搜索框架可实现比特级分组密码算法差分路线的精确搜索，对一批算法给出更紧致的差分分析安全界。考虑到基于 MILP 的方法尚不能用于 ARX 类算法的分析，付凯等人[22]在 2016 年 FSE 会议上提出刻画差分和线性掩码在模加运算中传递的 MILP 模型，首次实现了 ARX 类算法差分和线性路线的自动化搜索。模型应用于 SPECK 算法[40]，取得当时最优分析结果。进一步地，为了解决差分路线的精确搜索框架用于具有大 S 盒算法的适用性问题，Ahmed Abdelkhalek 等人[33]在 2018 年 FSE 会议上提出一种分而治之策略，即把大 S 盒的差分分布表划分为多个相对简单的子表，为每

个子表构建 MILP 模型后，将其有机组合，从而完成大 S 盒算法差分路线的自动化搜索。

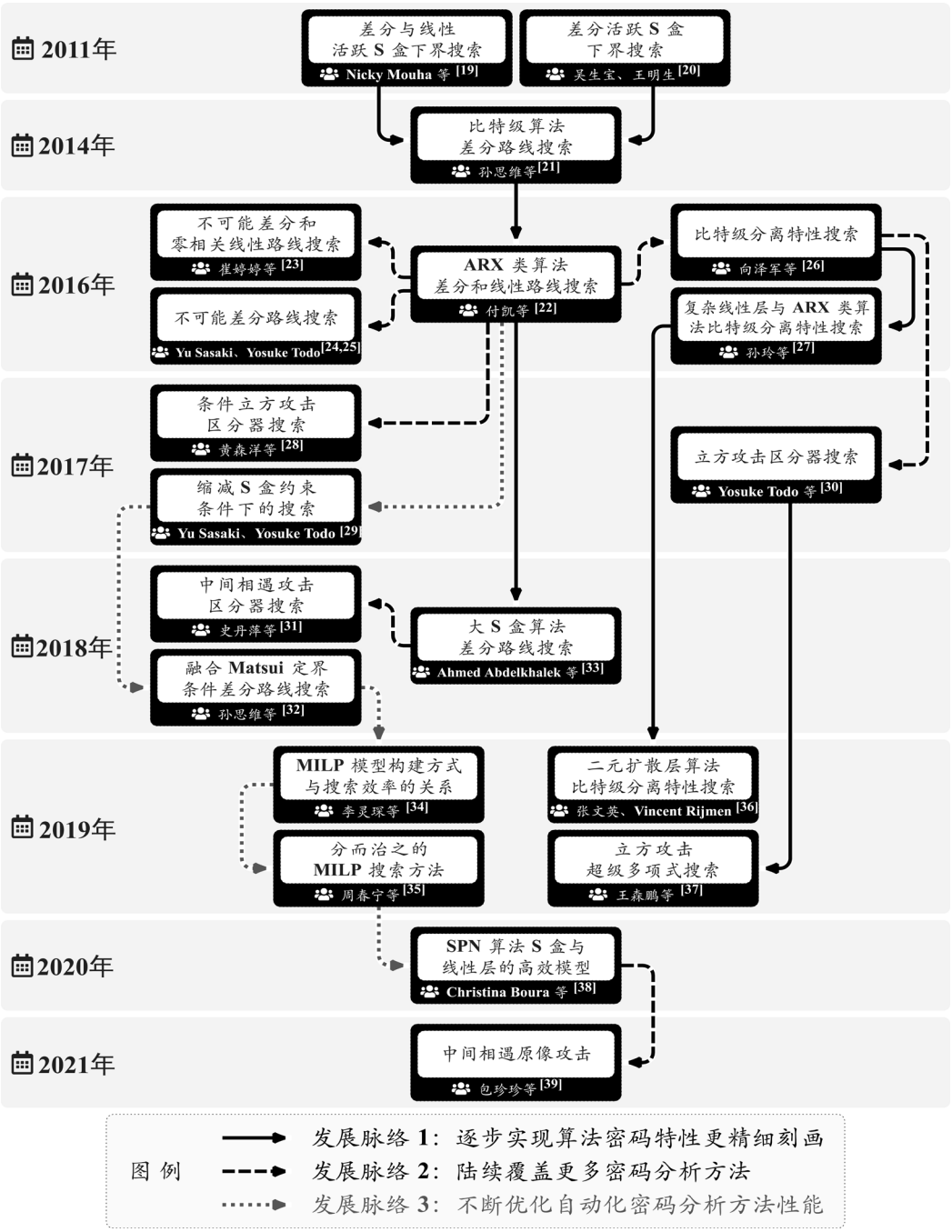

图 3　基于 MILP 的自动化密码分析方法研究进展

其次，在基于 MILP 方法的积分区分器搜索方面，精细化过程也有所体现。2016 年亚密会，向泽军等人[26]将文献[21]中的搜索框架延拓于积分分析范畴，首次给出基于比特级分离特性[41]自动化搜索积分区分器的方法。该方法被用于 6 个轻量级密码算法，算法区分器得到了不同程度的改进。注意到文献[26]中的方法无法覆盖具有复杂线性层的算法和 ARX 类算法，孙玲等人[27]随后提出了刻画比特级分离特性在复杂线性层和模加运算中传递的补充模型，融合这些模型后的 MILP 搜索方法可实现绝大多数对称密码算法积分区分器的自动化搜索。2019 年，张文英与 Vincent Rijmen[36]对比特级分离特性在二元扩散层中的传播特性建立了更加精细的 MILP 模型，使用优化后的模型改进了两个算法的积分区分器。

基于 SAT/SMT 的自动化方法出现时间稍晚于基于 MILP 的方法，但其发展过程也呈现出精细化逐渐提升的特点（图 4）。在差分和线性类区分器搜索方面，Nicky Mouha 和 Bart Preneel[42]于 2013 年首次将差分在 ARX 类算法中的传递规律刻画为 SMT 模型，通过调用 SMT 求解器实现了最优差分路线的自动化搜索。2015 年美密会，Stefan Kölbl 等人[43]将研究对象锁定为基于 AND-RX 结构的 SIMON 算法[40]，首先通过理论推导得到了计算给定差分/线性路线概率/相关度的精确公式；其次将这些公式转化为 SAT/SMT 模型，在求解器的帮助下，对 SIMON 算法的安全性进行了全面深入研究。为了拓展基于 SAT/SMT 方法的应用范围，刘韵雯等人[44]在 2016 年 ACNS 会议上构建了追踪线性掩码在 ARX 类算法中传递规律的 SAT 模型，使得 SPECK 算法和 Chaskey 算法[53]的线性路线轮数得到大幅度改进。2016 年的 ACISP 会议，宋凌等人[45]提出把由短轮路线构建长轮路线的思想与文献[42]中的模型相结合，在节约搜索时间的同时更易挖掘性质较好的路线。新方法用 SPECK 算法和 LEA 算法[54]，差分分析结果得到不同程度的改进。2018 年，孙玲等人[49]使用符合 SAT 语法的合取范式对差分在 S 盒中的传递规律建模，研发了带 S 盒算法差分闭包的自动化搜索工具，成功解决海量差分路线累积概率的计算问题。2018 年 SAC 会议，Ralph Ankele 和 Stefan Kölbl[48]构建了刻画差分在带 S 盒算法中传递的 SMT 模型，实现了一批分组密码算法大量差分路线的搜索。2019 年，刘瑜等人[50]使用 SMT 方法对（大）S 盒的差分和线性性质建模，在求解工具 STP 的辅助下，刷新了大批算法的（相关密钥）差分和线性路线。

随着研究的不断深入，基于 SAT/SMT 的搜索方法也延拓到积分分析领域，并再次展现出精细化程度逐渐提升的发展规律。2017 年亚密会，孙玲等人[46]发现在 ARX 类算法差分/线性路线的自动化搜索中，基于 SAT/SMT 的搜索模型普遍表现优于基于 MILP 的搜索模型，因而首次给出针对 ARX 类算法自动化搜索比特级分离特性的 SAT 模型，提出高效识别最优区分器的搜索策略，对系列 ARX 类算法的积分区分器进行了不同程度的改进。此外，考虑到自动化方法在搜索字级分离特性方面的空白，构建 SMT 模型，将 ISO/IEC 标准密码算法 CLEFIA 算法[55]的区分器长度拉长一轮。2018 年 SAC 会议，Zahra Eskandari 等人[47]基于 SAT 方法开发了一套搜索比特级分离特性的全自动化工具，使用者仅需学习简单的语法对算法结构进行

描述，便可灵活使用该工具。

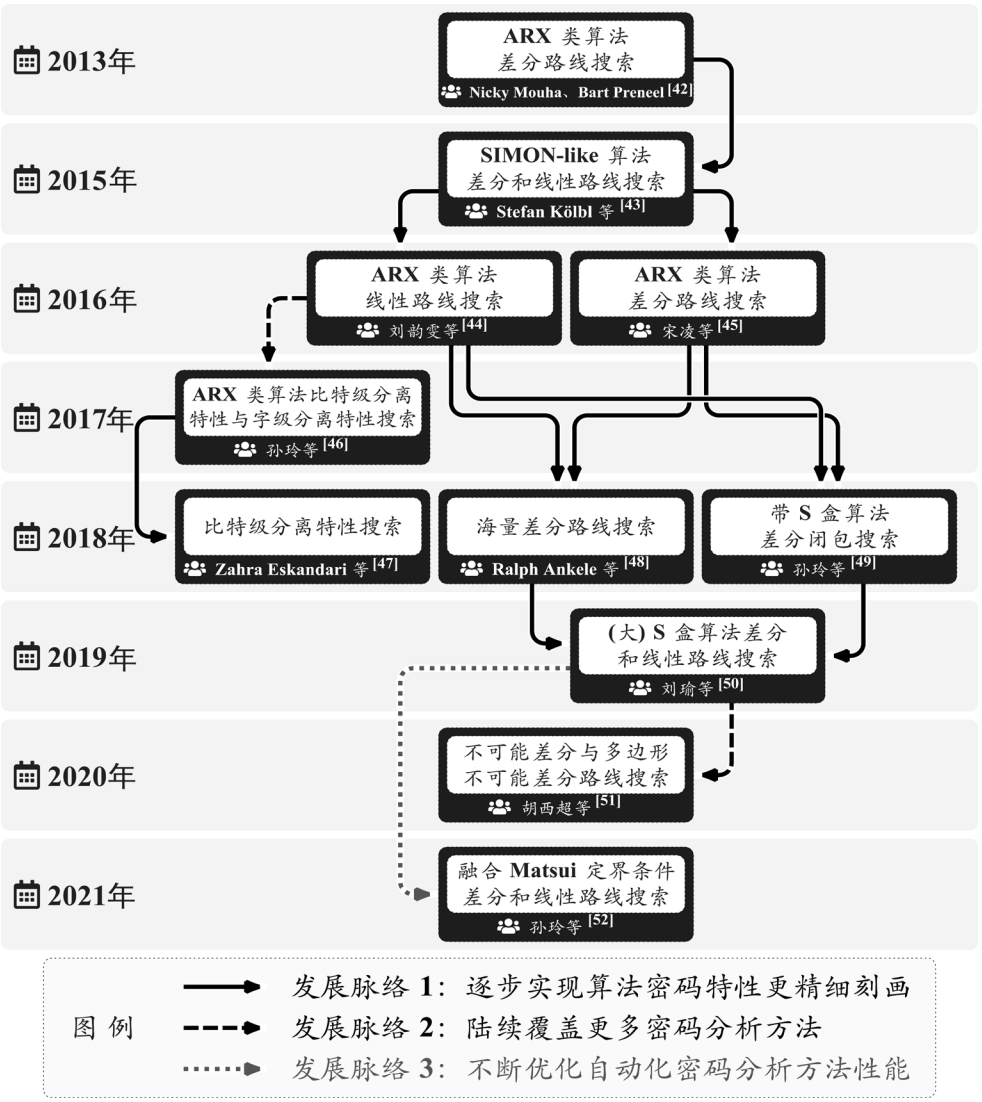

图 4　基于 SAT/SMT 的自动化密码分析方法研究进展

基于 CP 的自动化方法所依赖的底层数学问题为约束满足问题（Constraint Satisfaction Problem，CSP）。严格意义上讲，MILP 和 SAT/SMT 问题可看作 CSP 问题的特例。与 MILP 和 SAT/SMT 相比，CSP 支持更多变量类型和约束条件类别，因此建模更便捷。但值得注意的是，在自动化密码分析领域，基于 CP 的方法与前两种方法相比，出现时间更晚，且讨论度一直偏低（图 5）。基于 CP 的自动化方法最早于 2016 年由 David Gérault 等人[56]用于 AES

算法[61]的选择密钥差分攻击。2016 年印密会，David Gérault 和 Pascal Lafourcade[57]又构建了 Midori 算法[62]相关密钥差分路线 CP 搜索模型。由此可知，基于 CP 的方法也遵循逐步精细化的发展特点。

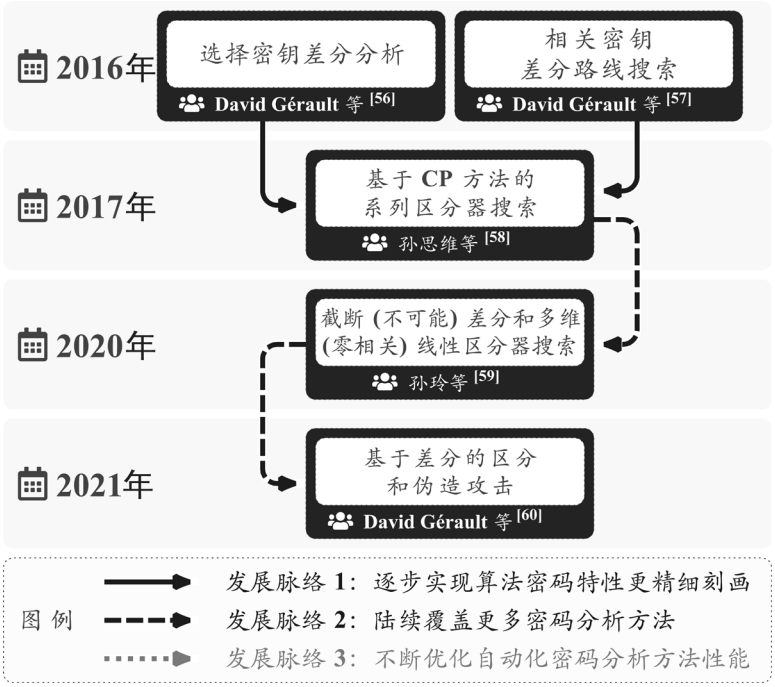

图 5　基于 CP 的自动化密码分析方法研究进展

从发展的角度来看，模型在不断更迭的过程中逐步实现对目标算法密码特性的精细刻画是必然要求，搜索得到的区分器在算法分析中的效果越来越好，有助于密码研究人员对算法安全性的精准把握。

（2）陆续覆盖更多密码分析方法。

如图 3 所示，在基于 MILP 方法的研究中，早期仅局限于经典差分与线性分析中区分器的搜索[19-22]，后来逐渐拓展到不可能差分分析[23-25]、零相关线性分析[23]、积分分析[26-27, 36]、中间相遇攻击[31,39]等领域区分器的搜索。同样，在基于 SAT/SMT 的自动化密码分析方法研究过程中（图 4），最先出现的也是与差分和线性分析相关的区分器搜索方法[42-45]。随后，基于 SAT/SMT 的方法被应用于积分分析[46-47]和不可能差分分析[51]中区分器的搜索。在基于 CP 的自动化方法研究方面（图 5），逐渐支持更多密码分析方法这一发展特点也有所体现。如今，密码研究人员仍致力于将自动化方法推广到更多的密码分析方法，进一步将算法设计人员从繁重复杂的分析任务中解放出来，使自动化方法在对称密码算法的设计与分析中发挥更

大的作用。

（3）不断优化自动化密码分析方法性能。

对自动化方法性能的优化体现于两个层面，其中之一着眼于从自动化方法依赖的底层数学问题寻求突破。初代自动化分析方法以 MILP 作为底层数学问题。基于 SAT/SMT 方法出现并不断发展的动因之一在于该方法在某类算法区分器搜索方面比 MILP 略胜一筹。基于 CP 的自动化搜索方法于 2016 年开始出现，彼时基于 MILP 和 SAT/SMT 的方法已相对成熟，而基于 CP 的方法仍风靡一时，主要原因在于其兼容更多求解器和求解策略，不同求解器和求解策略的组合可能会为自动化方法性能提升带来意想不到的效果。

除了从拓展底层数学问题角度优化自动化密码分析方法的性能，近年来，密码研究人员也开始探索在不改变底层数学问题的前提下改进搜索性能的可能性。2017 年日本学者 Yu Sasaki 和 Yosuke Todo[29]针对基于 MILP 问题差分和线性路线的搜索，给出一种最小化 S 盒模型不等式数量的新算法，但通过实验发现，最小化不等式数量并不一定有助于缩减搜索时间。2018 年 ISC 会议，孙思维等人[32]提出了融合了 Matsui 定界条件[63]的 MILP 差分路线搜索方法。2019 年，李灵琛等人[34]对 MILP 模型构建方式与搜索效率之间的关系进行了深入探索，并给出 GIFT 算法[64]分析结果的改进。2019 年 FSE 会议，周春宁等人[35]给出了分而治之的 MILP 搜索策略。2021 年 FSE 会议，孙玲等人[52]借助顺序计数器电路中的附加编码变量，提出一种针对 Matsui 算法定界条件的建模思想，给出一种基于 SAT 问题自动化搜索差分和线性路线的加速方案，使相关问题的自动化搜索效率大幅提升。

自动化密码分析方法在解放密码研究人员劳动力和提升密码算法设计水平方面功不可没。经过 10 多年的发展，针对差分分析、线性分析、积分分析等攻击方法的自动化模型已在众多算法的设计与分析中得到了广泛使用，不断更新的模型涵盖了更多的组件类型和攻击类型。然而，已有的基于数学问题求解器的自动化分析框架受现有计算能力的制约，只能处理分组长度较小、轮数较短的情况，无法应用于分组长度较大、轮数较长的算法，这成为现阶段制约自动化密码分析方法进一步应用的核心瓶颈。与此同时，人工智能在各行各业对传统计算方式的冲击激发了密码研究人员在人工智能领域寻求密码分析突破的研究热情，人工智能与密码分析交叉方向的研究重回聚光灯下。

3　人工智能时代下的密码分析

人工智能作为一门学科于 1956 年正式问世，在其漫长的研究过程中，由于硬件能力不足，经历过两次较大低谷期（图 6）。进入 21 世纪以来，得益于计算机性能的飞速提升，各式各样的机器学习算法在众多领域中焕发出了新的活力，人工智能技术取得了前所未有的高速发展。现今，人工智能几乎渗透到人们生活的方方面面，在语音、图像、自然语言处理、

人机对弈、自动驾驶、医疗健康等方面均取得了骄人成绩。虽然人工智能在侧信道攻击[65-66]方面取得了超越传统方法的成绩，但其在对称密码分析中应用的研究仍处于摸索阶段。在此背景下，各国、各共同体对基于人工智能的密码分析这一研究方向给予高度重视，对相关项目的立项工作给予重点支持。例如，欧盟“地平线 2020 计划”大幅增加人工智能领域的科研投入，大力支持新思路、新技术的早期联合科技攻关研究；新加坡新的“国家人工智能战略”将支持含基于机器学习的密码分析在内的更多利用人工智能技术的计划。

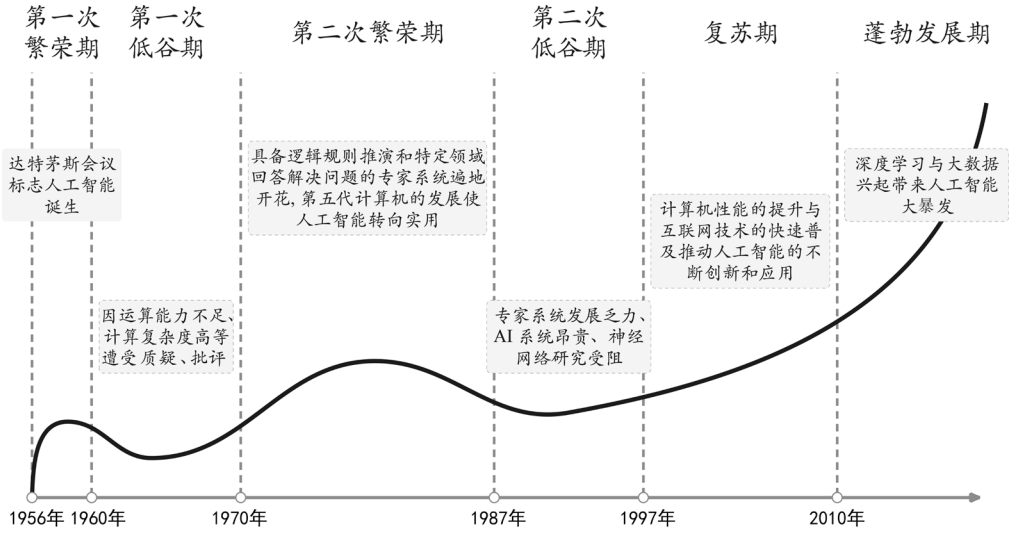

图 6　人工智能发展历程

密码学与机器学习关联性的讨论最早出现于 1991 年亚密会，Ronald L. Rivest[67]指出，攻击者在经典密码分析中通过对大量明/密文对的分析获取密钥的过程可以被描述为从输入-输出表现中学习一个未知函数的过程。从这一层面来讲，机器学习与密码分析共享许多概念和关注点，两种思想交叉融合，因此可视作姊妹研究领域。此后，部分密码研究人员[68-69]陆续进行了使用机器学习算法进行密码分析的尝试。目前使用机器学习方法进行密码分析较成功的例子出现在 2019 年美密会，Aron Gohr[70]首次尝试将深度学习技术用于轻量级分组密码算法 SPECK32/64 的分析，训练深度卷积神经网络学习算法在固定输入差分下输出差分的分布情况，将所得的分类器用作攻击的区分器，与传统方法相比更具优势。2020 年 ESORICS 会议，侯柏韬等人[71]给出了新的基于深度学习算法的线性分析框架，并将其用于缩减轮 DES 算法的攻击。2021 年欧密会，Adrien Benamira 等人[72]指出由于深度神经网络的可解释性是一项众所周知的艰巨任务，文献[70]中的攻击网络虽然为机器学习辅助密码分析开辟了新的可能性，但尚不清楚这种区分器是如何工作的，以及机器学习算法推导出的信息是什么。文献[72]

围绕这一核心问题，对文献[70]中神经网络区分器的内在工作原理进行了详细分析和彻底解释。

截至目前，基于人工智能的密码分析研究结果仍不断涌现。然而，为了避免这一研究方向成为一场虚无的狂欢，研究目标需要再度明确。若想使人工智能从根本上推动密码分析发展，密码研究人员不应只满足于人工智能算法对密码算法分析结果的改进，更重要的任务在于挖掘人工智能算法取得良好效果的背后究竟学习到密码算法的哪些特征，而这些特征又能反过来对密码研究人员设计和分析密码算法带来哪些好处。只有解决好人工智能在密码分析中的可解释性问题，才能保证人工智能和密码分析的交叉研究不会成为一座空中楼阁，进而从人工智能角度获取新型密码算法设计与分析的新方法、新思路。

4　总结

经过 30 多年的发展，对称密码算法分析理论已渐趋成熟，如多种多样的分析方法可对算法的安全性进行全面综合评估；便捷高效的自动化密码分析方法间接增强了密码科研人员的算法设计水平。相生相成的对称密码算法设计技术和分析技术增强了学术界与工业界对对称密码算法理论和实际安全性的信心。然而，现有的对称密码分析理论体系尚不完善，经典对称密码算法分析模型的推广与优化、新型密码分析模型的构建，以及自动化密码分析方法的延拓与创新等难题仍有待解决，这也必将需要一代又一代对称密码研究人员的不懈努力。

参考文献

[1] BIHAM E, SHAMIR A. Differential cryptanalysis of DES-like cryptosystems. In Advances in Cryptology - CRYPTO 1990, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings, pages 2-21, 1990.

[2] MATSUI M. Linear cryptanalysis method for DES cipher. In Advances in Cryptology - EUROCRYPT 1993, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings, pages 386-397, 1993.

[3] KNUDSEN L R. Truncated and higher order differentials. In Bart Preneel, editor, Fast Software Encryption: Second International Workshop. Leuven, Belgium, 14-16 December 1994, Proceedings, volume 1008 of Lecture Notes in Computer Science, pages 196-211. Springer, 1994.

[4] KALISKI JR B S, ROBSHAW M J B. Linear cryptanalysis using multiple approximations. In Yvo Desmedt, editor, Advances in Cryptology - CRYPTO 1994, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings, volume

839 of Lecture Notes in Computer Science, pages 26-39. Springer, 1994.

[5] LANGFORD S K, HELLMAN M E. Differential-linear cryptanalysis. In Yvo Desmedt, editor, Advances in Cryptology - CRYPTO 1994, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings, volume 839 of Lecture Notes in Computer Science, pages 17-25. Springer, 1994.

[6] KNUDSEN L R. DEAL-a 128-bit block cipher. In NIST AES Proposal, 1998.

[7] WAGNER D A. The boomerang attack. In Lars R. Knudsen, editor, Fast Software Encryption, 6th International Workshop, FSE 1999, Rome, Italy, March 24-26, 1999, Proceedings, volume 1636 of Lecture Notes in Computer Science, pages 156-170. Springer, 1999.

[8] BIHAM E, DUNKELMAN O, KELLER N. The rectangle attack - rectangling the Serpent. In Birgit Pfitzmann, editor, Advances in Cryptology - EUROCRYPT 2001, International Conference on the Theory and Application of Cryptographic Techniques, Innsbruck, Austria, May 6-10, 2001, Proceeding, volume 2045 of Lecture Notes in Computer Science, pages 340-357. Springer, 2001.

[9] KNUDSEN L R, WAGNER D A. Integral cryptanalysis. In Joan Daemen and Vincent Rijmen, editors, Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, February 4-6, 2002, Revised Papers, volume 2365 of Lecture Notes in Computer Science, pages 112-127. Springer, 2002.

[10] BAIGNÈRES T, JUNOD P, VAUDENAY S. How far can we go beyond linear cryptanalysis? In Pil Joong Lee, editor, Advances in Cryptology-ASIACRYPT 2004, 10th International Conference on the Theory and Application of Cryptology and Information Security, Jeju Island, Korea, December 5-9, 2004, Proceedings, volume 3329 of Lecture Notes in Computer Science, pages 432-450. Springer, 2004.

[11] BIRYUKOV A, CANNIÈRE C D, QUISQUATER M. On multiple linear approxi mations. In Matthew K. Franklin, editor, Advances in Cryptology-CRYPTO 2004, 24th Ann ual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 200 4, Proceedings, volume 3152 of Lecture Notes in Computer Science, pages 1-22. Springer, 2004.

[12] BIHAM E, DUNKELMAN O, KELLER N. New combined attacks on block ciphers. In Henri Gilbert and Helena Handschuh, editors, Fast Software Encryption: 12th International Workshop, FSE 2005, Paris, France, February 21-23, 2005, Revised Selected Papers, volume 3557 of Lecture Notes in Computer Science, pages 126-144. Springer, 2005.

[13] COLLARD B, STANDAERT F X. A statistical saturation attack against the block cipher PRESENT. In Marc Fischlin, editor, Topics in Cryptology - CT-RSA 2009, The Cryptographers Track

at the RSA Conference 2009, San Francisco, CA, USA, April 20-24, 2009. Proceedings, volume 5473 of Lecture Notes in Computer Science, pages 195-210. Springer, 2009.

[14] BLONDEAU C, GÉRARD B. Multiple differential cryptanalysis: Theory and practice. In Antoine Joux, editor, Fast Software Encryption-18th International Workshop, FSE 2011, Lyngby, Denmark, February 13-16, 2011, Revised Selected Papers, volume 6733 of Lecture Notes in Computer Science, pages 35-54. Springer, 2011.

[15] BOGDANOV A, RIJMEN V. Zero-correlation linear cryptanalysis of block ciphers. IACR Cryptol. ePrint Arch., 2011:123.

[16] BOGDANOV A, WANG M. Zero correlation linear cryptanalysis with reduced data complexity. In Anne Canteaut, editor, Fast Software Encryption - 19th International Workshop, FSE 2012, Washington, DC, USA, March 19-21, 2012. Revised Selected Papers, volume 7549 of Lecture Notes in Computer Science, pages 29-48. Springer, 2012.

[17] BOGDANOV A, LEANDER G, NYBERG K, et al. Integral and multidimensional linear distinguishers with correlation zero. In Xiaoyun Wang and Kazue Sako, editors, Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings, volume 7658 of Lecture Notes in Computer Science, pages 244-261. Springer, 2012.

[18] TODO Y. Structural evaluation by generalized integral property. In Elisabeth Oswald and Marc Fischlin, editors, Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I, volume 9056 of Lecture Notes in Computer Science, pages 287-314. Springer, 2015.

[19] MOUHA N, WANG Q, GU D, et al. Differential and linear cryptanalysis using mixed-integer linear programming. In Information Security and Cryptology - 7th International Conference, Inscrypt 2011, Beijing, China, November 30 - December 3, 2011. Revised Selected Papers, pages 57-76, 2011.

[20] WU S, WANG M. Security evaluation against differential cryptanalysis for block cipher structures. IACR Cryptol. ePrint Arch., 2011:551.

[21] SUN S, HU L, WANG P, et al. Automatic security evaluation and (related-key) differential characteristic search: Application to SIMON, PRESENT, LBlock, DES(L) and other bit-oriented block ciphers. In Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan,China, December 7-11, 2014. Proceedings, Part I, pages 158-178, 2014.

[22] FU K, WANG M, GUO Y, et al. MILP-based automatic search algorithms for differential and linear trails for Speck. In Fast Software Encryption - 23rd International Conference, FSE 2016, Bochum, Germany, March 20-23, 2016, Revised Selected Papers, pages 268-288, 2016.

[23] CUI T, JIA K, FU K, et al. New automatic search tool for impossible differentials and zero-correlation linear approximations. IACR Cryptol. ePrint Arch., 2016:689, 2016.

[24] SASAKI Y, TODO Y. New impossible differential search tool from design and cryptanalysis aspects. IACR Cryptol. ePrint Arch., 2016:1181.

[25] SASAKI Y, TODO Y. New impossible differential search tool from design and cryptanalysis aspects - revealing structural properties of several ciphers. In Advances in Cryptology-EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part III, pages 185-215, 2017.

[26] XIANG Z, ZHANG W, BAO Z, et al. Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers. In Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I, pages 648-678, 2016.

[27] SUN L, WANG W, WANG M. MILP-aided bit-based division property for primitives with non-bit-permutation linear layers. IACR Cryptol. ePrint Arch., 2016:811.

[28] HUANG S, WANG X, XU G, et al. Conditional cube attack on reduced-round Keccak sponge function. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II, volume 10211 of Lecture Notes in Computer Science, pages 259-288, 2017.

[29] SASAKI Y, TODO Y. New algorithm for modeling S-box in MILP based differential and division trail search. In Innovative Security Solutions for Information Technology and Communications-10th International Conference, SecITC 2017, Bucharest, Romania, June 8-9, 2017, Revised Selected Papers, pages 150-165, 2017.

[30] TODO Y, ISOBE T, HAO Y, et al. Cube attacks on non-blackbox polynomials based on division property. In Advances in Cryptology-CRYPTO 2017-37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part III, pages 250-279, 2017.

[31] SHI D, SUN S, DERBEZ P, et al. Programming the Demirci-Selçuk meet-in-the-middle

attack with constraints. In Thomas Peyrin and Steven D. Galbraith, editors, Advances in Cryptology - ASIACRYPT 2018-24th International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, December 2-6, 2018, Proceedings, Part II, volume 11273 of Lecture Notes in Computer Science, pages 3-34. Springer, 2018.

[32] ZHANG Y, SUN S, CAI J, et al. Speeding up MILP aided differential characteristic search with Matsui's strategy. In Information Security-21st International Conference, ISC 2018, Guildford, UK, September 9-12, 2018, Proceedings, pages 101-115, 2018.

[33] ABDELKHALEK A, SASAKI Y, Todo Y, et al. Youssef. MILP modeling for (large) S-boxes to optimize probability of differential characteristics. IACR Trans. Symmetric Cryptol., 2017(4):99-129, 2017.

[34] LI L, WU W, ZHENG Y, et al. The relationship between the construction and solution of the MILP models and applications. IACR Cryptol. ePrint Arch, 2019:49.

[35] ZHOU C, ZHANG W, DING T, et al. Improving the MILP-based security evaluation algorithm against differential/linear cryptanalysis using a divide-and-conquer approach. IACR Trans. Symmetric Cryptol., 2019(4):438-469.

[36] ZHANG W, RIJMEN V. Division cryptanalysis of block ciphers with a binary diffusion layer. IET Inf. Secur., 2019, 13(2):87-95.

[37] WANG S, HU B, GUAN J, et al. MILP-aided method of searching division property using three subsets and applications. In Steven D. Galbraith and Shiho Moriai, editors, Advances in Cryptology - ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security, Kobe, Japan, December 8-12, 2019, Proceedings, Part III, volume 11923 of Lecture Notes in Computer Science, pages 398-427. Springer, 2019.

[38] BOURA C, COGGIA D. Efficient MILP modelings for Sboxes and linear layers of SPN ciphers. IACR Trans. Symmetric Cryptol., 2020(3):327-361.

[39] BAO Z, DONG X, GUO J, et al. Automatic search of meet-in-the-middle preimage attacks on AES-like hashing. In Anne Canteaut and François-Xavier Standaert, editors, Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part I, volume 12696 of Lecture Notes in Computer Science, pages 771-804. Springer, 2021.

[40] BEAULIEU R, SHORS D, SMITH J, et al. The SIMON and SPECK families of lightweight block ciphers. IACR Cryptol. ePrint Arch., 2013:404.

[41] TODO Y, MORII M. Bit-based division property and application to SIMON family. In Thomas Peyrin, editor, Fast Software Encryption - 23rd International Conference, FSE 2016,

Bochum, Germany, March 20-23, 2016, Revised Selected Papers, volume 9783 of Lecture Notes in Computer Science, pages 357-377. Springer, 2016.

[42] MOUHA N, PRENEEL B. Towards finding optimal differential characteristics for ARX: Application to Salsa20. Technical report, Cryptology ePrint Archive, Report 2013/328, 2013.

[43] KÖLBL S, LEANDER G, TIESSEN T. Observations on the SIMON block cipher family. In Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I, pages 161-185, 2015.

[44] LIU Y, WANG Q, RIJMEN V. Automatic search of linear trails in ARX with applications to SPECK and Chaskey. In Applied Cryptography and Network Security-14th International Conference, ACNS 2016, Guildford, UK, June 19-22, 2016. Proceedings, pages 485-499, 2016.

[45] SONG L, HUANG Z, YANG Q. Automatic differential analysis of ARX block ciphers with application to SPECK and LEA. In Information Security and Privacy-21st Australasian Conference, ACISP 2016, Melbourne, VIC, Australia, July 4-6, 2016, Proceedings, Part II, pages 379-394, 2016.

[46] SUN L, WANG W, WANG M. Automatic search of bit-based division property for ARX ciphers and word-based division property. In Advances in Cryptology-ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I, pages 128-157, 2017.

[47] ESKANDARI Z, KIDMOSE A B, KÖLBL S, et al. Finding integral distinguishers with ease. In Carlos Cid and Michael J. Jacobson Jr., editors, Selected Areas in Cryptography-SAC 2018-25th International Conference, Calgary, AB, Canada, August 15-17, 2018, Revised Selected Papers, volume 11349 of Lecture Notes in Computer Science, pages 115-138. Springer, 2018.

[48] ANKELE R, KÖLBL S. Mind the gap - A closer look at the security of block ciphers against differential cryptanalysis. In Carlos Cid and Michael J. Jacobson Jr., editors, Selected Areas in Cryptography - SAC 2018 - 25th International Conference, Calgary, AB, Canada, August 15-17, 2018, Revised Selected Papers, volume 11349 of Lecture Notes in Computer Science, pages 163-190. Springer, 2018.

[49] SUN L, WANG W, WANG M. More accurate differential properties of LED64 and Midori64. IACR Trans. Symmetric Cryptol., 2018(3):93-123.

[50] LIU Y, LIANG H, LI M, et al. STP models of optimal differential and linear trail for S-box based ciphers. IACR Cryptol. ePrint Arch., 2019:25.

[51] HU X, LI Y, JIAO L, et al. Mind the propagation of states - new automatic search tool for impossible differentials and impossible polytopic transitions. In Shiho Moriai and Huaxiong Wang, editors, Advances in Cryptology - ASIACRYPT 2020 - 26th International Conference on the Theory

and Application of Cryptology and Information Security, Daejeon, The Republic of Korea, December 7-11, 2020, Proceedings, Part I, volume 12491 of Lecture Notes in Computer Science, pages 415-445. Springer, 2020.

[52] SUN L, WANG W, WANG M. Accelerating the search of differential and linear characteristics with the SAT method. IACR Transactions on Symmetric Cryptology, 2021(1):269-315, Mar. 2021.

[53] MOUHA N, MENNINK B, HERREWEGE A V, et al. Chaskey: An efficient MAC algorithm for 32-bit microcontrollers. In Antoine Joux and Amr M. Youssef, editors, Selected Areas in Cryptography - SAC 2014 - 21st International Conference, Montreal, QC, Canada, August 14-15, 2014, Revised Selected Papers, volume 8781 of Lecture Notes in Computer Science, pages 306-323. Springer, 2014.

[54] HONG D, LEE J K, KIM D C, et al. LEA: A 128-bit block cipher for fast encryption on common processors. In Yongdae Kim, Heejo Lee, and Adrian Perrig, editors, Information Security Applications-14th International Workshop, WISA 2013, Jeju Island, Korea, August 19-21, 2013, Revised Selected Papers, volume 8267 of Lecture Notes in Computer Science, pages 3-27. Springer, 2013.

[55] SHIRAI T, SHIBUTANI K, AKISHITA T, et al. The 128-bit blockcipher CLEFIA (extended abstract). In Alex Biryukov, editor, Fast Software Encryption, 14th International Workshop, FSE 2007, Luxembourg, Luxembourg, March 26-28, 2007, Revised Selected Papers, volume 4593 of Lecture Notes in Computer Science, pages 181-195. Springer, 2007.

[56] GÉRAULT D, MINIER M, SOLNON C. Constraint programming models for chosen key differential cryptanalysis. In Michel Rueher, editor, Principles and Practice of Constraint Programming-22nd International Conference, CP 2016, Toulouse, France, September 5-9, 2016, Proceedings, volume 9892 of Lecture Notes in Computer Science, pages 584-601. Springer, 2016.

[57] GÉRAULT D, LAFOURCADE P. Related-key cryptanalysis of Midori. In Orr Dunkelman and Somitra Kumar Sanadhya, editors, Progress in Cryptology-INDOCRYPT 2016 - 17th International Conference on Cryptology in India, Kolkata, India, December 11-14, 2016, Proceedings, volume 10095 of Lecture Notes in Computer Science, pages 287-304, 2016.

[58] SUN S, GÉRAULT D, LAFOURCADE P, et al. Analysis of AES, Skinny, and others with constraint programming. IACR Trans. Symmetric Cryptol., 2017(1):281-306.

[59] SUN L, GÉRAULT D, WANG W, et al. On the usage of deterministic (related-key) truncated differentials and multidimensional linear approximations for SPN ciphers. IACR Trans. Symmetric Cryptol., 2020(3):262-287.

[60] GÉRAULT D, PEYRIN T, TAN Q. Exploring differential-based distinguishers and forgeries for ASCON. IACR Trans. Symmetric Cryptol., 2021(3):102-136, 2021.

[61] DAEMEN J, RIJMEN V. The Design of Rijndael: AES-The Advanced Encryption Standard. Information Security and Cryptography. Springer, 2002.

[62] BANIK S, BOGDANOV A, ISOBE T, et al. Midori: A block cipher for low energy. In Tetsu Iwata and Jung Hee Cheon, editors, Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29-December 3, 2015, Proceedings, Part II, volume 9453 of Lecture Notes in Computer Science, pages 411-436. Springer, 2015.

[63] MATSUI M. On correlation between the order of S-boxes and the strength of DES. In Advances in Cryptology - EUROCRYPT 1994, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings, pages 366-375, 1994.

[64] BANIK S, PANDEY S K, PEYRIN T, et al. GIFT: A small present - towards reaching the limit of lightweight encryption. In Wieland Fischer and Naofumi Homma, editors, Cryptographic Hardware and Embedded Systems-CHES 2017-19th International Conference, Taipei, Taiwan, September 25-28, 2017, Proceedings, volume 10529 of Lecture Notes in Computer Science, pages 321-345. Springer, 2017.

[65] MAGHREBI H, PORTIGLIATTI T, PROUFF E. Breaking cryptographic implementations using deep learning techniques. In Claude Carlet, M. Anwar Hasan, and Vishal Saraswat, editors, Security, Privacy, and Applied Cryptography Engineering - 6th International Conference, SPACE 2016, Hyderabad, India, December 14-18, 2016, Proceedings, volume 10076 of Lecture Notes in Computer Science, pages 3-26. Springer, 2016.

[66] PICEK S, SAMIOTIS I P, KIM J, et al. On the performance of convolutional neural networks for side-channel analysis. In Anupam Chattopadhyay, Chester Rebeiro, and Yuval Yarom, editors, Security, Privacy, and Applied Cryptography Engineering - 8th International Conference, SPACE 2018, Kanpur, India, December 15-19, 2018, Proceedings, volume 11348 of Lecture Notes in Computer Science, pages 157-176. Springer, 2018.

[67] RIVEST R L. Cryptography and machine learning. In Hideki Imai, Ronald L. Rivest, and Tsutomu Matsumoto, editors, Advances in Cryptology-ASIACRYPT 1991, International Conference on the Theory and Applications of Cryptology, Fujiyoshida, Japan, November 11-14, 1991, Proceedings, volume 739 of Lecture Notes in Computer Science, pages 427-439. Springer, 1991.

[68] ALANI M M. Neuro-cryptanalysis of DES and triple-DES. In Tingwen Huang, Zhigang Zeng, Chuandong Li, and Chi-Sing Leung, editors, Neural Information Processing-19th International

Conference, ICONIP 2012, Doha, Qatar, November 12-15, 2012, Proceedings, Part V, volume 7667 of Lecture Notes in Computer Science, pages 637-646. Springer, 2012.

[69] LASRY G. A Methodology for the Cryptanalysis of Classical Ciphers with Search Metaheuristics. PhD thesis, University of Kassel, Germany, 2018.

[70] GOHR A. Improving attacks on round-reduced Speck32/64 using deep learning. In Alexandra Boldyreva and Daniele Micciancio, editors, Advances in Cryptology-CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2019, Proceedings, Part II, volume 11693 of Lecture Notes in Computer Science, pages 150-179. Springer, 2019.

[71] HOU B, LI Y, ZHAO H, et al. Linear attack on round-reduced DES using deep learning. In Liqun Chen, Ninghui Li, Kaitai Liang, and Steve A. Schneider, editors, Computer Security-ESORICS 2020-25th European Symposium on Research in Computer Security, ESORICS 2020, Guildford, UK, September 14-18, 2020, Proceedings, Part II, volume 12309 of Lecture Notes in Computer Science, pages 131-145. Springer, 2020.

[72] BENAMIRA A, GÉRAULT D, PEYRIN T, et al. A deeper look at machine learning-based cryptanalysis. In Anne Canteaut and François-Xavier Standaert, editors, Advances in Cryptology-EUROCRYPT 2021-40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part I, volume 12696 of Lecture Notes in Computer Science, pages 805-835. Springer, 2021.

第三部分
2021 年中国密码学会优秀博士学位论文主要成果

基于可分性的密码分析与设计方法研究

王森鹏

战略支援部队信息工程大学，郑州，450000

摘　要：在 2015 年欧密会上，积分区分器构造方法领域取得突破性进展，Todo 提出了更一般化的积分性质——可分性。随后，可分性与自动化搜索工具相结合，在积分攻击、立方攻击和代数次数估计等方面取得重要成果。由于可分性出现时间较短，相关研究成果还不够成熟，一些公开问题和理论问题还没有解决。本文对基于可分性的密码分析与设计方法进行研究，以方法研究为主体，以具体密码算法的安全性分析为实例，深入研究其理论和应用，为密码算法的分析与设计提供重要支撑。

关键词：可分性；积分攻击；立方攻击；自动化搜索

Research on the Division Property Based Analysis and Design Methods of Ciphers

WANG Senpeng

PLA SSF Information Engineering University，Zhengzhou，450000

Abstract: At EUROCRYPT 2015, a breakthrough was made in the construction of integral distinguishers. Todo proposed the concept of division property, which is a generalization of integral property. Then, combined with automatic searching tools, division property achieves important results in integral attack, cube attack, the estimation of algebraic degree and so on. Because division property is a new concept, the relevant research results are not mature enough. Some open and theoretical problems have not been solved. In this paper, we research the analysis and design methods of ciphers based on division property. With the methods research as main body and the security evaluation of specific ciphers as examples, we will deeply study the theories and applications of division property. In this way, this paper will provide important support for the analysis and design of ciphers.

Keywords: Division Property; Integral Attack; Cube Attack; Automatic Search

1 研究背景

密码学是伴随着战争而发展起来的科学，从古代到第二次世界大战结束，密码的存在主要服务于国防和军事需要，密码算法的分析与设计也没有严格的数学理论支撑。1949 年，信息论的创始人 Shannon 发表了《保密系统的通信理论》（*Communication Theory of Secrecy Systems*）[1]，为密码学奠定了坚实的理论基础。密码技术作为信息安全的核心技术，不仅可以用于信息传输与存储的加密保护，还可以实现信息的真实性认证和承诺的不可抵赖性等功能，于是密码学在商业民用领域也迎来了快速发展。

密码算法的设计和分析是密码学发展与进步的主旋律，密码算法设计者希望可以给出算法的可证明安全或针对目前已知攻击方法的安全性评估；密码算法分析者希望有好的分析方法来破译密码算法，二者在相互竞争中共同发展，所以对密码算法的分析与设计方法进行研究具有重要的意义。在 2015 年欧密会上，积分区分器构造方法领域取得突破性进展，Todo 提出了可分性（Division Property）的概念。随后，可分性与自动化搜索工具相结合，在积分攻击[2]、立方攻击[3]和代数次数估计[4]等方面取得重要成果。博士学位论文“基于可分性的密码分析与设计方法研究”围绕可分性的一些公开问题和理论问题进行研究。

2 研究现状

2.1 可分性与积分攻击

积分攻击是针对对称密码算法最有效的攻击方法之一，它是由 Daemen 等人[5]首次提出用于攻击 SQUARE 分组密码算法，所以刚开始也被称为 SQUARE 攻击。随后提出的饱和度攻击[6]和多重集攻击[7]采用的也是类似的思想。在 2002 年 FSE 会议上，Knudsen 和 Wagner[8]对此类攻击技术进行理论上的统一，称为积分攻击。积分攻击在分组密码算法的分析中被证实是十分有效的。

积分区分器具有如下的性质：通过选取特定的明文集合，在一般情况下，输入明文的某些比特固定为常量（称为常量比特），其余比特遍历全部可能取值（称为活跃比特），然后考虑输出密文的逐比特异或求和情况。若输出密文的某些比特异或求和值不恒等于 0，则称为未知比特；若某些比特的异或求和值为 0，则称为平衡比特，即存在积分区分器。当攻击者利用积分区分器进行密钥恢复攻击时，通常会在积分区分器前后各增加几轮，然后猜测涉及的密钥对密文进行解密，检测积分区分器位置的异或求和值是否为 0。如果异或求和值为 0，那么被猜测的密钥值便作为候选密钥；否则，肯定为错误密钥，通过这种方式恢复密钥信息。在进行密钥恢复攻击时，为了降低攻击的计算和存储复杂度，很多技术被提出。例如，Ferguson

等人[9]提出了部分和技术（Partial Sum Technique）；Todo 等人[10]将快速傅里叶变换（Fast Fourier Transform，FFT）应用到积分攻击中；Sasaki 和 Wang[11]将中间相遇技术（Meet in the Middle Technique）应用到 Feistel 结构密码算法的积分攻击中。

积分区分器的性质会影响到攻击的效果，传统构造积分区分器的方法主要有以下两种。

（1）使用积分性质的传播法则，如文献[8]将多重集的性质划分为以下 4 类。

• All 性质：若 $\mathbb{F}_2^n$ 中的每个元素在多重集 $\mathbb{X}$ 中出现的次数均相等且 $\mathbb{X} \neq \varnothing$，则称 $\mathbb{F}_2^n$ 上的多重集 $\mathbb{X}$ 满足 All 性质。

• Balance 性质：若多重集 $\mathbb{X}$ 中的所有元素的异或求和值等于 0，则称多重集 $\mathbb{X}$ 满足 Balance 性质。

• Constant 性质：若多重集 $\mathbb{X}$ 中的所有元素均相同，则称多重集 $\mathbb{X}$ 满足 Constant 性质。

• Unknown 性质：若多重集 $\mathbb{X}$ 不满足上述 3 类性质，则称多重集 $\mathbb{X}$ 满足 Unknown 性质。

通过考查多重集上述 4 类积分性质的传播法则，可以得到密码算法的积分区分器。但这种方法对含有非双射组件的密码算法适用性比较差。另外，该方法并没有有效利用轮函数的低代数次数性质，还有进一步改进的可能。

（2）构造积分区分器的传统方法是代数次数估计方法。根据高阶差分攻击[12]，如果分组密码算法的代数次数小于等于 d，当选择明文量为 2^{d+1} 时，存在积分区分器。Canteaut 和 Videau[13]给出迭代型轮函数的代数次数的上界值，然后文献[14]给出了更紧的界。但是，此方法对于代数次数比较高的密码算法，所需要的数据量往往非常大。

在 2015 年欧密会上，积分区分器构造方法领域取得突破性进展，日本学者 Todo[2]提出了更一般化的积分性质——可分性（Division Property）。可分性可以追踪介于 All 性质与 Balance 性质之间的积分性质，将密码学组件更多的代数信息考虑到积分区分器的构造过程中，更精确地刻画密码算法的积分特征，从而可以构造更好的积分区分器。随后，在 2015 年美密会上，Todo[15]利用可分性技术对 MISTY1 密码算法进行了分析，发现了 MISTY1 算法中 S 盒的可分性弱点，首次给出了全轮 MISTY1 算法的理论攻击结果。之后，Sun 等人[16]研究了一类集合上的可分性，为理解可分性提供了新的视角。在 2016 年美密会上，Boura 等人[17]对可分性的代数性质进行进一步的研究，给出了 S 盒的可分性刻画方法，并利用该技术成功地提高了 PRESENT 算法积分区分器的轮数。尽管基于可分性构造的积分区分器与传统积分区分器相比有一定优势，但对于基于非 S 盒构造的分组密码算法（如 SIMON），用经典可分性搜索得到的积分区分器与统计测试下获得的实验积分区分器之间仍然存在差距。于是，在 2016 年 FSE 会议上，Todo 和 Morii[18]提出了两种基于比特的可分性：二子集比特可分性（Conventional Bit-based Division Property）和三子集比特可分性（Bit-based Division Property Using Three Subsets），可以找到更精确的积分区分器。但直接穷举中间状态所有的可分性向

量时，对于分组长度为 n 比特的分组密码，刻画其比特可分性传播的计算复杂度和存储复杂度大约为 2^n，这极大地限制了其被应用到大分组规模的密码算法中。

为了解决传播复杂性的问题，在 2016 年亚密会上，Xiang 等人[19]利用混合整数线性规划（Mixed Integer Linear Programming，MILP）模型刻画了二子集比特可分性的传播性质，然后借助 MILP 求解器（如 Gurobi[20]），使得密码算法的基于二子集比特可分性的积分性质可以被自动化地搜索，极大地扩展了二子集比特可分性的应用。这种方法被应用到扩散层是比特拉线的分组密码中，取得了非常好的效果。随后，Sun 等人[21]解决了二子集可分性自动化搜索算法在 ARX 结构分组密码算法中的适用性问题。在 2017 年亚密会上，Sun 等人[22]提出了基于 SAT 方法的二子集比特可分性的自动化搜索工具和基于 SMT 方法的字级可分性的自动化搜索工具。

虽然三子集比特可分性可以更精确地刻画密码算法的积分性质，但由于三子集比特可分性的特殊性质，三子集比特可分性的传播问题无法直接转化为 MILP 问题，而在 Todo 和 Morii[18]的方法下，三子集比特可分性在分组规模为 n 比特的密码算法中传播的计算和存储复杂度大约为 2^n。在 CT-RSA 2019 会议上，Hu 等人[23]利用 STP 求解器给出了一种变型三子集比特可分性的自动化搜索方法，但该变型三子集比特可分性是原始三子集比特可分性的弱化版本。如何有效地刻画三子集比特可分性的传播仍然是一个公开问题，也是目前该领域研究的重点和难点。

2.2 可分性与立方攻击

立方攻击属于选择 IV 的密钥恢复攻击，由著名密码学家 Dinur 和 Shamir[24]于 2009 年在欧密会上首次提出。它可以被看成是高阶差分攻击[12]和选择 IV 攻击[25]的特殊形式，是对称密码算法分析领域重要的分析方法之一，对序列密码算法尤其有效。对于特定的密码算法，令 $x=(x_0,x_1,\cdots,x_{n-1})$ 为 n 个秘密变量，$v=(v_0,v_1,\cdots,v_{m-1})$ 为 m 个公开变量，那么密码算法某一输出比特可以表示为布尔函数 $f(x,v)$。立方攻击的核心思想是通过选取特定的立方集合，然后通过求取其对应的 $f(x,v)$ 的异或和来约简 $f(x,v)$ 的表达式，约简后的多项式被称为超级多项式。立方攻击的核心技术是恢复立方集合对应的超级多项式，目标是从超级多项式中恢复出密钥变量。

在立方攻击刚被提出时，攻击者将密码算法看成黑盒，通过线性检测的方法来恢复超级多项式。例如，在 2009 年欧密会上，Dinur 和 Shamir[24]提出了线性检测技术并恢复了 767 轮 Trivium 的 35 个线性超级多项式。在文献[26]中，二次项检测技术第一次被应用到针对 Trivium 的立方攻击，于是 709 轮 Trivium 的 41 个线性超级多项式和 38 个二次超级多项式被找到。在文献[27]中，Ye 等人提出了一种新的线性化技术，对于 802 轮的 Trivium，他们找到了 6 个

线性超级多项式和 2 个非线性超级多项式。上面所有的实验立方攻击均需要验证大量的立方集合去寻找合适的立方集合，当立方集合的维数超过 35 时，计算复杂度就会非常的高。所以其无法评估当立方集合维数较大时，密码算法抵抗立方攻击的能力。最近有很多变型的立方攻击被提出，如动态立方攻击[28]、条件立方攻击[29]、相关立方攻击[30]、确定立方攻击[31]以及基于二子集比特可分性的立方攻击[3][32]。

动态立方攻击的基本思想是动态的选取立方集合，然后进行密钥恢复攻击。2011 年 Dinur 和 Shamir[28]利用动态立方攻击的思想攻破了 Grain128，显示出动态立方攻击强大的分析能力。在 2015 年欧密会上，Dinur 等人[33]利用立方攻击的基本原理和分别征服法成功地攻击了减轮消息认证码 KECCAK-MAC 和认证加密算法 KEYAK。在 2017 年欧密会上，Huang 等人[29]提出了一种带条件的立方攻击方法，并成功运用到 Keccak-MAC 和 Keyak 上。在 2018 年欧密会上，Liu 等人[30]提出了相关立方攻击，对于一个立方集合，首先去找到一些低阶的超级多项式，称为 basis；然后去搜索超级多项式与这些 basis 的关系；最后可以得到一系列关于密钥变量的概率多项式。通过求解这些概率多项式，便可以恢复有关的密钥比特，该方法可以利用小规模的立方集合来恢复 835 轮 Trivium 的密钥。2018 年，Ye 等人[31]提出了一种新的立方攻击——确定立方攻击，他们的方法是基于 Liu 等人[34]在 2017 年美密会上提出的数值映射技术（Number Mapping Technology），在此基础上 Ye 等人提出一种特殊的立方集合——有用立方（Useful Cube）集合，该类型的立方集合满足输出布尔函数的每个代数项的代数次数总是小于等于立方集合维数。利用 37 维的有用立方集合，他们恢复了 838 轮 Trivium 的精确超级多项式。但是，文章的作者也指出，当立方集合的规模增加时并不能提高攻击的轮数。也就是说，确定立方攻击方法对于大维数的立方集合效果不好。

值得注意的是，在 2017 年美密会上，Todo 等人[3]不再将密码算法看作黑盒，首次将二子集比特可分性与立方攻击相结合。通过二子集比特可分性来确定哪些密钥比特是包含在超级多项式中的，然后恢复对应的超级多项式，从而提出了基于二子集比特可分性的立方攻击，并将其应用到序列密码算法 Trivium、Grain128a 及认证加密算法 ACORN 中，取得了当时最好的密钥恢复攻击结果。然后，在 2018 年美密会上，Wang 等人[32]提出了改进的基于二子集比特可分性的立方攻击，给出了 839 轮 Trivium 的密钥恢复攻击结果。对于基于二子集比特可分性的立方攻击，大规模立方集合对应的超级多项式可以被理论方法所恢复。但是二子集比特可分性理论不能保证立方集合对应的超级多项式是含有密钥的，也就是密钥恢复攻击有可能退化为区分攻击。三子集比特可分性可以研究异或求和值为 1 的积分区分器，那么如何利用三子集比特可分性给出确定的密钥恢复攻击结果，是目前密码分析领域的前沿课题。

2.3 可分性与 S 盒的安全指标

S 盒被广泛应用于分组密码算法的设计，在最常见的两种设计结构（SPN 结构和 Feistel 结构）中，S 盒是唯一的非线性组件，对密码算法的安全性起到至关重要的作用。S 盒最常见的是 8 比特 S 盒和 4 比特 S 盒。近年来，由于存储设备和通信终端轻量化与移动化的发展趋势，众多的轻量级对称密码算法被提出。由于 4 比特 S 盒比 8 比特 S 盒在硬件实现上具有更大的优势，因此很多轻量级密码算法均采用 4 比特 S 盒

S 盒的选取主要考虑两个方面的标准：软硬件实现效率和安全性。但在很多情况下，这两个标准是相互冲突的，必须根据密码算法设计者的设计策略进行平衡折中。对 S 盒的安全性要求依赖于设计策略，且每个安全性要求都是为了抵抗一种或几种攻击方法。由于差分攻击[35]和线性攻击[36]的强大攻击能力，S 盒优良的差分和线性性质是最基本的安全性要求，如差分均匀度 $Diff(S)$和线性度 $Lin(S)$就是 S 盒必须考虑的安全指标。特别地，如果 4 比特 S 盒可以达到差分均匀度 $Diff(S)=4$ 和线性度 $Lin(S)=8$，则称其为 Optimal S 盒。

利用 S 盒的等价关系将 S 盒划分成不同的等价类可以更好地研究 S 盒的安全特征，如在文献[37]中，Leander 和 Poschmann 将所有的 Optimal S 盒划分为 16 个不同的仿射等价类，这个结果可以被用于有效生成满足其他安全性要求的 Optimal S 盒。对于很多的结构，设计策略仅要求是 Optimal S 盒是不够的，还有其他的一些安全要求需要考虑。例如，在密码算法 Serpent [38]的 S 盒设计中就要求 1 比特输入差分应当至少产生 2 比特输出差分，这种类型的 Optimal S 盒被称为 Serpent 类型 S 盒，随后文献[37]将所有的 Serpent 类型 S 盒划分成 20 个不同的置换异或等价类。在 SAC 2011 会议上，Saarinen [39]将分支数和其他代数性质考虑到 S 盒的设计中，提出了 Golden 类型的 S 盒，并指出所有的 Golden 类型 S 盒可以被划分成 4 个不同的置换异或等价类。

当新的攻击方法被提出时，新的安全指标也应当被加入 S 盒的设计标准中。可分性是 Todo[2]在 2015 年欧密会上提出的一般化的积分性质。在 2015 年美密会上，Todo[15]研究了 MISTY1 密码算法中 S 盒的可分性弱点，首次给出了全轮 MISTY1 算法的分析结果，所以研究 S 盒针对可分性的安全指标具有重要的意义。在 2016 年美密会上，Boura 和 Canteaut[17]研究了 S 盒针对可分性的安全性准则，指出 n 比特 S 盒的非 0 输出组合函数的代数次数应当均为 $n-1$。但文献[37]的结果显示，所有非 0 组合函数的代数次数均为 3 的 4 比特 S 盒不可能是 Serpent 类型 S 盒，那么 Boura 和 Canteaut 给出的设计准则与 Serpent 类型 S 盒的设计要求是互斥的。文献[40]也给出了关于可分性的安全性设计准则，称为“Perfect”S 盒，这里的“Perfect”是只针对可分性而言。文献[40]曾想将此类 S 盒应用到 PRESENT 和 RECTANGLE 密码算法上，但由原始的 S 盒无法通过线性变换生成“Perfect”S 盒。随后，他们提出了“Almost Perfect”S 盒的概念，同时指出利用“Almost Perfect”S 盒去代替 PRESENT 和 RECTANGLE

密码算法的原始 S 盒，可以提高密码算法抵抗基于可分性的积分攻击的能力，但是他们没有分析替代 S 盒对差分和线性攻击造成的影响。如何在不减弱抵抗其他攻击能力的基础上（符合 S 盒的设计标准），通过替换 S 盒改善密码算法抵抗基于可分性的积分攻击的安全性，对于密码算法的设计具有重要意义。

3　本文的主要工作

针对基于可分性的密码分析与设计方法，本文的主要工作如下。

（1）三子集比特可分性的自动化搜索方法研究。

三子集比特可分性可以有效地刻画密码算法的积分性质，但其传播的计算复杂度和存储复杂度极大，目前还没有有效的解决方案，如何有效刻画三子集比特可分性的传播是一个公开问题。为了解决此公开问题，我们研究了三子集比特可分性的性质，提出了三子集比特可分性的自动化搜索算法，并将算法分别应用到 ARX（modular Addition、Rotation，Xor）结构、SPN（Substitution-Permutation Network）结构和 Feistel 结构等分组密码算法中，具体如下。

三子集比特可分性的性质研究方面主要包含三部分内容。第一，给出了由 S 盒的代数表达式直接求取三子集比特可分性传播法则的通用方法，揭示了三子集比特可分性的代数本质。第二，研究了三子集比特可分性的约简性质，可以利用二子集比特可分性识别出冗余向量，通过移除冗余向量的方式约简三子集比特可分性向量集合，降低了计算和存储复杂度。第三，给出了三子集比特可分性的快速传播性质，将三子集比特可分性转化为二子集比特可分性，然后通过考查二子集比特可分性的传播，快速得到输出比特的积分性质。

三子集比特可分性的自动化搜索算法方面主要有三部分内容。第一，给出了三子集比特可分性传播的 3 个终止法则，分别对应三子集比特可分性的 3 种结果：“不确定”“0”“1”。第二，提出了三子集比特可分性传播的分块技术，将轮函数划分为若干基本运算模块，使得每块的向量个数膨胀都比较小，利用约简技术移除冗余向量后，再进行下一个基本运算模块的可分性传播，提高了三子集比特可分性的传播效率，降低了计算和存储复杂度。第三，依据约简性质、终止法则和分块技术，提出了三子集比特可分性的自动化搜索算法。

将三子集比特可分性的自动化搜索算法应用到分组密码算法上。首先，应用到 ARX 结构密码算法上，对于 SIMON32 密码算法，找到了无法由二子集比特可分性得到的 15 轮积分区分器。同时，对于 18 轮的 SIMON64 密码算法，自动化搜索算法可以找到 23 个平衡比特，比之前最长的积分区分器多 1 个平衡比特。其次，应用到 SPN 结构密码算法上，对于 PRESENT 密码算法，当输入明文量为 2^{60} 时，自动化搜索算法得到的积分区分器比之前的积分区分器多 3 个平衡比特；当输入明文量为 2^{63} 时，自动化搜索算法得到的积分区分器比之前的积分区分器多 6 个平衡比特。对于 RECTANGLE 密码算法，当输入明文量为 2^{60} 时，自动化搜索

算法得到的积分区分器比之前最长的积分区分器多 11 个平衡比特。最后，应用到 Feistel 结构密码算法上，对于 LBlock 算法，找到了需要更少数据量的 16 轮积分区分器。

三子集比特可分性自动化搜索算法的提出具有十分重要的意义，一方面，它使得考查大分组规模密码算法的三子集比特可分性首次成为现实，可以更加精确地评估密码算法的积分区分器长度；另一方面，它是一个普适性的自动化搜索方法，可以便捷地搜索特定密码算法的积分区分器，极大地提升密码算法的分析效率。相关成果发表于会议 ASIACRYPT 2019[41]。

（2）密钥异或运算的三子集可分性传播性质研究。

在研究 SPECK32 密码算法的积分区分器时，通过随机选取 2^{10} 个密钥，发现存在一个选择明文量为 2^{30} 的 6 轮实验积分区分器。在随机性假设下，这个积分区分器对所有密钥均成立的概率很大，但这个积分区分器无法被现有方法证明（包括三子集比特可分性），理论积分区分器与实验积分区分器之间还存在差距。

为了消除这一差距，研究了密钥对三子集比特可分性传播的影响，给出了“密钥异或”运算的三子集比特可分性传播新性质，同时提出了“密钥绕过”技术来消除部分密钥比特对三子集比特可分性传播的影响。然后，将“密钥绕过”技术与三子集比特可分性的自动化搜索算法结合，提出了基于密钥绕过技术的三子集比特可分性自动化搜索算法，并将其应用到 SPECK、KATAN 和 KTANTAN 族分组密码算法中，得到了更好的积分区分器。对于 SPECK32 密码算法，证明了实验得到的输入明文量为 2^{30} 的 6 轮积分区分器对所有密钥都是成立的，消除了理论积分区分器与实验积分区分器之间的差距。对于 KATAN64 和 KTANTAN64 密码算法，我们的算法可以找到 73.6 轮的积分区分器，提高了积分区分器的轮数。

“密钥绕过”技术可以对“密钥异或”运算的三子集比特可分性传播进行更精确的刻画，这为理解三子集比特可分性的传播提供了新的思路。相关成果发表于期刊 ToSC[41]。

（3）基于三子集比特可分性的立方攻击方法研究。

传统的立方攻击是实验性的，由于受限于计算能力，只能考查小立方阶的立方攻击情况。在 2017 年美密会上，Todo 等人首次将二子集比特可分性应用到立方攻击中，可以考查大立方阶的情况。但是二子集比特可分性理论无法确定所恢复出的超级多项式是包含密钥的，且恢复超级多项式的计算复杂度超过实际的计算能力。如果超级多项式是常数的话，密钥恢复攻击就会退化为区分攻击。为了得到更精确的立方攻击结果，我们首次将三子集比特可分性与立方攻击结合，提出了基于三子集比特可分性的立方攻击方法。

首先，引入“相似多项式”的概念，给出了超级多项式代数正规型（Algebraic Normal Form，ANF）系数与三子集比特可分性的关系，于是可以通过考查相似多项式的三子集比特可分性传播来恢复超级多项式的代数正规型系数。

然后，将密码算法划分为公开迭代密码算法和秘密迭代码算法，对于公开迭代密码算法，证明了立方攻击中超级多项式的代数表达式能够由三子集比特可分性完全精确恢复，从而可

以给出确定的密钥恢复攻击结果。

最后，将基于三子集比特可分性的立方攻击方法应用到序列密码算法 Trivium 上，将 2017 年美密会上恢复 832 轮 Trivium 超级多项式的理论计算复杂度从 2^{77} 降为实际可实现；将 2018 年美密会上恢复 839 轮 Trivium 超级多项式的理论计算复杂度从 2^{79} 降为实际可实现；同时首次给出了 841 轮 Trivium 超级多项式的理论恢复攻击结果，是目前最好的结果。

基于三子集比特可分性的立方攻击是首个可以实际恢复大规模立方集合对应的精确超级多项式的方法，为立方攻击的发展提供了全新的视角。相关成果发表于会议 ASIACRYPT 2019[41]。

（4）S 盒针对可分性的安全指标研究。

S 盒被广泛应用于分组密码算法的设计中，在最常见的两种设计结构（SPN 和 Feistel）中，S 盒是唯一的非线性组件，对密码算法的安全性起到至关重要的作用。例如，在 2015 年美密会上，Todo 利用 MISTY1 密码算法中 S 盒的可分性弱点，首次给出了全轮 MISTY1 的理论攻击结果。所以当新的攻击方法被提出时，新的安全指标也应当被加入到 S 盒的设计准则中。

我们提出了 S 盒针对可分性的安全指标，称为(l, m)-可分类，并证明该指标是置换异或等价的（Permutation-Xor Equivalence，PXE）。然后，研究了 Optimal S 盒、Serpent 类型 S 盒、Golden S 盒、Platinum S 盒和典型密码算法中的 4 比特 S 盒的可分性安全指标。另外，为了证明安全指标的合理性和科学性，将其应用到 PRESENT 和 LBlock 密码算法上，在不改变它们抵抗其他攻击安全强度的前提下（符合 S 盒的设计标准），通过替换 S 盒提高其抵抗基于可分性的积分攻击的能力。由基于可分性的积分区分器自动化搜索算法可知，PRESENT 抵抗基于可分性的积分攻击的能力提高了 2 轮，LBlock 抵抗基于可分性的积分攻击的能力提高了 1 轮。所以，提出 S 盒针对可分性的安全指标，对密码算法的设计具有重要指导作用。相关成果发表于期刊 Chinese Journal of Electronics[42]。

4　展望

在基于可分性的密码分析与设计方法研究方面，仍有几个问题值得深入地研究和思考。

（1）由于目前基于可分性的积分区分器自动化搜索算法没有考虑密钥调度算法的影响，对于一些采用简单密钥调度算法的轻量级密码算法，是否存在更优的积分区分器还有待进一步的研究。

（2）基于密钥绕过技术的三子集比特可分性自动化搜索算法虽然可以比较精确地搜索密码算法的积分区分器，但其不能从理论上保证搜索到的积分区分器就是对所有密钥均成立的最优积分区分器。对于 SIMON32、SIMECK32、SPECK32、KATAN32 和 KTANTAN32 密

码算法，因为分组规模只有 32 比特，可以在固定密钥条件下考虑其所有明/密文对的情况，通过实验可以验证本文自动化搜索算法得到的积分区分器就是对所有密钥均成立的最优积分区分器。但对于更大规模的密码算法，由于计算能力的限制，无法通过实验的方法证明本文得到的积分区分器就是最优积分区分器。目前还没有关于积分区分器的可证明安全理论，如何获得密码算法针对积分分析的可证明安全是值得关注的重要课题。

（3）本文只研究了 S 盒的可分性安全指标，P 盒的可分性安全指标该如何刻画尚需进一步的研究。另外，对于其他结构的密码算法，如 ARX 结构，如何给出它们的可分性安全指标也是值得研究的重要问题。

参考文献

[1] SHANNON C. Communication theory of secrecy systems [J]. Bell system Technical Journal, 1949, 28(4): 656-715.

[2] TODO Y. Structural evaluation by generalized integral property [C]. EUROCRYPT 2015, LNCS 9056, Springer, 2015, 287-314.

[3] TODO Y, ISOBE T, HAO Y, et al. Cube attacks on non-blackbox polynomials based on division property [C]. CRYPTO 2017, LNCS 10403, Springer, 2017, 250-279.

[4] TODO Y. Division property: Efficient method to estimate upper bound of algebraic degree [C]. Mycrypt 2016, LNCS 10311, Springer, 2017, 553-571.

[5] DAEMEN J, KNUDSEN L, RIJMEN V. The block cipher Square [C]. FSE 1997, LNCS 1267, Springer, 1997, 149-165.

[6] LUCKS S. The saturation attack - a bait for Twofish [C]. FSE 2001, LNCS 2355, Springer, 2002, 1-15.

[7] BIRYUKOV A, SHAMIR A. Structural cryptanalysis of SASAS [C]. EUROCRYPT 2001, LNCS 2045, Springer, 2001, 394-405.

[8] KNUDSEN L, WAGNER D. Integral cryptanalysis [C]. FSE 2002, LNCS 2365, Springer, 2002, 112-127.

[9] FERGUSON N, KELSEY J, LUCKS S, et al. Improved cryptanalysis of Rijndael [C]. FSE 2000, LNCS 1978, Springer, 2000, 213-230.

[10] TODO Y, AOKI K. Fft key recovery for integral attack [C]. CANS 2014, LNCS 8813, Springer, 2014, 64-81.

[11] SASAKI Y, WANG L. Meet in the middle technique for integral attacks against feistel

ciphers [C]. SAC 2012, LNCS 7707, Springer, 2012, 234-251.

[12] LAI X. Higher order derivatives and differential cryptanalysis [J]. Communications and Cryptography, 1994, 276, 227-233.

[13] CANTEAUT A, VIDEAU M. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis [C]. EUROCRYPT 2002, LNCS 2332, Springer, 2002, 518-533.

[14] BOURA C, CANTEAUT A, DE CANNIERE C. Higher-order differential properties of Keccak and Luffa [C]. FSE 2011, LNCS 6733, Springer, 2011, 252-269.

[15] TODO Y. Integral Cryptanalysis on Full MISTY1 [C]. CRYPTO 2015, LNCS 9215, Springer, 2015, 413-432.

[16] SUN B, HAI X, ZHANG W, et al. New observation on division property [J]. Science China (Information Sciences), 2017, 9, 274-276.

[17] BOURA C, CANTEAUT A. Another view of the division property [C]. CRYPTO 2016, LNCS 9814, Springer, 2016, 654-682.

[18] TODO Y, MORII M. Bit-based division property and application to Simon family [C]. FSE 2016, LNCS 9783, Springer, 2016, 357-377.

[19] XIANG Z, ZHANG W, BAO Z, et al. Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers [C]. ASIACRYPT 2016, LNCS 10031, Springer, 2016, 648-678.

[20] Gurobi: http://www.gurobi.com/.

[21] SUN L, WANG W, LIU R, et al. MILP-aided bit-based division property for ARX-based block cipher [J]. SCIENCE CHINA Information Sciences, 2018, 61(1-3).

[22] SUN L, WANG W, WANG M. Automatic search of bit-based division property for ARX ciphers and word-based division property [C]. ASIACRYPT 2017, LNCS 10624, Springer, 2017, 128-157.

[23] HU K, WANG M. Automatic search for a variant of division property using three subsets [C]. CT-RSA 2019, LNCS 11405, Springer, 2019, 412-432.

[24] DINUR I, SHAMIR A. Cube attacks on tweakable black box polynomials [C]. EUROCRYPT 2009, LNCS 5479, Springer, 2009, 278-299.

[25] ENGLUND H, JOHANSSON T, TURAN M. A framework for chosen IV statistical analysis of stream ciphers [C]. INDOCRYPT 2007, LNCS 4859, Springer, 2007, 268-281.

[26] MROCZKOWSKI P, SZMIDT J. The cube attack on stream cipher Trivium and quadraticity tests [J]. Fundamental Informaticae, 2012, 114(3-4): 309-318.

[27] YE C, TIAN T. A new framework for finding nonlinear superpolies in cube attacks against trivium-like ciphers [C]. ACISP 2018, LNCS 10946, Springer, 2018, 172-187.

[28] DINUR I, SHAMIR A. Breaking Grain-128 with dynamic cube attacks [C]. FSE 2011, LNCS 6733, Springer, 2011, 167-187.

[29] HUANG S, WANG X, XU G, et al. Conditional cube attack on reduced-round Keccak sponge function [C]. EUROCRYPT 2017, LNCS 10211, Springer, 2017, 259-288.

[30] LIU M, YANG J, WANG W, et al. Correlation cube attacks: From weak-key distinguisher to key recovery [C]. EUROCRYPT 2018, LNCS 10821, Springer, 2018, 715-744.

[31] YE C, TIAN T. Deterministic cube attacks [EB/OL]. IACR Cryptology ePrint Archive, 2018:1028. Available at https://eprint.iacr.org/2018/1082.pdf.

[32] WANG Q, HAO Y, TODO Y, et al. Improved division property based cube attacks exploiting low degree property of superpoly [C]. CRYPTO 2018, LNCS 10991, Springer, 2018, 275-305.

[33] DINUR I, MORAWIECKI P, PIEPRZYK J, et al. Cube attacks and cube-attack-like cryptanalysis on the round-reduced Keccak sponge function [C]. EUROCRYPT 2015, LNCS 9056, Springer, 2015, LNCS 733-761.

[34] LIU M. Degree evaluation of NFSR-based cryptosystems [C]. CRYPTO 2017, LNCS 10403, Springer 2017, 227-249.

[35] BIHAM E, SHAMIR A. Differential cryptanalysis of DES-like cryptosystems [C]. CRYPTO 1990. LNCS 537, Springer, 1991, 2-21.

[36] MATSUI M. Linear Cryptanalysis method for DES cipher [C]. EUROCRYPT 1993, LNCS 765, Springer, 1994, 386-397.

[37] LEANDER G, POSCHMANN A. On the classification of 4 bit S-Boxes [C]. WAIFI 2007, LNCS 4547, Springer, 2007, 159-176.

[38] BIHAM E, ANDERSON R, KNUDSEN L. Serpent: A new block cipher proposal [C]. FSE 1998, LNCS 1372, Springer, 1998, 222-238.

[39] SAARINEN M. Cryptographic analysis of all 4×4-Bit S-Boxes [C]. SAC 2011, LNCS 7118, Springer, 2012, 118-133.

[40] DERBEZ P, FOUQUE P, LAMBIN B. Linearly equivalent S-boxes and the division property [EB/OL]. Cryptology ePrint Archive. Report 2019/097. Availabel at https://eprint.iacr.org/2019/097.

[41] WANG S, HU B, GUAN J, et al. MILP-aided method of searching division property using three subsets and applications [C]. ASIACRYPT 2019. LNCS, vol. 11923, pp. 398-428, Springer, 2019.

[42] WANG S, HU B, GUAN J, et al. Exploring secret keys in searching integral distinguishers based on division property [J]. IACR Transactions on Symmetric Cryptology, 2020(3), 288-304.

[43] WANG S, HU B, GUAN J, et al. Research on the security criterion of S-boxes against division property [J]. Chinese Journal of Electronics, 2021, 30(1): 85-91.

对称密码的可证明安全与若干关键问题研究

沈耀斌

上海交通大学，电子信息与电气工程学院，上海，200240

摘　要：对称密码是一类密码学算法，具有运行速度快、便于软硬件实现和易于标准化等特点。这类算法通常是信息安全中实现数据机密性、数据完整性以及数据认证性保护的核心密码算法，在互联网、物联网和金融等领域有着十分广泛的应用。可证明安全是对称密码研究中的重要内容，它的研究对象涵盖标准算法、各种工作模式、相关安全模型和底层算法结构等，其成果能够为对称密码的分析、设计与应用提供科学规范与理论依据。本文将围绕对称密码的可证明安全这一主题展开，介绍在国际标准算法、新型安全模型和底层算法结构所取得的一些研究成果。

关键词：对称密码；可证明安全；消息认证码国际标准；多用户安全；广义 Feistel 结构

On the Provable Security of Symmetric-Key Cryptography

SHEN Yaobin

School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University, Shanghai, 200240

Abstract: Symmetric-key cryptography is a class of cryptographic algorithms, which are fast and efficient for large amount of data, easy to implement in both software and hardware, and easy to standardize. They are the core cryptographic algorithms for data privacy, data integrity and authenticity in information security, which are widely used in the field of Internet, Internet of Things, Financial Technology, and so on.

The provable security is one of important aspects in the research of symmetric-key algorithms, which covers standard algorithms, modes of operation, security models, the structure of underlying primitives. The study of provable security can provide scientific guideline and fundamental basis for

the design, analysis and application of symmetric key algorithms. Centering around the provable security, this thesis will introduce some recent results on international standard algorithms, new security models, and underlying primitives.

Keywords: Symmetric-key Cryptography; Provable Security; International Standard of MACs; Multi-user Security; Generalized Feistel Structure

1　引言

随着互联网的高速发展和近些年物联网的兴起，大量的数据每时每刻都被产生，并借由手机、基站和路由器等通信设备在公众社会与各式各样的电子设备迅速传播开来。数据在存储、交换和处理的过程中，如果没有使用特定的技术进行保护，那么很容易受到敌手恶意的截取、篡改和伪造。尤其是在互联网高度普及的今天，大到国防军工、政府机要，小到个人通信、电子支付，都面临着数据泄露、篡改和伪造等威胁。因此，研究对数据进行保护的信息安全技术尤为重要。特别是“棱镜”事件的发生，让人们更深刻地意识到保护数据对国家安全、社会稳定和个人隐私的重要性，更揭示了研究信息安全技术的迫切需求。

根据国际标准 ISO/IEC 17999，信息安全技术的主要目标包括机密性、完整性和可用性 3 个方面。其中，机密性确保只有被授权的用户才可以访问信息；完整性防止信息被未授权的用户进行篡改和伪造；可用性确保被授权的用户在需要时可以访问信息和相关资产。密码学是解决信息安全问题、实现信息安全目标的有效手段之一。

密码学主要分为两个方面：一方面是以 RSA 为代表的非对称密码，也称为公钥密码；另一方面是以 DES 为代表的对称密码，也称为私钥密码。对称密码具有运行速度快、便于软硬件实现和易于标准化等特点，通常是信息安全中保护数据机密性、数据完整性以及数据认证性的关键密码算法，广泛应用于互联网、物联网和金融科技等领域。

可证明安全是对称密码研究中的重要内容，它的研究对象包括现行标准算法、各种工作模式、相关安全模型、底层算法结构，以及密钥编排方案等，其成果对对称密码的分析、设计和应用具有重要理论与现实指导意义。

可证明安全的概念是由 Goldwasser 和 Micali[1]于 1984 年提出，最早被用于公钥密码的研究。他们的论文提出了语义安全（Semantic Security）的定义，首次从计算理论的角度对加密方案的安全性进行了定义。这篇文章开启了现代密码学的可证明安全研究。更具体地说，这里的可证明安全具体指的是，在分析整个密码方案的过程中，通过归约的方式，将方案的安全性归结到底层密码学组件的安全性。也就是说，如果底层密码组件是安全的，那么通过该组件构造的密码方案也是安全的；反之，如果存在攻破密码方案的敌手，那么利用该敌手，我们可以构造出另一个敌手，来攻破底层密码组件的安全性。这里的密码组件一般指的是密

码学中最基本的组成单位，如数学困难问题、分组密码算法和压缩函数等。

可证明安全理论在对称密码中的应用始于 Luby 和 Rackoff[2]对于 DES 的研究。DES 使用的 Feistel 结构有几个非常好的特性，但是一直没有理论上的可证明安全，即使在某个合理的数学假设下。Luby 和 Rackoff 的工作定义了伪随机置换和强伪随机置换的概念，他们开创性地证明了如果轮函数是伪随机函数，那么 3 轮的 Feistel 结构是伪随机置换，4 轮的 Feistel 结构是强伪随机置换。这一结果表明，如果轮函数是伪随机函数，那么 DES 从理论上说是安全的。此结果的意义在于将 DES 的安全问题归约到轮函数的伪随机性上。纵观对称密码可证明安全的发展史，研究工作主要集中于对称密码算法和对称密码应用两个方面展开。在对称密码算法方面，主要研究对称密码结构的安全性，如 Feistel 结构、SPN 结构等，主要思路是将其安全性归约到底层子模块的随机性。在对称密码应用方面，主要研究如何以分组密码为基础，设计满足各种需求的密码方案。

1994 年，Bellare 等人[3]首次将 CBC-MAC 的安全性归约为分组密码的伪随机性，即证明了如果分组密码满足伪随机置换的性质，那么 CBC-MAC 对于固定长度的消息是安全的。此后，在对称密码算法方面，涌现了大量可证明安全的研究成果。现在设计可证明安全的对称密码算法已经成为一种潮流，几乎所有新提出来的对称密码算法都有安全证明，可证明安全已经成为衡量对称密码算法好坏的重要指标。

2 研究现状

下面对现行国际标准算法、多用户安全、广义 Feistel 结构和密钥编排方案的国内外研究现状进行逐一介绍和总结。

2.1 国际标准 ISO/IEC 9797-1

消息认证码（MAC）是对称密码的基础算法之一，可以用来保护通信双方的数据完整性和认证性。MAC 通常需要被证明具有伪随机函数（PRF）的性质，从而表明具有不可伪造的安全性质。攻击者对 PRF 安全的优势一般由 4 个参数来衡量，即分组块的大小n、询问的总次数q、最长消息的长度ℓ和所有询问消息的总块数σ。在密码学中，存在多种方式来构造一个 MAC，迭代使用分组密码是比较流行的一种方式[4]。基于分组密码的 MAC 的安全分析或安全证明通常是在假设底层分组密码是伪随机置换（PRP）进行的。大部分基于分组密码的 MAC 能达到生日界安全，即最多能抵抗$O\left(2^{\frac{n}{2}}\right)$次攻击者询问。

对于基于分组密码的 MACs，仅仅达到生日界安全并不总是足够，尤其当分组密码的块长度比较小时。在资源受限的环境中，往往会使用轻量级的分组密码，如 PRESENT[5]、

PRINCE[6]和 GIFT[7]，或者在金融领域，传统的分组密码 TDES 还在被使用。这些分组密码的块长度都只有 64 比特，生日界安全会具体化为2^{32}。这个安全度是脆弱的，因为攻击者只要做约 32GB 的数据询问，就能破解算法的安全性。例如，Bhargavan 和 Leurent[8]利用短块分组密码容易产生碰撞的弱点，提出了两个对 https 协议和 OpenVPN 协议的实用攻击。所以，在一些实际应用中，使用超生日界安全的 MACs 很有必要。

ISO/IEC 9797-1 是一个基于分组密码的 MACs 国际标准。 现行的 ISO/IEC 9797-1:2011[9]提供了 6 种不同机制的 CBC-MACs，分别称为 MAC 算法 1~6。这 6 个 MACs 的主要区别在于最后迭代和输出处理方式的不同。这些 MACs 在实际应用中被广泛地使用，因此十分重要。因为这些 MACs 都使用了单链的 CBC-MAC 结构，它们都遭受生日界的伪造攻击[10]。为了提高安全度到超生日界安全，ISO/IEC 9797-1:2011 建议将两个 MACs 的输出串联起来：

“if a MAC algorithm with a higher security level is needed, it is recommended to perform two MAC calculations with independent keys and concatenate the results (rather than XORing them).”

然而，这个建议是否能达到超生日界安全尚未得到严格分析，目前仍然存疑。

2.2 多用户模型下超生日界安全

Double-block Hash-then-Sum 构造。为了超越生日界安全，一系列基于分组密码的 MACs 被陆续提出来，包括 SUM-ECBC[11]、PMAC_Plus[12]、3kf9[13]和 LightMAC_Plus[14]。有趣的是，这些 MACs 都使用了一个类似的框架，即 Double-block Hash-then-Sum（缩写为 DbHtS）。在这个框架中，消息首先会被一个双块的杂凑函数映射为 $2n$ 比特的字符串；然后将两个 n 比特半块的加密值异或起来，产生标签值。Datta 等人[15]将这个框架抽象出来，并且划分为两类：①3 个密钥的 DbHtS 构造，除了杂凑函数的密钥，在最后的加密步骤，使用了两个分组密码的密钥（包括 SUM-ECBC、PMAC_Plus、3kf9 和 LightMAC_Plus）；②2 个密钥的 DbHtS 构造，除了杂凑函数的密钥，在最后的加密步骤，仅使用了一个分组密码的密钥（包括 2 个密钥的 DbHtS 变形，即 2k-SUM-ECBC、2k-PMAC_Plus、2k-LightMAC_Plus 和 2kf9）。在这个框架下，他们证明了 3 个密钥和 2 个密钥的 DbHtS 构造都能够以界$\frac{q^3}{2^{2n}}$实现超生日界安全，其中q表示询问的个数，n表示分组密码块的大小。Leurent 等人[16]展示了对所有 3 个密钥的 DbHtS 构造的复杂度为$2^{\frac{3n}{4}}$的攻击。最近，Kim 等人[17]给出了 3 个密钥的 DbHtS 构造的紧致可证明安全界$\frac{q^{\frac{4}{3}}}{2^n}$。

多用户安全。以上所有超生日界安全的结果只考虑了单用户情形。然而，作为在现实中被用得最广的密码学基础算法之一，MAC 通常被部署在拥有大量用户的应用中。例如，MAC

是安全传输协议 TLS、SSH 和 IPSec 的核心部件。这些传输协议被主流的网站使用，每天都有成千上亿的用户。一个很自然的问题是，用户的数量会以何种程度影响 DbHtS 构造的安全界？或者更确切地说，在多用户情形下，DbHtS 构造是否依然能够实现超生日界安全？

多用户安全的概念分别由 Biham 在对称密码分析的研究中提出[18]，由 Bellare 等人在公钥加密的研究中提出[19]。敌手可以有策略地将它的询问分散在使用独立密钥的用户中。如果敌手能够攻击其中至少一位用户，就认为敌手是成功的。之前一系列的工作已经表明[20-27]，当用户的数量上升时，衡量安全界会如何下降是个有挑战的技术问题，即便是在单用户的安全界已经知道的情况下。然而，目前关于消息认证码的多用户安全的研究却寥寥无几。值得关注的个例有 Chatterjee 等人[28]的工作、最近 Andrew 等人[29]的工作 Bellare 等人[22]的工作。前两个工作考虑了通用归约，在该归约下，DbHtS 构造的多用户安全会被限制在生日界$2^{\frac{n}{2}}$（甚至比生日界更差）。下面会对此进行详细的讨论。后面的工作考虑的是基于杂凑函数的 MAC 的多用户安全，与基于分组密码的 MAC 有很大的不同。

假设用户的数量为u，则使用从单用户到多用户安全的通用归约[28-29]，以上关于 2 个密钥的 DbHtS 构造的超生日界在多用户模型下会变为

$$\frac{uq^3}{2^{2n}} \tag{1}$$

假设敌手对于每个用户只做一次询问，则该安全界会变为

$$\frac{uq^3}{2^{2n}} \leqslant \frac{q^4}{2^{2n}} \tag{2}$$

还是只能达到令人担忧的生日界。甚至对于在单用户模型下，有更好安全界$\frac{q^{\frac{4}{3}}}{2^n}$的 3 个密钥的 DbHtS 构造，使用通用归约得到的多用户安全界变为

$$\frac{uq^{\frac{4}{3}}}{2^n} \leqslant \frac{q^{\frac{7}{3}}}{2^n} \tag{3}$$

这比生日界$2^{\frac{n}{2}}$还差。因此不依赖于通用归约，直接分析 DbHtS 构造的多用户安全是很有必要的。

2.3 分组密码结构

分组密码是对称密码的底层算法，在信息安全领域中有广泛的应用。除了用来对数据进行加密，分组密码作为底层组件，还可以用来构造消息认证码、杂凑函数、伪随机数生成器和流密码等。

整体结构是每个分组密码算法的重要特征，对于分组密码的安全强度、轮数选择和软硬件性能都有很大的影响。Feistel 结构是使用最广的一类分组密码结构。Feistel 结构可以把函

数（通常称为轮函数）转换成一个置换，最早是由 H. Feistel 在设计 Lucifer 分组密码时提出的，并因 DES 的使用而流行。Feistel 结构是由多轮迭代的 Feistel 置换构成的。在经典的 Feistel 结构中，使用了如下 Feistel 置换：$\Psi_{F_i}(A,B)=(B,A\oplus F_i(B))$，其中函数$F_i:\{0,1\}^n\to\{0,1\}^n$，$A$和$B$是长度为$n$的比特串。许多分组密码算法都使用了经典的 Feistel 结构，包括美国的标准算法 DES、苏联设计的标准算法 GOST，以及 Blowfish 和 Twofish。经典的 Feistel 结构又称为 Luby-Rackoff 结构。在轮函数是伪随机函数的的假设下，Luby 和 Rackoff 于 1988 年证明了 3 轮的 Feistel 结构是伪随机置换，4 轮的 Feistel 结构是强伪随机置换[2]。后面一系列的工作继续对 Luby-Rackoff 结构进行了深入的研究，主要是从两个方面入手：一方面是证明了更好的安全界[30-36]；另一方面是在保证相同安全界的前提下，降低了结构的复杂度[37-40]。

广义的 Feistel 结构（GFNs）。上面经典 Feistel 结构可以通过不同的方法进行广义化。具体地说，将域保持函数F_i替换成扩展函数或压缩函数，可以得到不平衡 Feistel 结构[41]。交替使用扩展函数和压缩函数，可以得到交替的 Feistel 结构[42-43]。此外，将输入分为多于两个的分组（或者分支），可以得到多线 Feistel 结构，包括 1 类 Feistel 结构、2 类 Feistel 结构和 3 类 Feisel 结构[44]。这 3 类 Feistel 结构有不同的分支关系。和经典 Feistel 结构相比，广义 Feistel 结构具有更好的灵活性，从而也得到了更广泛的应用，包括超轻量级分组密码[45]、全域安全的加密算法[46]和宽置换[47]等。

分析广义 Feistel 结构的信息论意义上的安全模型与分析经典 Feistel 结构类似，文献[42-44,46,48-50]给出了生日界安全的结果，文献[35-36]给出了超生日界安全的结果。与本文更紧密相关的是，Hoang 和 Rogaway（缩写为 HR）[35]使用 coupling 技术，证明了上述类型的广义 Feistel 结构渐进意义上的最优安全。详细地说，在轮数足够大的情况下，上述类型的广义 Feistel 结构都是 CCA 安全的。对于任意的$\varepsilon>0$，都能抵抗$2^{n(1-\varepsilon)}$次攻击者询问。这个结果尽管看上去很不错，但为了达到渐进意义的n比特安全，它需要很大的轮数。

基于可调分组密码的 GFN。可调置换和可调分组密码由 Liskov 等人[51]率先提出来，前者定义了一族以参数调柄为索引的置换，后者定义了一族带密钥的可调置换。GFN 的轮函数可以替换为可调分组密码或可调置换，从而带来更多的可能性。

作为一个具体的例子，Coron 等人[52]提出了可以把带$\omega(\omega>n)$比特调柄的n比特可调置换转换成带$(\omega-n)$比特调柄的$2n$比特可调置换 GFN，即它用调柄空间换来了输入空间。因为调柄的扩展通常都比较容易[52,53]，这提供了一种扩展可调置换和可调分组密码的域的方法。在本文中，我们用TGF^r表示r轮的 Coron 等人的构造的变形。Coron 等人证明了当$r=2$时，TGF^r能实现生日界$2^{\frac{n}{2}}$的 CCA 安全；当$r=3$时，TGF^r能实现最优2^n的 CCA 安全。然而，注意到底层可调置换的输入大小实际上是比$2n$比特还大（n比特的输入加上ω比特的调柄）。最近 Lee B 和 Lee J[54]指出，对于此类可调置换，传统意义上的最优2^n安全实际上只是生日界。

在 Lee B 和 Lee J 的$2^{\frac{4n}{3}}$安全的可调分组密码构造的启发下，一个有趣的问题是对于$r \geqslant 4$轮的TGF^r，类似地超越2^n安全的结果是否能被证明。

2.4 密钥编排方案

一个分组密码算法通常包括轮函数和密钥编排方案两个部分。作为分组密码算法的重要组件之一，密钥编排方案并没有受到应有的关注。密钥编排方案通常输入一个主密钥，产生轮密钥，这些轮密钥在每轮中会被使用。以 AES-128 为例，主密钥是一个 128 比特的字符串，轮密钥的总长度为$11 \times 128 = 1408$比特。AES-128 的密钥编排方案可以被看作是从$\{0,1\}^{128}$到$\{0,1\}^{1408}$的映射。

如何科学地设计分组密码的密钥编排方案是一个重要的课题，但没有被很好的研究。总体来说，一个好的密钥编排方案需要遵循何种实际和必须的设计理念目前还不是很清楚。为了抵抗现存的攻击，密钥编排方案不应该有一些性质，如避免半弱密钥、等价密钥、对称性质和互补性质，以及实际密钥信息的不充足[55-56]。此外，密钥编排方案需要抵抗简单的猜测-确定攻击、中间相遇攻击、相关密钥攻击、滑动攻击和不变子空间攻击。从可证明安全的角度来考虑密钥编排方案是另一个方向。Chen 等人[57]使用了通过正形同构实现的密钥编排方案来最小化两轮的 Even-Mansour 结构，该结构只使用了一个n比特的主密钥和一个n比特的置换。他们证明这样的类 AES 构造能够实现超生日界安全。最近，Guo 和 Wang[58]使用了同样的正形同构密钥编排方案来得到一个生日界安全的 4 轮密钥交替的 KAF 结构，该结构只使用了一个n比特的主密钥和一个n比特的置换。他们宣称这个 4 轮的构造是理论意义上最小的，因为移除这个构造的任意一个部分都会破坏它的安全性。

除了提供必要的密码学安全，密钥编排方案的效率也很重要，尤其对于轻量级分组密码。轻量级分组密码通常被部署在资源受限的环境中，如 RFID 标签和传感器网络。在这些轻量级密码中，为了优化软件和硬件的效率，密钥编排方案通常是被高度简化。有些密钥编排方案使用了低扩散的逐轮迭代[5,59-60]，或者对主密钥进行简单的置换或线性操作。特别是在某些轻量级分组密码，它们的密钥编排方案是超轻量级的甚至不存在，直接在每轮中使用主密钥[61-62]。

3 本文的工作和成果

本文以可证明安全为切入点，深入研究了对称密码的若干个问题，包括现行标准算法的安全分析、多用户模型下消息认证码的安全证明、广义 Feistel 结构的安全分析，以及无密钥编排方案的密钥交替 Feistel 结构等，取得了以下研究成果。

3.1 ISO/IEC 9797-1:2011 的安全分析

首先，本文揭示了 ISO/IEC 9797-1：2011[9]的建议是错的，即串联组合器并不能超越生日界安全。对于串联的任意两个 MACs，本文提出了使用生日界复杂度的伪造攻击。令人惊讶的是，对于使用10^*填充的串联的两个 MACs 算法 1，本文的伪造攻击仅需要进行 3 次询问。这些攻击很好地说明了 ISO/IEC 9797-1:2011 推荐的串联组合器最多只能达到生日界安全。

其次，为了寻找修补方法，本文回顾了这个标准的发展历程。有趣的是，上一版本的标准 ISO/IEC 9797-1:1999[63]建议将两个 MACs 的输出异或起来。更具体地说，该旧版的标准推荐了两个 MACs 算法来实现超生日界安全，它们在标准文档中被称为 MAC5 和 MAC6。由于每个 MAC 算法可以使用 3 个填充方法，即 Pad1、Pad2 和 Pad3，其中 Pad1 会导致简单的伪造攻击。因此，在 ISO/IEC 9797-1:1999 中，一共有 4 个具体的 MACs 宣称能实现超生日界安全。Joux 等人[64]提出了对使用 Pad2 的 MAC5 的生日界伪造攻击，其中 MAC5 是将两个原始 CBC-MAC 简单地异或起来。这是目前仅有的对 CBC-MAC 异或组合器的生日界攻击，也应该是 ISO/IEC 9797-1：2011 没有使用异或组合器来超越生日界安全的主要原因。另外，Yasuda[11]证明了使用两个合适填充方法的 MAC6 确实能够超越生日界安全。因此，使用 Pad2 或 Pad3 的 MAC5 的可证明安全分析仍然是个公开问题，这也由 Rogaway 在给 CRYPTOREC 的报告中重点指出[65]。

本文衡量了将两个 MACs 异或起来的操作对 ISO/IEC 9797-1：2011 的影响，本文的研究结果表明该操作能有效地提高安全度。以 XMAC1 表示将 ISO/IEC 9797-1:2011 的两个 MACs 算法 1 异或起来的算法，该算法也即 ISO/IEC 9797-1:1999 中的 MAC5。本文给出了 XMAC1 的两个可证明安全界，包括使用填充方法 Pad2 和 Pad3 的 XMAC1。本文证明了使用填充方法 Pad3 的 XMAC1 能以安全界$O(\frac{\sigma q^2 \ell}{2^{2n}})$超越生日界安全。值得注意的是，本文的结果表明了 XMAC1 是第一个仅使用两个密钥来超越生日界安全的 CBC 类型的 MAC。当使用填充方法 Pad2 时，本文证明了 XMAC1 能以安全界$O(\frac{\sigma^2}{2^n})$来达到生日界安全。结合 Joux 等人的攻击以及不考虑常数的影响，这个界是紧致的。

此外，与之前的版本相比，ISO/IEC 9797-1:2011 推荐的 MAC 算法 5 是一个新引进来的单链 CBC 类型 MAC，通常被称为 CMAC。将异或的两个 MAC 算法 5 表示为 XMAC5，本文也证明了 XMAC5 能以安全界$O(\frac{\sigma q^2 \ell}{2^{2n}})$超越生日界安全。该工作发表在 FSE 2020 上[66]，并获得该届会议最佳论文奖。此外，还促使 ISO/IEC 国际组织编写了该标准的补篇文档“ISO 9797-1 AMENDMENT 1”[67]。

3.2 DbHtS MACs 在多用户模型下的超生日界安全

本文考虑了 DbHtS 构造在多用户模型下的安全性，尤其关注 2 个密钥的 DbHtS 构造。2 个密钥的 DbHtS 构造包括 2k-PMAC_Plus、2k-LightMAC_Plus 和 2kf9，总共只使用了 2 个分组密码的密钥。假设每个密钥的长度$k = n$，那么为了抵抗类似 Biham 对于 DES 的密钥碰撞攻击，2 个密钥是可能超越生日界安全的最少数量的密钥。

本文给出了在多用户模型下证明 2 个密钥的 DbHtS 构造的超生日界安全的通用框架。本文的框架便于使用，并且相较于之前的通用归约方法，能够得到更好的安全界。在这个框架下，本文只需要证明抽象出来的双块杂凑函数满足两个性质，即ϵ_1-规则和ϵ_2-规则几乎通用。第一个性质意味着当密钥随机从密钥空间选取时，对于一个消息，它的杂凑值等于任意一个固定的字符串的概率是很小的；第二个性质意味着当密钥随机从密钥空间选取时，对于任意两个不同的消息，它们的杂凑值发生碰撞的概率是很小的。DbHtS 构造的杂凑部分通常都能够满足这两个性质。

本文通过将该框架应用到 2 个密钥的 DbHtS 构造，展示了其可用性。更具体地说，本文证明了 2k-SUM-ECBC、2k-PMAC_Plus 和 2k-LightMAC_Plus 在多用户模型下仍然是超生日界安全的。本文证明的安全界和用户的数量无关，因此说明了 2 个密钥的 DbHtS 构造的安全界不会随着用户数量的增长而下降。另外，Datta 等人[15]使用了域分割函数来简化单用户安全的分析，同时使这些构造更复杂。而在这 3 个构造的证明过程中，本文没有依赖于域分割函数。因此本文的结果也表明了在不使用域分割函数的情况下，这 3 个构造在单用户和多用户模型下都能够超生日界安全。

更进一步地，本文发现 2kf9 在单用户模型下存在一个严重的缺陷。Datta 等人[15]证明了不使用域分割函数的 2kf9 是超生日界安全的，基于该结果，他们声称其他 3 个使用 2 个密钥的 DbHtS 构造在不使用域分割函数的情况下也能够超生日界安全。然而，不做任何询问，本文可以以概率 1 成功地伪造一个标签值。该缺陷在于，对于任意一个单块的消息，不使用域分割函数的 2kf9 总是输出 0。有人可能会觉得如果我们重新在 2kf9 使用域分割函数，那么它能够实现超生日界安全。然而，本文的攻击表明，即使在使用域分割函数的情况下，2kf9 也不能够超生日界安全。本文进一步研究了是否能够使用常见的办法对 2kf9 进行修缮，从而超生日界安全。不幸的是，对于这些变形的 2kf9，类似的攻击总是存在。该工作发表在 2021 年美密会上[68]。

3.3 提高广义 Feistel 结构的安全界

对于上述提到的 GFNs，本文要么提高了现有的 coupling 分析，要么提供了新的 coupling 分析（如果之前不存在）。具体地说，在 Lampe 和 Seurin[69]，以及 Nachef 等人[70]的启发下，

本文提高了 HR[35,71]的 coupling 分析，证明了以下结果。

- 对于不平衡 Feistel 的$UBF^r[m,n]$，当$n \geqslant m$时，本文证明了$(2\lceil\frac{n}{m}\rceil+2)t+2\lceil\frac{n}{m}\rceil+1$轮能够有安全界$\frac{2q}{t+1}\left(\frac{4\lceil\frac{n}{m}\rceil q+4q}{2^n}\right)^t$。这个界和 HR 的$\frac{2q}{t+1}\left(\frac{\left(3\lceil\frac{n}{m}\rceil+3\right)q}{2^n}\right)^t$相当，但所需轮数是 HR 的界$(4\lceil\frac{n}{m}\rceil+4)t$的一半。当$n<m$时，本文证明了$4t+2\lceil\frac{n}{m}\rceil+1$轮能够有安全界$\frac{2q}{t+1}\left(\frac{4\lceil\frac{n}{m}\rceil q}{2^n}\right)^t$，该安全界和 HR 的安全界是一样的，但所需轮数比 HR 的轮数小得多。

- 对于交替 Feistel 的$ALF^r[m,n]$，本文证明了$(12\left\lceil\frac{n}{m}\right\rceil 2)t+5$轮能够有安全界$\frac{2q}{t+1}\left(\frac{6\lceil\frac{n}{m}\rceil q+3q}{2^n}\right)^t$，而 HR 需要$\left(12\left\lceil\frac{n}{m}\right\rceil+8\right)t$轮来达到安全界$\frac{2q}{t+1}\left(\frac{\left(6\lceil\frac{n}{m}\rceil+3\right)q}{2^n}\right)^t$。对于数字版本的交替 Feistel，本文能得到相同的改善。

- 对于多线 GFNs 的$Feistel1^r[k,n]$和$Feistel2^r[k,n]$，本文分别证明了$(k^2+k-2)t+1$轮能有安全界$\frac{2q}{t+1}\left(\frac{2k(k-1)q}{2^n}\right)^t$、$2kt+1$轮能够有安全界$\frac{2q}{t+1}\left(\frac{2k(k-1)q}{2^n}\right)^t$，而 HR 分别需要$(2k^2+2k)t$轮来达到安全界$\frac{2q}{t+1}\left(\frac{2k\left(k^2-k+1\right)q}{2^n}\right)^t$，$(2k+2)t$轮来达到安全界$\frac{2q}{t+1}\left(\frac{2k(k-1)q}{2^n}\right)^t$。

- 对于 3 类 GFN 的$Feistel3^r[k,n]$，本文证明了$(k+2)t+1$轮能够有安全界$\frac{2q}{t+1}\left(\frac{4(k-1)^2q}{2^n}\right)^t$，而 HR 需要$(k+4)t$轮来达到安全界$\frac{2q}{t+1}\left(\frac{4(k-1)^2q}{2^n}\right)^t$。

对于基于可调分组密码的 GFN 的$TGF^r[\omega,2n]$，本文提供了 coupling 分析，证明了$4t+2$轮能够有$2\times\left(\frac{q}{t+1}\left(\frac{30q}{2^{2n}}\right)^t\right)^{\frac{1}{2}}$安全界。这展示了$TGF^r[\omega,2n]$的超越$2^n$安全界的结果。此外，当轮数$t$增加时，该安全界会接近$2^{2n}$，这带来了高安全度、双倍长度的分组密码。例如，当使用 Deoxys-BC-256 算法时，10 轮能够实现安全界$2^{\frac{4\times 128}{3}}\approx 2^{170}$。尽管效率没有很高，但高安全度使它适合某些应用。

改善的核心要点。本文对 HR[35]的改善源于对 coupling 概率更细致化的分析。为了进一步说明，考虑以使用压缩函数的非平衡 Feistel 为例，其中域为$\{0,1\}^n$和$\{0,1\}^m(n\geqslant m)$。HR 将该构造视为$2\lceil\frac{n}{m}\rceil+2$轮的大块，然后分析每个大块。在每个大块的内部 coupling 失败的概率最多为$\frac{3\lceil\frac{n}{m}\rceil\ell}{2^n}$。因为每个大块之间的事件都是独立的，最后 coupling 失败的概率是容易得到的。

然而，经过更仔细的研究，事实上，$\left\lfloor \frac{n}{m} \right\rfloor + 1$轮（只需大块一半的大小）已经足够让 coupling 成功。看起来 HR 使用额外的$\left\lfloor \frac{n}{m} \right\rfloor + 1$轮是为了建立不同大块之间的强独立性，从而可以进行模块化的论证（如上所述，他们可以只关注单个大块内部的情况），但本文可以进行更有针对性的分析：

首先，如上所述，本文缩小了每大块。本文更细致化的分析表明了即使当大块变小时，不同大块之间的事件某种程度上还是相互独立的；

其次，本文在每个构造的开始都预先增加了一些轮数，因此在这些轮数之后，两个赋值过程中（将会在 coupling 分析体现）的中间值会在某种程度上变得随机并且无碰撞。这对于 coupling 的分析尤为重要。

因此，最终本文能够使用接近一半的轮数来达到相当的安全界。该工作发表在 FSE 2020 上[72]。

3.4 无密钥编排方案的密钥交替 Feistel 结构

本文从可证明安全的角度，研究了如何设计一个尽可能轻量级的密钥编排方案这个有趣的问题，回顾了 Guo 和 Wang 的 4 轮 KAF。尽管通过线性正形同构实现的密钥编排方案在某些应用可以是有效率的，当轻量级密码应用在很多资源受限的环境中，该方案仍然不是令人十分满意。本文优化了 Guo 和 Wang 的结构，提出了一个新的、使用超轻量级（不需要）密钥编排方案的 4 轮 KAF。有趣的是，本文发现通过对第 1 轮一个小的改动，即在第一个轮函数后面使用一个 1 比特翻转函数，他们构造中的正形同构可以被移除。本文证明改进后的构造是生日界安全。和 Guo 和 Wang 的结构相比，本文的构造有两个优势。第一，即比较突出的一点是，密钥编排方案是超轻量级的（实际上不需要），从而不会消耗计算和内存。不需要依赖于任何的轮密钥生成函数，本文提出的方案只需要在相应的轮将主密钥异或进去。第二，在大部分应用中，1 比特翻转比 Guo 和 Wang 构造中使用的线性正形同构来得更有效，因为它只需要使用 1 比特移位而不是加法或域上的乘法。本文相信这个构造是理论上最小的（比 Guo 和 Wang 构造更轻量级），因为移除 1 比特翻转或任意一个组件都会令这个构造不安全。据本文所知，这是首个可证明安全的、没有使用任何密钥编排方案、使用相同轮函数和n比特主密钥的密钥交替的 Feistel 结构。

另外，本文研究了同样的 1 比特翻转是否适用于使用相同轮函数、单密钥的 3 轮密钥交替 Feistel。这次本文发现该 3 轮的构造不是伪随机置换（PRP），展示了只使用 4 个加密询问的区分攻击。此外，本文证明了使用合适密钥编排方案的 3 轮密钥交替的 Feistel 可以实现 PRP 安全。该文发表在 Science China Information Sciences[73]。

4　总结与展望

4.1　总结

对称密码是密码学的一类基础算法，可以用来保护数据的机密性、完整性和认证性等，是实现信息安全目标的核心算法。本文以可证明安全理论为切入点，对对称密码算法展开深入细致的研究，研究内容包括国际标准算法、消息认证码、相关安全模型、底层算法结构以及密钥编排方案等，取得了以下成果。

（1）**国际标准 ISO/IEC 9797-1：2011 的安全分析。**本文指出了国际标准 ISO/IEC 9797-1:2011 中使用串联操作来提供高安全度 MAC 的建议是无效。本文的攻击表明，串联操作并不能提高 MACs 的安全度，甚至仅需 3 个询问，就能攻破串联后的 MACs。此外，本文提出了修补建议，证明了将两个 MACs 异或起来的操作能有效地超生日界安全。

（2）DbHtS MACs **在多用户模型下的超生日界安全。**本文提出了在多用户模型下证明 DbHtS 这类 MACs 的超生日界安全框架，并将该框架运用到密钥减少的 DbHtS 构造，包括 2k-SUM-ECBC、2k-LightMAC_Plus 和 2k-PMAC_Plus。本文的证明框架能有效地超越生日界安全，而之前的通用归约最多只能达到生日界安全。此外，本文还以一个询问就攻击了 FSE'19 提出的 2kf9 构造，从而否定了 FSE'19 中的超生日界安全结果。本文还对 2kf9 几个常用的变形进行了分析，本文的分析表明，这几个变形最多只能达到生日界安全。

（3）**提高广义 Feistel 结构的安全界。**本文有效地优化了 Hoang 和 Rogaway（2010 年美密会）的 coupling 分析，从而提高了多个广义 Feistel 结构的安全界，包括非平衡 Feistel 结构、交替 Feistel 结构、1 类 Feistel 结构、2 类 Feistel 结构和 3 类 Feistel 结构。这些广义 Feistel 结构在对称密码算法有重要的应用，包括 Skijack 算法、Bear 和 Lion 算法、CAST-256 算法、RC6 算法和 MARS 算法等。此外，本文给出了基于可调分组密码的 Feistel 结构的 coupling 分析，从而证明了渐进意义上 $2n$比特安全。这提供了一种设计高安全度、双倍长度的分组密码算法的方法。

（4）**无密钥编排方案的密钥交替 Feistel 结构。**本文改进了 2018 年亚密会提出来的 4 轮密钥交替 Feistel 结构，提出了不使用密钥编排方案，只使用一个n比特主密钥和一个轮函数的密钥交替 Feistel 结构。本文的结构与 2018 年亚密会的结构保有相同的安全度，但不需要使用密钥编排方案，从而节约了计算和内存。此外，本文证明了使用合适密钥编排方案的 3 轮密钥交替 Feistel 结构是一个伪随机置换。

4.2　展望

随着新技术、新应用的不断涌现，对称密码是一个不断发展、欣欣向荣的密码学分支。

在本文工作的基础上，结合当今的前沿研究成果，本文将对称密码的未来可能的研究方向概括如下。

（1）**轻量级对称密码算法。**随着万物互联时代的到来，原先在常规设备上表现良好的密码算法可能在轻量级设备（包括嵌入式系统、RFID 和传感器网络）上面临水土不服的困境，如所需运行内存过大、消耗的计算能力难以承受等。因此，分析和设计能在资源受限的设备上使用的轻量级密码算法很有必要。特别是美国 NIST 正在进行的轻量级密码算法征集竞赛，更加突显了分析和设计轻量级对称密码算法的重要性。

（2）**抗泄露对称密码算法。**传统意义上的密码可证明安全主要考虑只看到密码算法输入输出的攻击者，忽略了算法在现实执行中产生的物理信息泄露，这导致了很多理论上安全的密码算法在实际应用中被轻易破解。抗泄露密码学用信息论意义上的泄露函数来刻画物理泄露信息，在此基础上设计可证明安全的密码算法和协议，使得算法和协议的安全性独立于现实中的硬件实现与攻击者的侧信道攻击，因此具有更广的安全性。如何设计和分析在合理物理信息泄露下可证明安全的对称密码算法是未来一个重要的研究方向。

（3）**后量子安全的对称密码算法。**随着量子计算技术的发展，以及 SIMON、Shor 和 Grover 等量子算法的提出，后量子密码成为密码学的研究热点。目前后量子对称密码算法的研究仍处于初步阶段，尚未形成完整的体系。现有的后量子对称密码可证明安全理论的研究主要集中在将经典环境下的安全概念和安全模型移植到量子环境下，并围绕基本的分组密码结构的安全性归约展开。如何设计后量子可证明安全的工作模式，如后量子安全的消息认证码、可调分组密码和杂凑函数等，是一个重要的研究方向。

参考文献

[1] GOLDWASSER S, MICALI S. Probabilistic encryption[J]. Journal of computer and system sciences, 1984, 28(2): 270-299.

[2] LUBY M, RACKOFF C. How to construct pseudorandom permutations from pseudorandom functions[J]. SIAM Journal on Computing, 1988, 17(2).

[3] BELLARE M, KILIAN J, ROGAWAY P. The Security of Cipher Block Chaining[C]. in:Desmedt Y. LNCS: CRYPTO’94: vol. 839. Springer, Heidelberg, 1994: 341-358.DOI: 10.1007/3-540-48658-5_32.

[4] BELLARE M, KILIAN J, ROGAWAY P. The Security of the Cipher Block Chaining Message Authentication Code[J/OL]. J.Comput.Syst.Sci.,2000,61(3):362-399.https://doi.org/10.1006/jcss. 1999.1694. DOI: 10.1006/jcss.1999.1694.

[5] BOGDANOV A, KNUDSEN L R, LEANDER G, et al. PRESENT: An Ultra-Lightweight Block Cipher[C]. in: Paillier P, Verbauwhede I. LNCS: CHES 2007: vol. 4727.Springer, Heidelberg, 2007: 450-466. DOI: 10.1007/978-3-540-74735-2_31.

[6] BORGHOFF J, CANTEAUT A, GÜNEYSU T, et al. PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract[C]. in:Wang X, Sako K.LNCS:ASIACRYPT 2012: vol. 7658. Springer, Heidelberg, 2012: 208-225. DOI: 10.1007/978-3-642-34961-4_14.

[7] BANIK S, PANDEY S K, PEYRIN T, et al. GIFT: A Small Present - Towards Reaching the Limit of Lightweight Encryption[C]. in: Fischer W, Homma N. LNCS:CHES 2017: vol. 10529. Springer, Heidelberg, 2017: 321-345. DOI: 10.1007/978-3-319-66787-4_16.

[8] BHARGAVAN K, LEURENT G. On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN[C]. in: Weippl E R, Katzenbeisser S, Kruegel C, et al. ACM CCS 2016. ACM Press, 2016: 456-467. DOI:10.1145/2976749.2978423.

[9] ISO/IEC:Information technology - Security techniques - Message Authentication Codes (MACs) - Part 1: Mechanisms using a block cipher[R]. ISO/IEC 9797-1:2011. 2011.

[10] PRENEEL B, VAN OORSCHOT P C. MDx-MAC and Building Fast MACs from Hash Functions[C]. in: Coppersmith D. LNCS: CRYPTO'95: vol. 963. Springer, Heidelberg,1995: 1-14. DOI: 10.1007/3-540-44750-4_1.

[11] YASUDA K. The Sum of CBC MACs Is a Secure PRF[C]. in: Pieprzyk J. LNCS:CT-RSA 2010: vol. 5985. Springer, Heidelberg, 2010: 366-381. DOI: 10.1007/978-3-642-11925-5_25.

[12] YASUDA K. A New Variant of PMAC: Beyond the Birthday Bound[C]. in: Rogaway P. LNCS: CRYPTO 2011: vol. 6841. Springer, Heidelberg, 2011: 596-609. DOI:10.1007/978-3-642-22792- 9_34.

[13] ZHANG L, WU W, SUI H, et al. 3kf9: Enhancing 3GPP-MAC beyond the Birthday Bound[C]. in: Wang X, Sako K. LNCS: ASIACRYPT 2012: vol. 7658. Springer, Heidelberg, 2012: 296-312. DOI: 10.1007/978-3-642-34961-4_19.

[14] NAITO Y. Blockcipher-Based MACs: Beyond the Birthday Bound Without Message Length[C]. in: Takagi T, Peyrin T. LNCS: ASIACRYPT 2017, Part III: vol. 10626. Springer, Heidelberg, 2017: 446-470. DOI: 10.1007/978-3-319-70700-6_16.

[15] DATTA N, DUTTA A, NANDI M, et al. Double-block Hash-then-Sum: A Paradigm for Constructing BBB Secure PRF[J]. IACR Trans. Symm. Cryptol., 2018, 2018(3):36-92. DOI:10.13154/tosc.v2018.i3.36-92.

[16] LEURENT G, NANDI M, SIBLEYRAS F. Generic Attacks Against Beyond-Birthday-

Bound MACs[C]. in: Shacham H, Boldyreva A. LNCS: CRYPTO 2018, Part I: vol. 10991. Springer, Heidelberg, 2018: 306-336. DOI: 10.1007/978-3-319-96884-1_11.

[17] KIM S, LEE B, LEE J. Tight Security Bounds for Double-Block Hash-then-Sum MACs[C]. in: Canteaut A, Ishai Y. LNCS: EUROCRYPT 2020, Part I: vol. 12105. Springer, Heidelberg, 2020: 435-465. DOI: 10.1007/978-3-030-45721-1_16.

[18] BIHAM E. How to decrypt or even substitute DES-encrypted messages in 228 steps[J]. Information Processing Letters, 2002, 84(3): 117-124.

[19] BELLARE M, BOLDYREVA A, MICALI S. Public-Key Encryption in a Multi-user Setting:Security Proofs and Improvements[C]. in: Preneel B. LNCS: EUROCRYPT 2000: vol. 1807. Springer, Heidelberg, 2000: 259-274. DOI: 10.1007/3-540-45539-6_18.

[20] TESSARO S. Optimally Secure Block Ciphers from Ideal Primitives[C]. in: Iwata T, Cheon J H. LNCS: ASIACRYPT 2015, Part II: vol. 9453. Springer, Heidelberg, 2015: 437-462. DOI: 10.1007/978-3-662-48800-3_18.

[21] MOUHA N, LUYKX A. Multi-key Security: The Even-Mansour Construction Revi sited[C]. in: Gennaro R, Robshaw M J B. LNCS: CRYPTO 2015, Part I: vol.9215.Springer, Heidelberg, 2015: 209-223. DOI: 10.1007/978-3-662-47989-6_10.

[22] BELLARE M, BERNSTEIN D J, TESSARO S. Hash-Function Based PRFs: AMAC and Its Multi-User Security[C]. in: Fischlin M, Coron J S. LNCS: EUROCRYPT 2016, Part I: vol. 9665. Springer, Heidelberg, 2016: 566-595. DOI: 10.1007/978-3-662-49890-3_22.

[23] HOANG V T, TESSARO S. Key-Alternating Ciphers and Key-Length Extension:Exact Bounds and Multi-user Security[C]. in: Robshaw M, Katz J. LNCS:CRYPTO 2016, Part I: vol. 9814. Springer, Heidelberg, 2016: 3-32. DOI: 10 . 1007/978-3-662-53018-4_1.

[24] LUYKX A, MENNINK B, PATERSON K G. Analyzing Multi-key Security Degra dation[C]. in: Takagi T, Peyrin T. LNCS: ASIACRYPT 2017, Part II: vol. 10625.Springer, Heidelberg, 2017: 575-605. DOI: 10.1007/978-3-319-70697-9_20.

[25] BOSE P, HOANG V T, TESSARO S. Revisiting AES-GCM-SIV: Multi-user Secu rity,Faster Key Derivation, and Better Bounds[C]. in: Nielsen J B, Rijmen V. LNCS:EURO CRYPT 2018, Part I: vol. 10820. Springer, Heidelberg, 2018: 468-499. DOI:10.1007/978-3-3 19-78381-9_18.

[26] BELLARE M, TACKMANN B. The Multi-user Security of Authenticated Encrypti on:AES-GCM in TLS 1.3[C]. in: Robshaw M, Katz J. LNCS: CRYPTO 2016, Part I:vol. 9 814. Springer, Heidelberg, 2016: 247-276. DOI: 10.1007/978-3-662-53018-4_10.

[27] HOANG V T, TESSARO S, THIRUVENGADAM A. The Multi-user Security of GCM,

Revisited:Tight Bounds for Nonce Randomization[C]. in: Lie D, Mannan M, Backes M, et al. ACM CCS 2018. ACM Press, 2018: 1429-1440. DOI: 10.1145/3243734.3243816.

[28] CHATTERJEE S, MENEZES A, SARKAR P. Another Look at Tightness[C]. in: Miri A, Vaudenay S. LNCS: SAC 2011: vol. 7118. Springer, Heidelberg, 2012: 293-319.DOI: 10.1007/978-3-642-28496-0_18.

[29] MORGAN A, PASS R, SHI E. On the Adaptive Security of MACs and PRFs[C]. in:Moriai S, Wang H. LNCS: ASIACRYPT 2020, Part I: vol. 12491. Springer, Heidelberg,2020: 724-753. DOI: 10.1007/978-3-030-64837-4_24.

[30] PATARIN J. Pseudorandom Permutations Based on the D.E.S. Scheme[C]. in: LNCS:ESORICS'90. AFCET, 1990: 185-187.

[31] MAURER U M. A Simplified and Generalized Treatment of Luby-Rackoff Pseudorandom Permutation Generator[C]. in: Rueppel R A. LNCS: EUROCRYPT'92: vol. 658. Springer, Heidelberg, 1993: 239-255. DOI: 10.1007/3-540-47555-9_21.

[32] MAURER U M, PIETRZAK K. The Security of Many-Round Luby-Rackoff Pseudo-Random Permutations[C]. in: Biham E. LNCS: EUROCRYPT 2003: vol. 2656.Springer, Heidelberg, 2003: 544-561. DOI: 10.1007/3-540-39200-9_34.

[33] VAUDENAY S. Decorrelation: A Theory for Block Cipher Security[J]. Journal of Cryptology, 2003, 16(4): 249-286. DOI: 10.1007/s00145-003-0220-6.

[34] PATARIN J. Security of Random Feistel Schemes with 5 or More Rounds[C]. in: Franklin M. LNCS: CRYPTO 2004: vol. 3152. Springer, Heidelberg, 2004: 106-122. DOI: 10.1007/978-3-540-28628-8_7.

[35] HOANG V T, ROGAWAY P. On Generalized Feistel Networks[C]. in: Rabin T. LNCS:CRYPTO 2010: vol. 6223. Springer, Heidelberg, 2010: 613-630. DOI: 10.1007/978-3-642-14623-7_33.

[36] PATARIN J. Security of balanced and unbalanced Feistel Schemes with Linear Non Equalities[Z]. Cryptology ePrint Archive, Report 2010/293. http://eprint.iacr.org/2010/293. 2010.

[37] SADEGHIYAN B, PIEPRZYK J. A Construction for Super Pseudorandom Permutations from A Single Pseudorandom Function[C]. in: Rueppel R A. LNCS:EUROCRYPT'92: vol. 658. Springer, Heidelberg, 1993: 267-284. DOI: 10.1007/3-540-47555-9_23.

[38] PATARIN J. How to Construct Pseudorandom and Super Pseudorandom Permutations from one Single Pseudorandom Function[C]. in: Rueppel R A. LNCS: EUROCRYPT'92: vol. 658. Springer, Heidelberg, 1993: 256-266. DOI: 10.1007/3-540-47555-9_22.

[39] NANDI M. The Characterization of Luby-Rackoff and Its Optimum Single-Key

Variants[C]. in: Gong G, Gupta K C. LNCS: INDOCRYPT 2010: vol. 6498.Springer, Heidelberg, 2010: 82-97.

[40] NANDI M. On the Optimality of Non-Linear Computations of Length-Preserving Encryption Schemes[C]. in: Iwata T, Cheon J H. LNCS: ASIACRYPT 2015, Part II: vol. 9453. Springer, Heidelberg, 2015: 113-133. DOI: 10.1007/978- 3-662-48800-3_5.

[41] SCHNEIER B, KELSEY J. Unbalanced Feistel Networks and Block Cipher Design[C]. in: Gollmann D. LNCS: FSE'96: vol. 1039. Springer, Heidelberg, 1996: 121-144.DOI: 10.1007/3-540-60865-6_49.

[42] ANDERSON R J, BIHAM E. Two Practical and Provably Secure Block Ciphers:BEARS and LION[C]. in: Gollmann D. LNCS: FSE'96: vol. 1039. Springer, Heidelberg,1996: 113-120. DOI: 10.1007/3-540-60865-6_48.

[43] LUCKS S. Faster Luby-Rackoff Ciphers[C]. in: Gollmann D. LNCS: FSE'96: vol. 1039. Springer, Heidelberg, 1996: 189-203. DOI: 10 . 1007/3-540-60865-6_53.

[44] ZHENG Y, MATSUMOTO T, IMAI H. On the Construction of Block Ciphers Provably Secure and Not Relying on Any Unproved Hypotheses[C]. in: Brassard G. LNCS:CRYPTO'89: vol. 435. Springer, Heidelberg, 1990: 461-480. DOI: 10.1007/0-387-34805-0_42.

[45] SHIBUTANI K, ISOBE T, HIWATARI H, et al. Piccolo: An Ultra-Lightweight Blockcipher[C]. in: Preneel B, Takagi T. LNCS: CHES 2011: vol. 6917. Springer, Heidelberg,2011: 342-357. DOI: 10.1007/978-3-642-23951-9_23.

[46] MORRIS B, ROGAWAY P, STEGERS T. How to Encipher Messages on a Small Domain[C]. in: Halevi S. LNCS: CRYPTO 2009: vol. 5677. Springer, Heidelberg,2009: 286-302. DOI: 10.1007/978-3-642-03356-8_17.

[47] GUERON S, MOUHA N. Simpira v2: A Family of Efficient Permutations Using the AES Round Function[C]. in: Cheon J H, Takagi T. LNCS: ASIACRYPT 2016, Part I: vol. 10031. Springer, Heidelberg, 2016: 95-125. DOI: 10.1007/978-3-662-53887-6_4.

[48] NAOR M, REINGOLD O. On the Construction of Pseudorandom Permutations: Luby-Rackoff Revisited[J]. Journal of Cryptology, 1999, 12(1): 29-66. DOI: 10.1007/PL00003817.

[49] BLACK J, ROGAWAY P. Ciphers with Arbitrary Finite Domains[C]. in: Preneel B.LNCS: CT-RSA 2002: vol. 2271. Springer, Heidelberg, 2002: 114-130. DOI: 10.1007/3-540-45760-7_9.

[50] BELLARE M, RISTENPART T, ROGAWAY P, et al. Format-Preserving Encryption[C]. in: Jacobson Jr. M J, Rijmen V, Safavi-Naini R. LNCS: SAC 2009: vol. 5867. Springer, Heidelberg, 2009: 295-312. DOI: 10.1007/978-3-642-05445-7_19.

[51] LISKOV M, RIVEST R L, WAGNER D. Tweakable Block Ciphers[C]. in: Yung M. LNCS:

CRYPTO 2002: vol. 2442. Springer, Heidelberg, 2002: 31-46. DOI: 10.1007/3-540-45708-9_3.

[52] CORON J S, DODIS Y, MANDAL A, et al. A Domain Extender for the Ideal Cipher[C]. in: Micciancio D. LNCS: TCC 2010: vol. 5978. Springer, Heidelberg,2010:273-289. DOI: 10.1007/978-3-642-11799-2_17.

[53] MINEMATSU K, IWATA T. Tweak-Length Extension for Tweakable Blockciphers[C]. in: Groth J. LNCS: 15th IMA International Conference on Cryptography and Coding:vol. 9496. Springer, Heidelberg, 2015: 77-93. DOI: 10.1007/978-3-319-27239-9_5.

[54] LEE B, LEE J. Tweakable Block Ciphers Secure Beyond the Birthday Bound in the Ideal Cipher Model[C]. in: Peyrin T, Galbraith S. LNCS: ASIACRYPT 2018, Part I: vol. 11272. Springer, Heidelberg, 2018: 305-335. DOI: 10.1007/978-3-030-03326-2_11.

[55] RIJMEN V, DAEMEN J. The Design of Rijndael: AES[J]. The Advanced Encryption Standard. Springer, Berlin, 2002.

[56] YAN H, LUO Y, CHEN M, et al. New observation on the key schedule of RECTANGLE[J]. Science China Information Sciences, 2019, 62(3): 32108.

[57] CHEN S, LAMPE R, LEE J, et al. Minimizing the Two-Round Even-Mansour Cipher[C]. in: Garay J A, Gennaro R. LNCS: CRYPTO 2014, Part I: vol. 8616.Springer, Heidelberg, 2014: 39-56. DOI: 10.1007/978-3-662-44371-2_3.

[58] GUO C, WANG L. Revisiting Key-Alternating Feistel Ciphers for Shorter Keys and Multi-user Security[C]. in: Peyrin T, Galbraith S. LNCS: ASIACRYPT 2018, Part I: vol. 11272. Springer, Heidelberg, 2018: 213-243. DOI: 10.1007/978- 3-030-03326-2_8.

[59] SUZAKI T, MINEMATSU K, MORIOKA S, et al. TWINE : A Lightweight Block Cipher for Multiple Platforms[C]. in: Knudsen L R, Wu H. LNCS: SAC 2012: vol. 7707. Springer, Heidelberg, 2013: 339-354. DOI: 10.1007/978-3-642-35999-6_22.

[60] WU W, ZHANG L. LBlock: A Lightweight Block Cipher[C]. in: Lopez J, Tsudik G. LNCS: ACNS 11: vol. 6715. Springer, Heidelberg, 2011: 327-344. DOI: 10.1007/978-3-642-21554-4_19.

[61] GUO J, PEYRIN T, POSCHMANN A, et al. The LED Block Cipher[C]. in: Preneel B, Takagi T. LNCS: CHES 2011: vol. 6917. Springer, Heidelberg, 2011: 326-341. DOI: 10.1007/978-3-642-23951-9_22.

[62] KNUDSEN L R, LEANDER G, POSCHMANN A, et al. PRINTcipher: A Block Cipher for IC-Printing[C]. in: Mangard S, Standaert F X. LNCS: CHES 2010: vol. 6225. Springer, Heidelberg, 2010: 16-32. DOI: 10.1007/978-3-642-15031-9_2.

[63] ISO/IEC:Information technology - Security techniques - Message Authentication Codes (MACs) - Part 1: Mechanisms using a block cipher[R]. ISO/IEC 9797-1:1999. 1999.

[64] JOUX A, POUPARD G, STERN J. New Attacks against Standardized MACs[C]. in: Johansson T. LNCS: FSE 2003: vol. 2887. Springer, Heidelberg, 2003: 170-181.DOI: 10.1007/978-3-540-39887-5_13.

[65] ROGAWAY P. Evaluation of some blockcipher modes of operation[J]. Cryptography Research and Evaluation Committees (CRYPTREC) for the Government of Japan, 2011.

[66] SHEN Y, WANG L. On Beyond-Birthday-Bound Security: Revisiting the Development of ISO/IEC 9797-1 MACs. IACR Transactions on Symmetric Cryptology. 2019 Jun 11:146-68.

[67] ISO/IEC:Information technology - Security techniques - Message Authentication Codes (MACs) - Part 1: Mechanisms using a block cipher - Amendment 1[R]. ISO/IEC 9797-1:2011/CD AMD 1.

[68] SHEN Y, WANG L, GU D, et al. Revisiting the security of DbHtS MACs: beyond-birthday-bound in the multi-user setting. In Annual International Cryptology Conference 2021 Aug 16 (pp. 309-336). Springer, Cham.

[69] LAMPE R, SEURIN Y. Security Analysis of Key-Alternating Feistel Ciphers[C]. in: Cid C, Rechberger C.LNCS: FSE 2014: vol. 8540. Springer, Heidelberg, 2015: 243-264. DOI: 10.1007/978-3-662-46706-0_13.

[70] NACHEF V, PATARIN J, VOLTE E. Feistel Ciphers - Security Proofs and Crypta nalysis[M/OL]. Springer, 2017. https://doi.org/10.1007/978-3-319-49530-9. DOI:10.1007/978-3-319-49530-9.

[71] HOANG V T, ROGAWAY P. On generalized Feistel networks[Z]. Cryptology ePrint Archive, Report 2010/301. http://eprint.iacr.org/2010/301. 2010.

[72] SHEN Y, GUO C, WANG L. Improved security bounds for generalized Feistel networks[J]. IACR Transactions on Symmetric Cryptology. 2020 May 7:425-57.

[73] SHEN Y, YAN H, WANG L, et al. Secure key-alternating Feistel ciphers without key schedule[J]. Science China Information Sciences. 2021 Jan;64(1):1-3.

模糊提取器的构造与安全性证明

温云华

上海交通大学，计算机科学与工程系，上海，200240

摘　要：密码算法的密钥通常假设是均匀随机的，在密码算法运行过程中也经常有随机数参与，因此安全可靠的随机数在密码学中起着至关重要的作用。然而，在现实生活中，能够直接产生安全可靠随机数的随机源是稀有的，噪声随机源却大量存在。例如，人的生物信息，它们有较高的最小熵但不够均匀随机且每次采样结果并不完全相同，而是存在一定的误差。模糊提取器可以从噪声随机源中提取出安全可靠的随机数为密码系统所用。本文将对模糊提取器的研究进展和最新技术进行简要介绍。

关键词：模糊提取器；鲁棒性；可重用性；线性级错误

Fuzzy Extractor：Construction and Security Proofs

WEN Yunhua

School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University, Shanghai, 200240

Abstract: In general, the secret keys of cryptographic schemes are presumed to be uniformly generated. At the same time, uniformly random strings are always demanded during the operations of cryptographic schemes. Therefore, uniformly random strings play a vital role in cryptography. However, random sources outputting such good strings are rare in real life. In contrast, there do exist plenty of imperfect noisy random sources, such as biometric information, which have enough entropy but are not uniformly random. Upon different samplings from these noisy sources, the samples are not identical and suffer from some noises. Fuzzy extractors can extract uniformly random strings from noisy random sources for cryptographic systems. In this report, we will briefly introduce the research progress and latest techniques in the field of fuzzy extractor.

Keywords: Fuzzy Extractor; Reusability; Robustness; Linear Fraction of Errors

1 引言

密码学是信息安全的核心与基石，众多现实应用的安全性很大程度上依赖于其底层密码算法的安全性，而密码算法的安全性则依赖于其密钥的安全性。密码学中著名的 Kerckhoff 准则说：“一个密码系统的安全性都应该基于密钥的安全性，而不是基于算法细节的安全性”。密码算法一般要求密钥是均匀随机产生的。此外，密码算法在运行过程中，还需要均匀产生的随机数参与以保证安全性。例如，公钥加密算法[1]或签名算法[2]都需要随机数参与，以保证加密算法的 CPA 安全和签名算法的不可伪造性。

因此，均匀随机的字符串在密码学中起着非常重要的作用。虽然利用确定性的数学算法可产生伪随机数，但由于其具有可预测性，难以确保密码系统的安全。另外，物理世界中的随机源具有一定的不可预测性，若其可以产生均匀随机的字符串，该随机源则可以为密码系统所用。然而，在现实生活中，能够直接产生均匀随机且精确再生的字符串的随机源几乎没有，噪声随机源却大量存在。例如，人的生物信息[3]、物理不可克隆函数[4]、量子信息[5]等均是噪声随机源。它们有很高的最小熵，但不是均匀随机的，而且每次的采样结果虽然相近，但都有一些小的偏差（噪声）。这使得噪声随机源无法直接为密码系统所用，造成了极大的浪费。

为了将噪声随机源应用到密码系统中，Dodis 等人[6]在 2004 年提出了模糊提取器的概念，旨在从噪声随机源中提取均匀随机的字符串。模糊提取器 FE=(Gen, Rep)有两个算法：生成算法和再生算法（图 1）。生成算法 Gen 输入字符串 w（噪声随机源的一次采样），输出一个字符串 R 和一个公开的帮助串 P；再生算法 Rep 输入 w'（噪声随机源的另一次采样）和公开帮助字符串 P，输出一个字符串 R'。模糊提取器的正确性要求如果两次采样 w 和 w'的距离足够近，那么 $R'=R$；模糊提取器的安全性要求如果随机源有足够多的熵，那么 R 是均匀随机的。

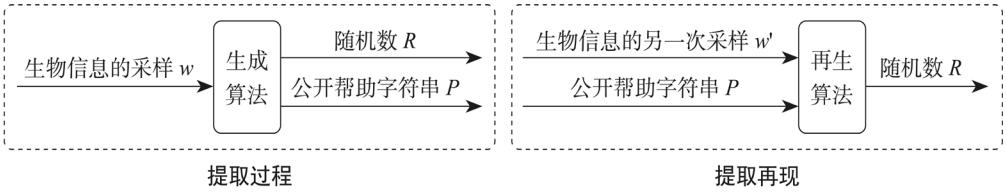

图 1 模糊提取器工作示意图

模糊提取器可以将噪声随机源转化成均匀随机且精确再生的字符串。这一良好性质使模糊提取器可以应用到密码系统中。例如，利用模糊提取器，用户可以从自己的生物信息（如指纹）中提取安全可靠的随机数 R 作为对称加密算法的密钥进行加/解密操作，且用户不需要存储密钥，当每次需要密钥时，只需要输入自己的生物信息，调用模糊提取器的再生算法就

可以将密钥 R 恢复出来，从而可以解决密钥的生成和存储问题。此外，模糊提取器还可以应用于公钥加密、身份认证[7]和密钥协商[8]等。

因此，进一步研究模糊提取器，设计性能更优的模糊提取器，可以更好地将生物信息、物理不可克隆函数等噪声随机源应用到密码系统中。本文将介绍模糊提取器国内外研究的主要研究进展，并简要介绍我们接下来需要进一步研究探讨的问题和难点。

2　研究现状

Dodis 等人[6]在欧密会 2004 上提出模糊提取器的概念以来，有不少与模糊提取器相关的研究成果发表。考虑到具体的应用场景，鲁棒模糊提取器（Robust Fuzzy Extractor）[7]和可重用模糊提取器（Reusable Fuzzy Extractor）[9]的概念相继提出。此外，随着量子计算机的发展，研究后量子安全的密码算法成为一个研究热点，后量子安全模糊提取器的相关研究也吸引了密码学者的目光。下面将分别介绍鲁棒模糊提取器、可重用模糊提取器和后量子安全模糊提取器的研究现状。

2.1　鲁棒模糊提取器的研究现状

模糊提取器只考虑了被动攻击的敌手，没有考虑主动攻击的敌手。如果主动攻击的敌手篡改了公开帮助字符串 P，那么用户可能在无意识的情况下得到一个错误的 R'进行加/解密操作，这会造成进一步的损失。为了解决上述问题，Boyen 等人[7]在 2005 年提出鲁棒模糊提取器的概念。模糊提取器的鲁棒性有两种安全性定义，分别为“应用前”鲁棒性和“应用后”鲁棒性。应用前鲁棒性保证了如果敌手在看到公开帮助字符串 P 后提交一个篡改的 P'，再生算法 Gen 输入 P'会以压倒性的概率输出终止符号，而不会输出一个错误的 R'。

在实际应用中，应用前鲁棒性是远远不够的。由于用户会将 R 参与到密码方案的运行过程中，R 的信息会部分地、甚至全部地泄露给敌手。在这种情况下，敌手不仅看到了公开帮助字符串 P，还得到了所提取字符串 R 的信息。由于敌手获取了更多关于 W 的信息，因此敌手就极有可能提交一个篡改的 P'，使得模糊提取器以大概率输出一个错误的字符串 R'。应用后鲁棒性可以解决上述问题。应用后鲁棒性保证了即使敌手在看到了公开帮助字符串 P 和所提取的字符串 R 后，再提交一个篡改的 P'，再生算法输入 P'也只能输出终止符号，而不会输出一个错误的 R'。

在欧密会 2005 上，Boyen 等人[7]提出了一种将模糊提取器转化成应用前鲁棒模糊提取器的通用方法，该方法使用哈希函数。其安全性证明需要将哈希函数当作随机谕言机，故该方法的安全性是建立在 Random Oracle 模型上的。

2006 年，Dodis 等人[8]首次在标准模型下构造了应用后鲁棒模糊提取器。在他们的构造

中，设输入噪声随机源的采样 w 是长度为 n、最小熵为 m 的比特串，模糊提取器可以从该噪声随机源中提取一个长度为 $l=(2m-n)/3$ 的均匀随机的比特串。由于 $n>m$，所提取出的随机串的长度不会超过最小熵的 1/3。2008 年，Kanukurthi 和 Reyzin[9]构造了一个升级版的应用后鲁棒模糊提取器，该提取器在给定相同输入的情况下，可以输出更长的随机串，为 $l=(2m-n)/2$ 比特。

文献[10,12]证明了在平凡模型下，如果输入 W 的熵率(m/n)小于一半，即使在不考虑纠错的情况下，模糊提取器的鲁棒性也是不可能实现的。为了解决这一问题，Cramer 等人[13]在 2008 年提出了一个新的密码学原语——“代数操作检测码”（Algebraic Manipulation Detection Code，AMD 码），并在共同参考字符串（Common reference String，CRS）模型下，利用 AMD 码构造了一个（应用后）鲁棒模糊提取器。CRS 模型指共同参考字符串固化在硬件中，任何人都不能对 CRS 进行篡改。他们所提出的鲁棒模糊提取器打破了在平凡模型下噪声随机源最小熵需要大于其长度的一半的界限。但是，在安全性归约中，作者需要 $w'-w$ 是公开的，这造成了巨大的熵损，不但缩短了所提取的字符串的长度，还提高了对噪声随机源最小熵的要求。此外，他们的方案要求噪声随机源与 CRS 是独立的。

在许多场景下，噪声随机源并不独立于 CRS。2021 年，Feng 等人[14]证明了信息论意义下，在 CRS 模型下的鲁棒模糊提取器如果允许噪声随机源与 CRS 相关，那么要求噪声随机源在条件上 CRS 后的条件熵至少是其长度的一半。为了降低对噪声随机源最小熵的要求，他们构造了一个计算安全的、CRS 模型下的鲁棒模糊提取器，该模糊提取器允许噪声随机源与 CRS 相关且只要求噪声随机源的条件熵大于 $\omega(\log_2 n)$即可。

2.2 可重用模糊提取器的研究现状

模糊提取器只能保证从一个噪声随机源中提取一个密钥的安全性，而无法保证从一个噪声随机源中提取多个不同密钥的安全性。在现实生活中，用户可能会在不同的机构使用不同的密钥进行密码操作。然而，人的生物信息是独一无二，与生俱来，不能被更改或创造的。如何从同一个生物信息中提取多个不同的安全可靠的密钥是一个急需解决的问题。为了解决这个问题，Boyen[9]在 2004 年提出可重用模糊提取器的概念。

Boyen[9]构造了两个可重用模糊提取器，并定义了“外向安全性”（Outsider Security）和“内向安全性”（Insider Security）。假设 $w_1,\cdots,w_n$ 是噪声随机源 W 多次采样的结果，调用模糊提取器的生成算法可以得到多组输出$(P_1, R_1),\cdots,(P_n, R_n)$。外向安全性保证了敌手看到了所有公开帮助字符串（$P_1,\cdots,P_n$）后，提取的字符串 R_i 是伪随机的；内向安全性保证了即使敌手看到了所有的公开帮助字符串（$P_1,\cdots,P_n$），还看到了除 R_i 之外的所有的提取出的字符串（$R_1,\cdots,R_{i-1}$，$R_{i+1},\cdots,R_n$），R_i 对于敌手而言仍然是伪随机的。上述两个安全模型都要求不同采样间的偏差 w_i-w_j 是由敌手动态选取的。在内向安全模型中，敌手获取了更多的信息，显然内向安全性

比外向安全性要强，但实现起来也更加困难。在文献[9]中，实现内向安全性的可重用模糊提取器是建立在随机谕言模型上的，此外还要求噪声随机源采样的偏差与噪声随机源是独立的。

由于噪声随机源的熵是有限的，在信息论意义下，源源不断地从噪声随机源中提取均匀随机的字符串是不可能的。因此，构造可重用模糊提取器一般都要诉诸于计算复杂性假设。Dodis 等人[12]提出可重用提取器（无须纠错）可以用在抗泄露密码学（Leakage-Resilient Cryptography）[15]中。可重用提取器本就复杂困难，可重用模糊提取器与可重用提取器相比又加入了纠错功能，这使得可重用模糊提取器构造更加困难。目前存在的许多模糊提取器[16]都不满足可重用性。

直到 2016 年欧密会上，一个新的可重用模糊提取器才由 Canetti 等人[21]提出。Canetti 等人定义的可重用模糊提取器的安全模型与 Boyen[9]所定义的内向安全性类似，不同之处是他们的安全模型对多次采样 $w_1,\cdots,w_n$ 的相关性没有任何限制，是目前安全性最强的一种安全模型。然而，他们的构造需要一个重要的模块“Composable Digital Locker”（CDL）[22]。目前 CDL 只有两种实现方法：一种实现方法是用随机谕言机来构造 CDL；另一种实现方法基于非标准假设，一个 DDH 假设的变种假设[23]。此外，他们的方案只能容忍亚线性级错误且对噪声随机源的分布有结构上的要求。模糊提取器容忍亚线性级错误指的是给定一个长度为 m 的输入，模糊提取器只能纠 t 个错误，其中 t 要满足 $t/n=o(c)$，c 是一个常数。要知道，有许多生物信息的错误率是线性比例的，如人的虹膜信息错误率为 20%~30%[24]。因此，纠亚线性级错误这一特性限制了该模糊提取器的应用。

类似于 Canetti 等人[21]的工作，Alamelou 等人[25]基于 CDL 构造了一个可重用模糊提取器。不同之处是他们提出了一个全新的密码学原语“Reusable Pseudoentropic Isometry”（RPI），并基于 CDL 构造了 RPI，然后以 RPI 为基础组件构造了一个可以纠线性级错误的可重用模糊提取器。由于该模糊提取器使用了 CDL 作为基础组件，该模糊提取器基于的是非标准假设。此外，该模糊提取器对噪声随机源的输入也有结构上的要求。具体来讲，他们的方案要求输入的噪声随机源采样可以划分为多个块，每块所在的字符集要足够大且拥有足够多的熵。

2.3　后量子安全模糊提取器的研究现状

随着量子技术的发展，如果大规模量子计算机问世，基于数论困难问题的密码方案将不再安全。LWE（Learning With Errors）问题和 LPN（Learning Parity With Noise）问题具有公认的抗量子特性。许多密码学者研究基于 LWE 和 LPN 困难问题的密码算法。类似地，如何设计基于 LWE 假设和 LPN 假设的模糊提取器也吸引了密码学者的目光。Fuller 等人[26]构造了第一个基于 LWE 假设的模糊提取器，该模糊提取器既不满足鲁棒性，也不满足可重用性。除此之外，为了减少熵损，他们不再利用安全梗概（Secure Sketch）来纠错，而是设计了一个 Decode 算法实现纠错功能。他们构造的模糊提取只能纠对数级的错误（如果输入 w 是长度为

n 的字符串，只能纠 $O(\log_2 n)$个错误）。此外，由于他们需要将噪声随机源作为噪声向量嵌入到 LWE 问题中，因此他们的方案要求输入的噪声随机源服从 LWE 问题的噪声分布。不同于常见的 LWE 问题，他们基于的 LWE 问题的噪声是在一个小区间上服从均匀分布而不是服从高斯分布。因此，他们的模糊提取器要求噪声随机源在小区间内服从均匀分布或某些固定位置上服从均匀分布，这样的要求与噪声随机源的实际情况是不相符的。

2017 年，Apon 等人[27]证明了 Fuller 等人[26]所提出的方案不满足可重用性。他们对 Fuller 等人人的方案进行升级得到了一个弱可重用模糊提取器。弱可重用模糊提取器需要限制敌手只能看到公开帮助字符串（$P_1,P_2,\cdots,P_n$），不能看到任意一个 R_i，才能保证提取随机数的安全性。此外，Apon 等人提出了一个将弱可重用模糊提取器转化为可重用模糊提取器的通用方法，然而该方法基于随机谕言机。同时，他们[27]基于 LWE 假设构造了一个可重用模糊提取器。不幸的是，与文献[26]中的方案类似，Apon 等人的可重用模糊提取器只能容忍对数级的错误。也就是说，如果输入字符串的长度为 n，其纠错个数 $t=O(\log_2 n)$。此外，他们构造的模糊提取器要求噪声随机源的分布服从 LWE 问题中的噪声分布，也就是离散高斯分布，这在实际中也很难成立。

Herder 等人[28]构造了第一个基于 LPN 假设的模糊提取器，然而该模糊提取器是在随机谕言机下安全的，且该模糊提取器既没有考虑可重用性也没有考虑鲁棒性。类似于 Fuller 等人[26]的方案，他们以物理不可克隆函数作为 LPN 问题的噪声向量，不同之处是他们的模糊提取器调用了一个 Project 函数，因此可以纠正线性级错误。2020 年，Li 等人[29]构造了一个基于 LPN 假设的可重用模糊提取器，该模糊提取器可以纠线性级错误。

已有的可重用模糊提取器在运行过程中，通常需要加入完美随机数以保证安全性，然而产生完美随机数在现实生活中是困难的。为了解决这一问题，2021 年，Cui 等人[30]构造了不依赖于完美随机数的可重用模糊提取器，该模糊提取器在运行时不再依赖于完美随机数，而是非完美随机数，因此更加实用。此外，他们的可重用模糊提取器还满足鲁棒性。他们分别给出了基于 DDH 假设和 LPN 假设的鲁棒可重用模糊提取器的实例化方案。

3 工作内容和成果

博士论文“模糊提取器的构造与安全性证明”围绕着模糊提取器进行了深入研究。取得代表性成果如下。

（1）构造了一个可以提取更长字符串的鲁棒模糊提取器。鲁棒模糊提取器保证了任意敌手对公开帮助字符串进行篡改都会导致模糊提取器输出 $\perp$。以往鲁棒模糊提取器的安全性定义都是统计意义下的。统计意义下的安全性固然很好，但是太严格，使得模糊提取器所提取的字符串可能太短而无法应用。为此，我们将鲁棒模糊提取器统计意义下的安全性放松到计

算安全性。计算安全性已经足够密码学的应用。同时，我们设计了一个特殊的“认证”方案，并基于该“认证”方案构造了一个计算安全的鲁棒模糊提取器。与 CDFPW 方案相比，在输入相同噪声随机源的情况下，我们的模糊提取器可以提取更长的字符串，且我们的方案对噪声随机源的要求比较低，对某些噪声随机源，CDFPW 方案无法从中提取，而我们的方案可以从中提取。相关成果发表在 The Computer Journal 上[31]。

（2）构造了一个基于标准假设的、可以纠线性级错误的可重用模糊提取器。可重用模糊提取器可以从相同的噪声随机源中提取多个均匀随机的字符串。可重用模糊提取器所基于的假设越标准，方案越安全；可重用模糊提取器容错率越高，对随机源的要求越低，方案的应用范围就越广。已有的可重用模糊提取器要么是基于随机谕言模型的，要么是基于非标准假设的，要么只能纠亚线性级错误。为了解决该问题，我们以安全梗概（Secure Sketch）和提取器（Extractor）为组件构造了一个可以纠线性级错误的可重用模糊提取器。利用安全梗概和提取器的同态性可将方案的安全性紧致归约到 DDH 假设上。此外，我们的方案是高效的，与传统的非可重用模糊提取器相比，我们只增加了两个群运算和一个哈希运算。相关成果发表在 Designs、Codes and Cryptography 上[32]。

（3）提出鲁棒可重用模糊提取器的通用构造方法。模糊提取器的鲁棒性保证了任何对公开帮助字符串 P 的改动，都会被用户监测到；模糊提取器的可重用性保证了对一个噪声随机源进行多次提取，所提取的字符串仍具有伪随机性。同时满足鲁棒性和可重用性的模糊提取器既可以抵御主动攻击的敌手，又可以从一个噪声随机源（如生物信息）中提取多个密钥应用在不同的机构中，因而有着更广泛的应用前景。然而，目前鲜有模糊提取器既考虑鲁棒性，又考虑可重用性。为此，我们形式化定义了鲁棒可重用模糊提取器（robustly reusable Fuzzy Extractor，rrFE）的安全模型，并提出了两个 rrFE 的通用构造方法。

- 我们构造了第一个基于标准假设的鲁棒可重用模糊提取器。我们提出了一个新的密码学原语对称密钥封装机制（Symmetric Key Encapsulation Mechanism，SKEM），并定义了其 key-shift 安全性，然后基于 SKEM、有损代数过滤器（Lossy Algebraic Filter）、提取器和安全梗概设计了一个 rrFE。通过对组件实例化，我们获得了第一个基于标准假设的、可以纠正线性级错误的鲁棒可重用模糊提取器。相关成果发表在 ASIACRYPT 2018 上[33]。
- 我们构造了第一个在非配对群上的、基于标准假设的鲁棒可重用模糊提取器。我们以支持辅助输入认证加密方案、哈希函数和安全梗概为组件设计了一个 rrFE。通过对组件实例化，我们获得了第一个在非配对群上的，基于 DDH 假设的鲁棒可重用模糊提取器。不仅如此，该模糊提取器还非常高效，且可以纠正线性级错误。相关成果发表在 PKC 2019 上[34]。

（4）构造了第一个可以纠线性级错误的、基于 LWE 假设的可重用模糊提取器和第一个基于 LWE 假设的鲁棒可重用模糊提取器。随着量子计算机的发展，传统的基于数论困难问题的密码方案将不再安全。目前，LWE 问题具有公认的抗量子特性。已有模糊提取器的安全性大

部分基于数论困难问题，目前虽然有少量基于 LWE 问题的模糊提取器[26]，但都只能纠对数级错误，且对噪声随机源的分布有着特殊的要求。为此，我们研究如何基于 LWE 假设构造模糊提取器并取得如下成果。

- 我们构造了一个具体的模糊提取器方案，该模糊提取器是第一个基于 LWE 假设的、可以纠线性级错误的可重用模糊提取器。相关成果发表在澳密会 2018 上[35]。
- 我们提出了第三个鲁棒可重用模糊提取器的通用构造方法，该方法以 unique-input key-shift 安全的伪随机函数、哈希函数和安全梗概为组件。通过对组件实例化，我们获得了第一个基于 LWE 假设的鲁棒可重用模糊提取器。相关成果发表在 PKC 2019 上[34]。

4 总结与展望

本文介绍了模糊提取器的研究进展，并简要介绍了我们所取得的研究成果。现有的关于模糊提取器的研究主要集中在两方面：一方面是降低对噪声随机源的要求，如对噪声随机源熵的大小、不同采样间的误差率和不同采样间的相关性等指标的要求；另一方面是提高模糊提取器的安全性，如是否基于标准假设、是否具有鲁棒性、是否具有可重用性、是否具有后量子安全性等。结合模糊提取器的研究现状和我们在该领域的积累，将未来的研究方向概括如下。

（1）已有模糊提取器在安全性和对噪声随机源的要求这两方面难以取得平衡。以可重用模糊提取器为例，允许噪声随机源不同采样之间可以任意相关的模糊提取器所基于的假设不标准、纠错率比较低且对噪声随机源有结构上的要求；而基于标准假设的可重用模糊提取器，对噪声随机源不同采样间的相关性有结构上的要求且需要噪声随机源有足够多的最小熵。一个重要的研究方向是如何平衡安全性和对噪声随机源的要求，设计更加实用的模糊提取器。

（2）随着量子计算机的发展，构造后量子安全的模糊提取器也是一个急需解决的问题。目前对后量子安全模糊提取器的研究还比较少，已有的后量子安全模糊提取器要么要求噪声随机源服从 LWE 或 LPN 问题中的噪声分布，要么要求有足够大的最小熵，缺乏高效、实用的后量子安全模糊提取器。一个重要的研究方向是设计实用的、后量子安全的模糊提取器。

参考文献

[1] CRAMER R, Shoup V. A Practical Public Key Cryptosystem Provably Secure Against Adaptive Chosen Ciphertext Attack. In CRYPTO 1998, pp. 13-25.

[2] WATERS, B. Efficient identity-based encryption without random oracles. In: EUROCRYPT

2005, pp. 114-127.

[3] KANG D, JUNG J, KIM H, et al. Efficient and Secure Biometric-Based User Authenticated Key Agreement Scheme with Anonymity. Security and Communication Networks, 2018: 9046064:1-9046064:14.

[4] SUZUKI M, UENO R, HOMMA N, et al. Efficient Fuzzy Extractors Based on Ternary Debiasing Method for Biased Physically Unclonable Functions. IEEE Trans. on Circuits and Systems,2019, 66-I(2): 616-629.

[5] BENNETT C H, BRASSARD G. Quantum cryptography: Public key distribution and coin tossing. Theor. Comput. Sci.2014, 560: 7-11.

[6] DODIS Y, REYZIN L, Smith AD. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. In EUROCRYPT 2004, pp.523-540.

[7] BOYEN X, DODIS Y, KATZ J, et al. Secure Remote Authentication Using Biometric Data. In: EUROCRYPT 2005, pp.147-163.

[8] DODIS Y, KATZ J, REYZIN L, et al. Robust Fuzzy Extractors and Authenticated Key Agreement from Close Secrets. In: CRYPTO 2006, pp. 232-250.

[9] BOYEN X. Reusable cryptographic fuzzy extractors. In: CCS 2004, pp. 82-91 (2004).

[10] KANUKURTHI B, REYZIN L. An Improved Robust Fuzzy Extractor. In: SCN.2008, pp:156-171.

[11] DODIS Y, SPENCER J. On the (non)Universality of the One-Time Pad. In: FOCS 2002，pp. 376-385.

[12] DODIS Y, WICHS D. Non-malleable extractors and symmetric key cryptography from weak secrets. In: STOC 2009, pp. 601-610.

[13] CRAMEr R, DODIS Y, FEHR S, et al. Detection of Algebraic Manipulation with Applications to Robust Secret Sharing and Fuzzy Extractors. In: EUROCRYPT 2008, pp. 471-488.

[14] FENG H, TANG Q. Computational Robust (Fuzzy) Extractors for CRS-Dependent Sources with Minimal Min-entropy. In TCC2021, pp.689-717.

[15] NAOR M, SEGEV G. Public-Key Cryptosystems Resilient to Key Leakage. In: CRYPTO 2009, pp.18-35.

[16] BUHAN I, DOUMEN J, HARTEL P.H, et al. Fuzzy extractors for continuous distributions. In: AsiaCCS 2007, pp. 353-355.

[17] SIMOENS K, TUYLS P, PRENEEl B. Privacy Weaknesses in Biometric Sketches. In: IEEE Symposium on Security and Privacy 2009, pp. 188-203.

[18] BLANTON M, ALIASGARI M. On the (Non-)reusability of Fuzzy Sketches and

Exractors and Security in the Computational Setting. In: SECRYPT2011, pp. 68-77.

[19] YAO J, LI K, ZHANG M, et al. A Robust Fuzzy Extractor without ECCs. In: Inscrypt 2012, pp. 60-68.

[20] BLANTON M, ALIASGARI M. Analysis of Reusability of Secure Sketches and Fuzzy Extractors. IEEE Trans. Information Forensics and Security.2013, 8(9): 1433-1445.

[21] CANETTI R, FULLER B, PANETH O, et al. Reusable Fuzzy Extractors for Low-Entropy Distributions. In: EUROCRYPT 2016, pp. 117-146.

[22] BITANSKY N, CANETTI R. On Strong Simulation and Composable Point Obfuscation In: CRYPTO 2010, pp. 520-537.

[23] CANETTI R, FULLER B, PANETH O, et al. Reusable Fuzzy Extractors for Low-Entropy Distributions. In: CRYPTO2021,pp. 34, 2.

[24] CHEON J H, JEONG J, KIM D, et al. A Reusable Fuzzy Extractor with Practical Storage Size: Modifying Canetti et al.'s Construction In: ACISP 2018, pp. 28-44.

[25] ALAMELOU Q, BERTHIER P, CACHET C, et al. Pseudoentropic Isometries: A New Framework for Fuzzy Extractor Reusability. In AsiaCCS 2018, pp. 673- 684.

[26] FULLER B, MENG X, REYZIN L. Computational Fuzzy Extractors. In: ASIACRYPT 2013, pp. 174-193.

[27] APON D, CHO C, ELDERAWY K, Efficient, Reusable Fuzzy Extractors from LWE. In: CSCML 2017, pp. 1-18.

[28] HERDER C, REN L, VAN DIJK M, et al, Devadas S. Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions. IEEE Trans. Dependable Sec. Comput. 2017,14(1): 65-82.

[29] LI Y, LIU S, GU D , et al. Reusable Fuzzy Extractor Based on the LPN Assumption[J]. The Computer Journal 2020(12), 12.

[30] CUI N, LIU S, WEN Y, et al. Pseudorandom Functions from LWE: RKA Security and Application. In: ACISP 2019, pp. 229-250.

[31] WEN Y, LIU S, HU Z, et al. Computational Robust Fuzzy Extractor. Comput J.2018, 61(12), pp. 1794-1805.

[32] WEN Y, LIU S, HAN S.Reusable fuzzy extractor from the decisional Diffie-Hellman assumption [J]. Des. Codes Cryptogr.2018, 86(11):2495-2512

[33] WEN Y, LIU S. Robustly Reusable Fuzzy Extractor from Standard Assumptions. In: ASIACRYPT 2018, pp. 459-489.

[34]　WEN Y, LIU S, GU D. Generic Constructions of Robustly Reusable Fuzzy Extractor. In: PKC 2019, pp. 349-378.

[35]　WEN Y, LIU S. Reusable Fuzzy Extractor from LWE. In: ACISP 2018, pp.13-27.

实用化量子保密查询协议的设计与应用

魏春艳

北京邮电大学，网络空间安全学院，北京，100876

摘　要：在某些通信场景下，人们不仅要保护传递的信息不被外部敌手窃取，还要保护通信双方的隐私不被对方获取。对称私有信息检索就是这样一类密码任务，它致力于保护数据检索中用户和服务器两方的隐私，本质上是“*N* 传 1”不经意传输。本文主要研究借助量子力学原理来实现对称私有信息检索，即量子保密查询。量子密码协议的优势在于其安全性受量子力学原理保护，与敌手的计算能力无关。尽管目前很多量子密码协议已被提出，但除量子密钥分配之外真正实用的协议很少，不利于搭建功能完备的量子通信网络。量子保密查询是当前实用潜力较为突出的一类量子密码协议，但仍存在如不能抵御参与者的不诚实测量攻击、不能容忍信道噪声等问题。本文简要介绍针对这些问题取得的研究进展。

关键词：量子密码；不经意传输；量子保密查询；量子密钥分配；匿名认证密钥交换

Design and Application of Practical Quantum Private Query Protocols

WEI Chunyan

School of Cyberspace Security,
Beijing University of Posts and Telecommunications, Beijing, 100876

Abstract: In some communication scenarios, people want to protect not only the transmitted information against outside adversaries but also the participants' privacy against each other. Symmetrically private information retrieval is exactly one of such cryptographic tasks. It aims to protect the privacy of both database holder and the user in the database retrieval, and it is in essence a variant of 1-out-of-*N* oblivious transfer. In this thesis, we mainly study the topic of realizing the task of symmetrically private information retrieval with the principles of quantum mechanics, i.e., quantum private query. The main advantage of quantum cryptographic protocols is that their security is protected by the principles of quantum mechanics, irrelevant with the adversaries' computation

capability. Though many kinds of quantum cryptographic protocols have been proposed, few are truly practical nowadays, except for the quantum key distribution, which cannot satisfy the requirement of establishing full-featured quantum communication networks. Quantum private query is one kind of quantum cryptographic protocol which shows great potential in practicality, but it still suffers certain problems, such as incapability of resisting participants' dishonest-measurement attacks, intolerance of the channel noise, and so on. We here will briefly introduce some research progress about dealing with these problems.

Keywords: Quantum Cryptography; Oblivious Transfer; Quantum Private Query; Quantum Key Distribution; Anonymous Authenticated Key Exchange

1　研究意义

“5G”新基建和大数据分析、人工智能等技术的迅猛发展极大地促进了电子商务、电子政务、自媒体的发展。这为人们的工作和生活带来了很多便利，也使得人们的隐私安全面临严重的泄露风险。不良服务商可以轻易获取用户的喜好、行动轨迹、联系方式甚至银行账号密码等私密信息。因此，如何在网络活动中保护自身的隐私安全成为人们当下最为关心的问题之一。

密码学是保障隐私安全的核心技术。然而，随着量子算法和量子计算的发展，现行密码体制（尤其是 RSA、ECC 等公钥密码体制）面临潜在的安全威胁[1-2]。现阶段，人们迫切希望发展能够对抗量子计算攻击的密码体制。将量子力学引入密码学发展起来的量子密码体制就具有这一优势，其安全性受量子物理基本原理保护，与敌手的计算能力无关。也就是说，量子密码借助量子力学原理来发现潜在的窃听，能够达到“信息论安全性”。基于这一优势，目前量子通信和量子密码已经成为与国家信息安全密切相关的战略竞争领域。建设安全高效、功能完备的量子通信网络既是量子通信技术发展的必然趋势，也是国家战略的迫切需求。

早在 20 世纪 70 年代，哥伦比亚大学的学者 Wiesner 就提出了量子货币的想法，阐述了量子效应可被用来保护信息安全的思想。1984 年，IBM 公司的 Bennett 和加拿大学者 Brassard 共同提出了第一个量子密钥分配（Quantum Key Distribution，QKD）协议（BB84 协议）[3]。紧接着，E91、B92、GV95、SARG04 等多个各具特色的 QKD 协议陆续被提出。针对这些协议的安全性证明方面也硕果累累，充分论证了量子力学可在密钥分配中实现“信息论”安全性。随后，考虑到理想的单光子源、单光子探测器、随机数生成器和无噪无损信道等理想设备在现有技术条件下难以获取，一些学者提出了诱骗态光源 QKD、测量设备无关 QKD 和设备无关 QKD 等协议。近年来，QKD 实验和商用化都取得了可观的进展，目前已经实现了百公里级的量子通信。在我国，借助卫星中继，最近已经实现了上千公里地面站间量子密钥的

分发。QKD 技术已经步入成熟商用阶段。

受 QKD 技术发展的鼓舞，人们迫切希望借助量子力学武装更多密码协议来实现更高的安全性，陆续提出了量子秘密共享、量子密钥协商、量子身份认证、量子签名、量子保密比较、量子比特承诺、量子掷币、量子匿名排序、量子匿名投票、量子保密查询等多种量子密码协议。然而，受到 No-go 定理理论限制及设备不完美等因素的影响，目前除 QKD 之外真正实用的量子密码方案仍极度匮乏，多数方案甚至连实验验证也难以完成。大家知道，一个成熟完备的通信系统需要具备认证、密钥分配、签名等多项功能。因此，在 QKD 以外，能否发掘出新的实用量子密码协议来满足认证、签名、数据检索等应用要求，已经成为关系到未来量子通信网络建设兴衰成败的关键问题。

量子保密查询（Quantum Private Query，QPQ）是目前除 QKD 之外实用潜力较为突出的一类量子密码协议[4-5]。这类协议能够容忍信道损失且借助现有 QKD 技术就能实现，部分协议已经得到了实验验证。本文主要研究解决这类协议在安全性和实用性方面存在的一些问题，力争进一步促进其实用化进程。另外，QPQ 本质上实现的是“多对一”不经意传输（Oblivious Transfer，OT）。鉴于 OT 在经典密码中可被用来实现其他安全多方计算任务，我们希望借鉴构造实用 QPQ 方案的思路来构造出实用的量子不经意传输（Quantum Oblivious Transfer，QOT）方案，进而借助其实现更多密码任务，以便为量子通信网络提供更多实用的量子密码协议。

2 量子保密查询的研究现状

量子保密查询是对称私有信息检索（Symmetrically Private Information Retrieval，SPIR）的量子方案。SPIR 最初由 Gentner 等人[6]于 2000 年提出，主要为了保护信息检索中用户和数据库拥有者两方的隐私。它的一个具体的应用场景如下：某个金融大鳄 Alice 希望从一个按条付费数据库的拥有者 Bob 那里检索一条她感兴趣的股票信息，检索条目在数据库中的位置（检索地址）能够揭示出 Alice 的购买倾向，因此 Alice 不希望任何人包括 Bob 获知该检索地址；另外，数据库方 Bob 希望 Alice 只能得到她购买的那个条目，而不能得到其余条目。本质上，SPIR 实现的是“多对一”的不经意传输，区别在于 SPIR 中传递的信息是数据库中的条目。人们在经典密码中已经提出了一些 SPIR 协议，但这些协议的安全性要么基于数学难题假设，容易遭受量子计算攻击的威胁；要么基于多数据库场景下的强假设，即要求协议中存在不止一个数据库，每个数据库都拥有完全一样的条目（能够保持同步更新）且数据库互不通信，这在某些应用场景下是不现实的。

随着量子密码尤其是 QKD 的发展，人们迫切希望借助量子力学原理来实现 SPIR。然而，正如量子比特承诺、量子掷币等两方安全计算任务受 No-go 定理[7-9]限制不能理想实现一样，SPIR 在量子密码中也不能理想实现[10]。更为实际地，其量子方案 QPQ 对安全性要求如下[11]：

①如果 Bob 通过某种欺骗试图获取用户 Alice 的检索地址，其欺骗行为将以非零的概率被 Alice 发现（欺骗敏感性）；②用户 Alice 除了她检索的条目，通常还可以额外获得几个条目。这里，Alice 额外得到的几个条目是随机的（Alice 不能控制额外获得的数据库条目所在的位置），一般不是她需要的。而 Bob 通常不敢冒着被发现欺骗的危险去攻击，因为一旦被发现将损害自己的声誉，甚至可能会面临十分严厉的惩罚。因此，这种安全性虽不理想但依然可以满足很多场景下的应用需求。人们在该安全要求下提出了很多 QPQ 协议。

2008 年，意大利学者 Giovannetti 等人[11]提出了第一个 QPQ 协议（GLM 协议）。在该协议中，Bob 将数据库信息编码到一个酉操作上，收到用户 Alice 的查询量子态后，他将该操作作用到查询态上然后返回给 Alice，Alice 通过测量返回的态获取想要的数据库条目。与经典 SPIR 方案相比，该协议的通信复杂度和计算复杂度都实现了指数级的下降。随后，作者给出了 GLM 协议的一个安全性证明及实验验证[12]。2011 年，波兰学者 Olejnik[13]将以上方案进行了改进，进一步降低了通信复杂度。2011 年，Wang 等人[14]给出了 GML 协议的一个变体。2014 年，Yu 等人[15]也提出了一种以纠缠态为查询态的 QPQ 协议。这类将数据库信息编码到酉操作上的 QPQ 协议在理论上意义非凡，尤其在通信复杂度和计算复杂度上优势明显，但它们实用性不强。一方面，将整个数据库（尤其是当数据库规模较大时）编码到一个酉操作上，那么该酉操作必然维数很大，在现有条件下难以实现；另一方面，这类协议不能容忍信道损失，即一旦存在信道损失的情形，将威胁到双方的隐私。例如，Bob 在收到查询态后对其进行测量就能够获得用户的检索地址，然后他可以声称没有收到该查询态（称其在传递过程中丢失），从而让用户重发。这样 Bob 既能获得用户隐私也不会被发现欺骗，显然不满足用户隐私的安全要求。

2011 年，瑞士日内瓦大学的 Jacobi 等人[4]基于 SARG QKD[16]提出了一个 QPQ 方案（J 方案），这是第一个基于 QKD 的 QPQ 协议。这类协议一般分为如下 3 个步骤来实现。

第一步：量子不经意密钥分配。双方共享一个不经意生密钥 K，它被 Bob 完全获得，但仅有部分比特被 Alice 获得，且 Bob 不知道这部分比特的位置。

第二步：经典后处理。双方对生密钥 K 来进行经典后处理（如按位相加）来得到一个最终密钥 K_f。这个过程主要是压缩 Alice 在不经意密钥中获得的比特，一般最终她在 K_f 中仅获得 1 个比特或几个比特。

第三步：检索。Bob 将数据库用 K_f 加密后发送给 Alice，Alice 用她知道的最终密钥比特恢复出想要的数据库条目。

这里只有第一步是量子过程，该过程可以用成熟的量子技术 QKD 来实现，且这类协议实现难度与数据库规模无关，能容忍信道损失。因此具有突出的实用潜力，已成为当前量子密码的研究热点之一。我们从如下几个方面来介绍基于 QKD 的 QPQ 的研究进展。

2.1 针对量子不经意密钥分配的研究

在某些数据库规模下，J 方案中用户可获得的数据条目数要么过大，不利于保护数据库安全性；要么过小，导致失败概率过大。为了解决该问题，2012 年，Gao 等人[17]使用调整载体态夹角的方法给出了一个灵活方案，它在任意数据库规模下都能使 Alice 预期获得的数据库条目数为一个特定的值。随后，Yang 等人[18]基于 B92 协议也给出了一个解决方案。2013 年，Zhang 等人[19]严格论证了将反直观 QKD 技术引入 QPQ 的可行性。2015 年，Yang 等人[20]提出了一个仅使用一种量子态来分发不经意密钥的 QPQ 方案。2016 年，Xu 等人[21]基于单光子干涉理论提出了两个 QPQ 协议。

近年来，为提高安全性和实用性，一些具有特殊性质的 QPQ 协议陆续被提出。在原始的 QPQ 协议中，一般将 1 个数据库条目模型化为 1 个比特，而实际应用中其往往为 1 个多比特的信息。用户采用逐比特查询的方式来获得完整的信息需要进行多次查询，不够实用，且任意一次查询中检索地址的泄露都意味着用户隐私完全泄露，不利于保护用户隐私。2014 年，Wei 等人[22]借助一种不均衡态 BB84 协议，提出了一个量子保密块查询方案，满足了用户在 1 次查询中可获得 1 个多比特条目的需求。随后，Shi 等人[23]使用二进制矩阵操作的后处理方法提出了另一种实现方案。最近，Pei 等人[24]基于置换技术也给出了块查询一个解决方案。2014 年，Yang 等人[25]给出了一个带检测的 QPQ 方案，提高了用户隐私安全性。Yu 等人[26]针对用户 Alice 难以及时发现 Bob 传送假数据的问题，也给出了一个带检测的方案。随后，Liu 等人[27]给出了一个基于环回差分相移 QKD 的 QPQ 协议，不仅实现了理想的数据库安全性，还将失败概率降为 0。Li 等人[28]基于被动环回差分相移 QKD 也给出了一个 QPQ 协议，改进了 MZ 干涉仪两臂中脉冲序列的长度差难以在现有技术条件下高速、稳定调节的状况。2016，Wei 等人[29]基于双路 QKD 协议提出了一个能够抵抗联合测量攻击的 QPQ 方案，即通过强制 Alice 在知道哪些载体态可被联合测量前就将载体态返回给 Bob，从而割裂了用户进行联合测量攻击的必要条件。2019 年，Liu 等人[30]借助干涉中光路与干涉结果之间的不确定关系构造了一个 QPQ 协议，该方案能够容忍信道损失，也能借助基于单光子干涉的通信系统实现。最近，Yan 等人[31]给出了一个经典 Bob 的 QPQ 方案，降低了数据库方的操作难度。

针对现有技术条件下光源、信道等设备的不完美性带来的挑战，人们也取得了一些研究进展。2016 年，Wang 等人[32]使用“无消相干态”作为载体态来分发密钥，给出了一个抗集体噪声的 QPQ 方案。2017 年，Yang 等人[33]基于四粒子“无消相干态”给出了一个抗集体噪声的协议，其接收者只需要执行单粒子测量就可以提取载体态的信息。2019 年，Li 等人[34]借助四粒子态也给出了一个普适的可对抗集体噪声的方案。另外，受 QKD 中借助 Bell 检测等方式来克服“侧信道”威胁启发，设备无关 QKD 和测量设备无关 QKD 也陆续被引入用于实现 QPQ 以克服不完美设备带来的安全威胁。2017 年，Zhao 等人[35]分析了不完美探测器带来

的安全隐患，给出了一个测量设备无关的 QPQ 协议。2011 年，Maitra 等人[36]提出了一个设备无关的 QPQ 方案，即使在设备由不诚实参与方提供的情况下，也能够保证方案的安全性。2018 年，Basak 等人[37]针对这个方案对 CHSH 测试和三方量子伪心灵感应策略进行了比较分析。2018 年，Roy 等人[38]提出了使用 qutrit 作为载体态的测量设备无关 QPQ 协议，与 qubit 作为载体态的方案相比，它具有更高的数据库安全性，但需要执行额外的测量来保证用户隐私安全。

2.2　针对经典后处理的研究

QPQ 的经典后处理方式对协议的执行效率和安全性都有重要的影响，它一般要实现两个目标：一是压缩 Alice 的优势使她在最终密钥中仅得到 1 个比特或几个比特；二是通过纠错来保证用户所传递信息的正确性。

针对第一个目标，继 J 协议中给出“按位相加”的后处理方式后，几个致力于提高协议效率的后处理方式被提出[39-41]。2013 年，Panduranga Rao 等人[40]给出了两个高效的后处理方案，能快速将 Alice 得到的最终密钥比特数压缩到 1 个或几个，大大提高了 QPQ 协议的效率。然而，2015 年，Gao 等人[42]发现这两种高效的后处理方式均存在严重的安全隐患，用户 Alice 可以借助多次查询获得远超预期的数据库条目数。2018 年，Wei 等人[43]给出了一种“窄移位相加”的后处理方法，既能实现理想的数据库安全性，又能将失败概率降为 0。

针对第二个目标，Gao 等人[42]在 2015 年研究后处理中的信息泄露问题时，提出了一种对不经意密钥的纠错方案，并针对其带来的信息泄露问题给出了解决方案。同年，Chan 等人[44]也给出了一种 QPQ 的纠错方案，其中 Bob 根据一个校验矩阵发送一些生密钥比特的和给 Alice，以便 Alice 可以估计出错误率并选择合适的（最不可能出错的）最终密钥比特来检索数据库。随后，Wei 等人[45]指出这些方案还存在一定的安全缺陷，即至少一方的隐私会遭到威胁，进而给出了一个带纠错的 QPQ 方案，该方案在降低错误率的同时兼顾了参与双方的隐私保护。

2.3　实验进展和应用推广

2014 年，卡尔加里大学 Chan 等人[44]完成了对 Gao 等人提出的方案[17]的一个验证性实验。2019 年，Li 等人[46]提出了一个适用于量子无线网络的 QPQ 方案，通过让用户节点和服务器节点之间预先共享纠缠态及引入多个协助第三方的方法实现了任意用户可向任意服务器进行检索的目标。最近，Kon 等人[47]研究了量子环境下保密数据库查询中多服务器的情形，指出在“服务器间不通信”这个假设下，协议可以实现信息论安全性。此外，2018 年，Xu 和 Luo 等人[48-49]分别实现了欺骗敏感的最近零查询等任务，这表明实用 QPQ 的研究能够有效地推动实用化量子安全多方计算协议的研究。

3 本文的工作和成果

3.1 抗不诚实测量攻击的实用 QPQ 协议

量子测量是量子密码中提取信息的主要手段。它不是唯一的，不同的测量方法往往能给测量者提供不同的信息量。因此，在量子多方安全计算中，不诚实的参与者往往不使用规定的测量方式，而采用一些更有利的测量方法来获得更多信息和优势。例如，QPQ 中用户的不诚实测量攻击会对数据库安全性带来严重威胁。一方面，用户的不诚实测量攻击会导致数据库条目严重泄露。以 J 方案[4]为例，当数据库条目数 N=10^4 时，诚实用户能获得的数据库条目数为 $\overline{n}=2.44$，而不诚实 Alice 借助对多个量子比特联合地进行最优无错区分可获得多达 500 个条目，这显然十分不利于保护数据库安全性。另一方面，不诚实测量攻击很难被检测出，这主要是由于 Alice 作为参与者借助量子存储能够逃避一般的检测。例如，Alice 将收到的量子比特存储起来以便后面对其执行不诚实测量攻击时，若 Bob 提出让 Alice 公布一些测量结果来检验其是否诚实，Alice 可以从存储器中取出这些检验粒子执行规定的测量来回应 Bob，对余下的粒子依然可以执行不诚实测量攻击。因此，一般的检测方法对这种参与者攻击无效。

不诚实测量攻击本质上是一种延迟测量攻击（图 1），即 Alice 必须存储载体态直到从 Bob 那里得到相应的 B 信息才能进行不诚实测量攻击。要阻止用户的这种攻击，就需要将执行延迟测量需要的两个要素割裂开来，即让 Alice 在拥有载体态时不能获得 B 信息，而在获得 B 信息时不能拥有载体态。2016 年，Wei 等人[29]基于双路 QKD 提出了一个 QPQ 方案，该方案能够阻止用户的延迟测量，从而能够有效地抵抗联合测量攻击。但是这种双向传输量子态的协议容易遭受 Trojan 木马攻击的威胁。虽然通过放置滤波器等方式可以抵御部分 Trojan 木马攻击，但想要阻止所有类型的 Trojan 木马攻击是不现实的。为此，采用让用户 Alice 预先对测量结果进行部分披露（预承诺）的方法来阻止用户的延迟测量，从而达到抵抗不诚实测量攻击的目的。具体来说，若 Alice 没有测量量子比特而是将其存储起来以便执行不诚实测量攻击，或者没有诚实地测量量子比特，则无法披露正确的信息，其欺骗行为会在后面的检测中被 Bob 发现。只有在 Bob 检测确认 Alice 的预承诺值无误后，他才会公开 B 信息。因此，Alice 无法同时拥有载体态和 B 信息。该协议无须双路 QKD，传输距离更远且不会引入 Trojan 木马攻击的威胁，也降低了用户的操作难度，为抵抗用户的不诚实测量攻击提供了更为实用、安全的方法。

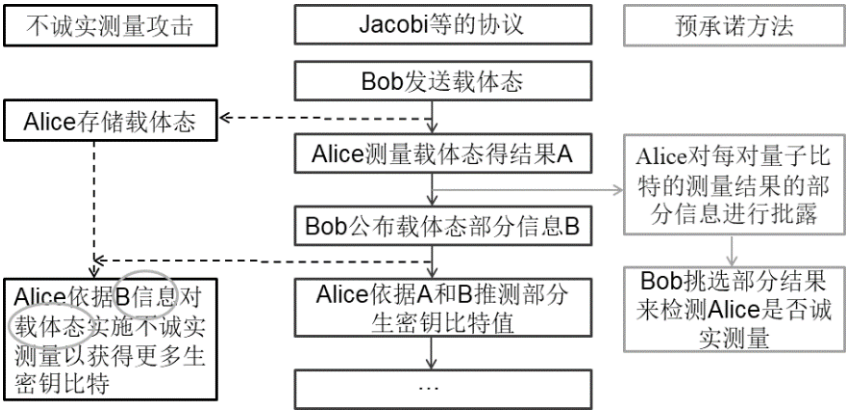

图 1　不诚实测量攻击方式及抗用户不诚实测量攻击的实现思路

3.2　具有 IDS-ZF 属性的 QPQ 协议

自 Jacobi 等人提出第一个基于 QKD 的 QPQ 方案起，人们一直渴望实现理想的数据库安全性和零失败（Ideal Database Security and Zero Failure，IDS-ZF）概率。但是，在已有的实用 QPQ 方案中，高数据库安全性通常伴随着高失败概率；反之亦然。例如，在 J 方案[4]中考虑一个 10^5 比特的数据库时，在后处理的按位相加过程中，如果选择 7 个生密钥比特相加来获得 1 个最终密钥比特，则 Alice 一次检索平均可从数据库中获得 6.10 比特，同时失败概率达到 0.002（参见文献[4]中表 1），但如果将 8 个生密钥比特累加来获得 1 个最终密钥比特，则 Alice 一次检索可获得的比特数将降至 $10^5\times0.25^8=1.53$ 个，但失败概率将会增至 $(1-0.25^8)^{10^5}=21.74\%$。为打破这一僵局，Liu 等人[27]基于 RRDPS-QKD 方案给出了一种解决方案（RRDPS-QPQ 方案），但是该方案需要借助现有条件下难以获取的理想单光子源，且当数据库规模很大时，需要传递稳定的单光子长脉冲序列，因此实现起来较为困难。更重要的是，在更为常见的基于 BB84 类 QKD 提出的 QPQ 方案中，该问题依然没有得到解决。2015 年，Gao 等人[42]提出了“移位相加”的后处理方法，可在压缩 Alice 优势的同时保证失败概率为 0。但在该方法中，不诚实 Alice 可以选择最有利于自己的移位，使得即便将 20 多个生密钥比特相加来得到 1 个最终密钥比特，也很难将 Alice 获得的最终密钥比特数降到 5 个以下，即理想的数据库安全性不能高效地实现。

2018 年，Wei 等人[43]分析了 Liu 等人的方案[27]在弱相干光源下的安全性，并提出了一种改进方案（图 2）。该方案中仅需使用弱相干光源，且传递短脉冲序列（l 值较小）作为载体态，因此实现起来较为容易。更为重要的是，由于方案中传递的脉冲序列较短，使得在后处理中可以采用一种“窄移位相加”的方法来快速压缩 Alice 获得的比特数。例如，当数据库长

度 N=8、脉冲序列长度 l+1=5 时（图 2），假定 Alice 想要得到第 6 个数据库条目 x_6，她借助干涉线路在每 4 个生密钥比特中可以确切地获得 1 个比特，因此移位数值可在较小的范围里（–3~+3 之间）选择。在图 2 中，她为 3 个子比特串分别声明移位 1、﹣1 和 0 使得移位后她获得的比特处于第 6 个位置（标灰处）。显然，她能得到第 6 个最终密钥比特并直接用它恢复出 x_6。

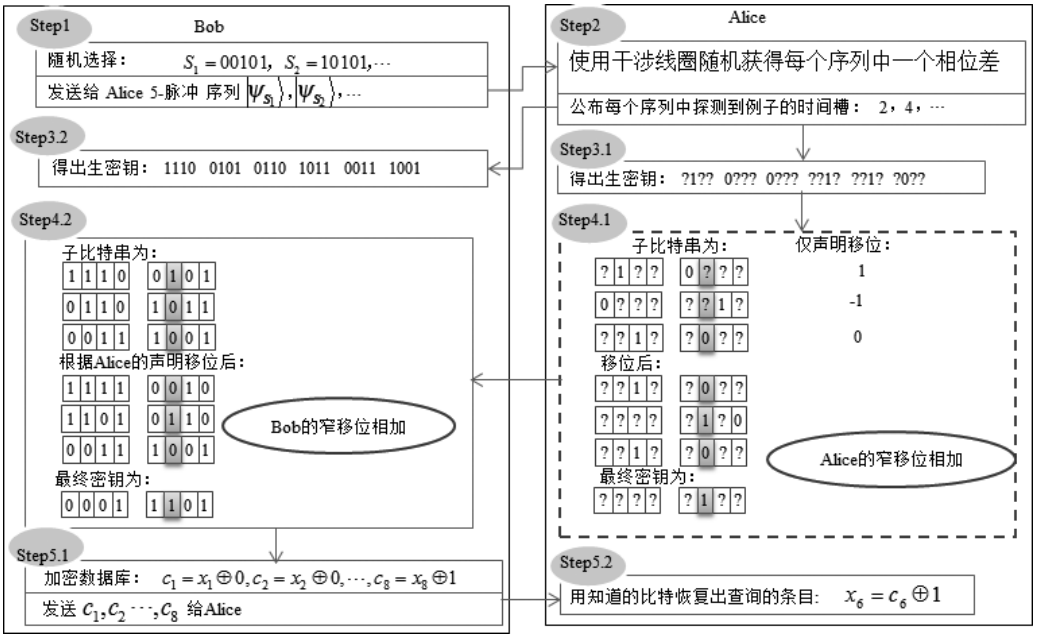

图 2　改进的 PPDPS-QPQ 方案概图（“？”表示 Alice 的未知比特）

事实上，“窄”的移位意味着 Alice 选择移位的范围显著缩小，从“移位相加”中的{0,1,3,…, n–1}降为{–l,–(l–1),…,–2,–1,0,1,3,…,l}（注意，数据库长度 N 一般远大于脉冲序列长度 l+1），故 Alice 优势显著下降。因此“窄移位相加”可以快速压缩用户获得的最终密钥比特数，实现理想的数据库安全性（图 3），且保证失败概率为 0。该协议打破了实用 QPQ 协议中高数据库安全性总是伴随着高失败概率的僵局。而且，通过在一般的基于 QKD 的 QPQ 中增加一个“块筛选”过程来保证用户在每 l 个生密钥比特中至少获得 1 个比特，就可将“窄移位相加”技术应用于所有实用 QPQ 协议的后处理过程，得到具有 IDS-ZF 属性的实用 QPQ 协议的一般性构造[43]。

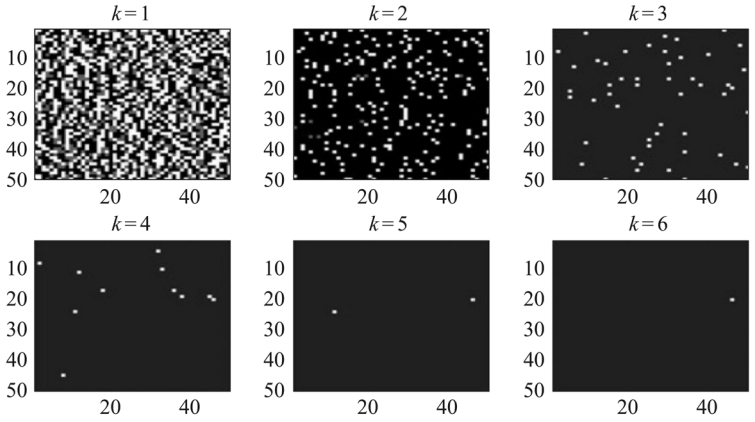

图 3　弱相干光源下压缩 Alice 优势的窄移位相加过程的仿真（这里，k 表示按位相加的生密钥比特数，脉冲长度 l=8，数据库长度 N=2500，且脉冲序列的光子数均值为 0.1。图中方框表示数据库，其中每个条目都用一个小方块表示，分别用黑色、白色、灰色表示未知条目、已知条目、奇偶相关条目。显然，Alice 获得的比特数及奇偶相关比特数均随 k 的增加而快速减少。当 k=6 时，Alice 仅知道 1 个最终比特，实现了理想的数据库安全性）

3.3　有噪信道下的 QPQ 协议

多数已有的 QPQ 协议无法适用于信道存在噪声的环境中。用户 Alice 在噪声环境中检索到的数据库条目可能会出错，而且两个参与方都可以借助噪声来掩盖自己的欺骗，不利于保护参与者的隐私。因此，处理噪声问题需要全面考虑纠错、用户隐私和数据库安全性。然而，目前仅有的两个针对 QPQ 的纠错方案[42,44]都缺少这种全面的考虑，而且它们也没有估计出多大程度的错误率能够被容忍。事实上，在量子两方安全计算中，噪声问题很少被讨论。据我们所知，到目前为止相关研究尚未得到关于可容忍错误率的结论。

为解决该问题，Wei 等人[45]提出了一个实用的 QPQ 协议（如图 4 所示），其主要贡献如下。

（1）分析了目前仅有的两个针对 QPQ 的纠错方法，发现它们没有全面考虑纠错和对双方隐私的保护，至少一个参与方的隐私面临泄露的威胁。

（2）找到了在噪声环境下区分外部攻击者和内部攻击者的方法。在无噪声的情形下，用户隐私是在“欺骗敏感”的意义下被保护的，即若 Bob 通过欺骗来获取用户隐私，他可能会提供错误的数据库条目给 Alice，从而被 Alice 以非零概率发现。但是，当噪声存在时，数据库条目出现的错误既可能来源于不诚实 Bob 的欺骗，又可能源于信道噪声，因此不能直接将其归咎于 Bob 的欺骗。为解决该问题，在噪声环境下，将“欺骗敏感”安全性赋予如下现实意义：①如果 Bob 的欺骗引入的错误率小于ε（ε是生密钥错误率的上界，可被提前估计），

Bob 将不能获得充分的优势来提取用户隐私；②如果他引入的错误率高于ε，Bob 的欺骗将会以显著的概率被发现。在这一修正的欺骗敏感安全性下讨论噪声问题，且从信息论角度量化了 Bob 的优势，以便可以判断它对于提取用户隐私是不是“充分”的。

（3）提出适用于噪声环境的 QPQ 方案，该协议在后处理过程中增加了一个“筛错”（图 4 中 Step10）的步骤，通过筛错检验错误率可将外部窃听行为和不诚实 Bob 的欺骗进行区分。它能够适用于有噪信道，也就是说，通过筛错一方面可以显著降低用户检索的数据库条目的出错概率；另一方面能够实时地检验 Bob 是否诚实（以往的协议大多靠事后发现检索条目出错来判定 Bob 欺骗，不能实时检验 Bob 是否诚实）。通过选取合适的阈值ε可以兼顾对双方隐私的保护。

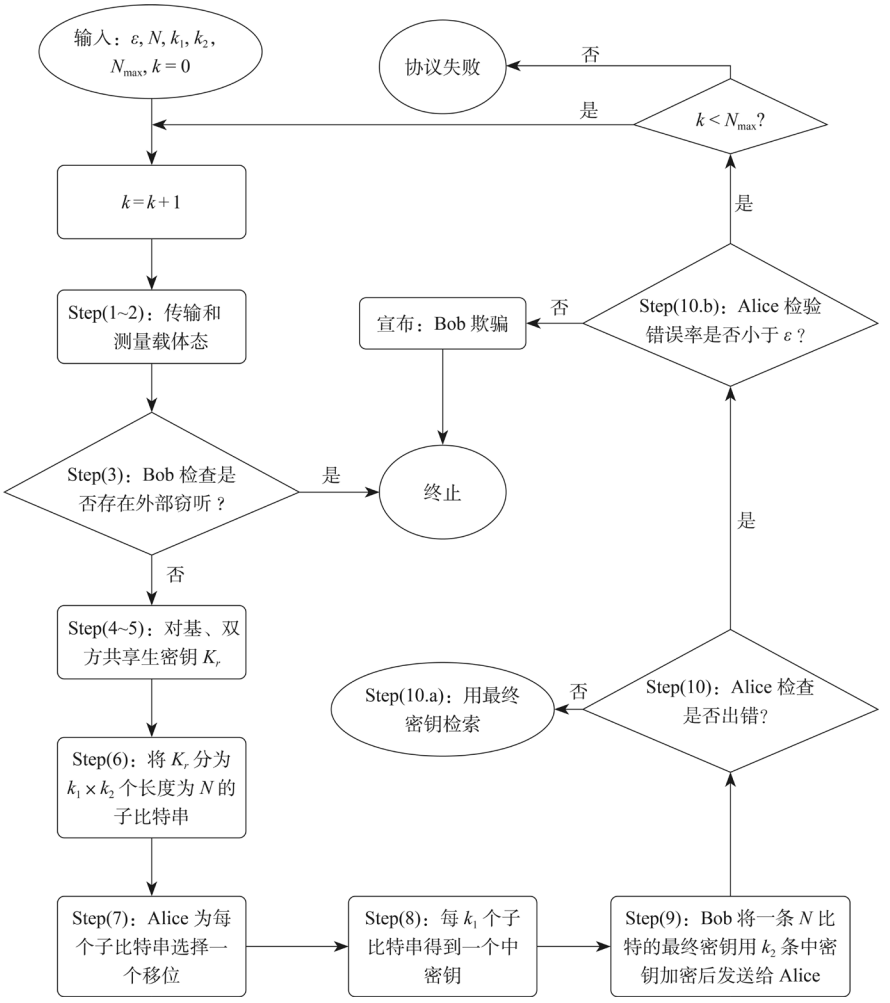

图 4　协议的简要框图

（4）分析得出最终的数据库条目的出错概率、用户隐私安全性和数据库安全性，三者两两处于一种“此消彼长”的关系中。根据具体的安全性、可靠性要求来均衡它们，可以得到可容忍错误率的上界。例如，当数据库长度 N=10000 时，假定人们希望：①不诚实 Alice 通过个体攻击至多可以获得 60 个数据库条目，通过联合测量攻击至多获得 600 个数据库条目；②不诚实 Bob 关于检索地址（总共包含 13.2877 比特）至多可得到 4 比特信息；③检索的数据库条目的最终出错概率被降至 0.01 以下。那么，为同时满足这些目标，应使得 ε<0.04。

3.4　量子匿名认证密钥交换协议

OT 是一种基础密码原语，它在经典密码中可被用于实现各类安全多方计算任务。然而，它的量子方案 QOT（包括 QPQ）至今仍很少被用于构造其他密码方案，即在量子密码中 QOT 的基础构件角色尚未被充分挖掘。这主要存在两个方面的原因。①当 QOT 用于实现其他密码任务时，一般要求其传输的信息是多比特的，而现阶段量子不经意传输方案中一般传递的信息是 1 个比特。传递多比特信息的量子不经意块传输（Quantum Oblivious Block Transfer，QOBT）方案目前还较少。②仅有的几个“多传 1”的 QOBT 方案[22-24]还很难被用于构造其他密码方案，因为他们还存在如下问题。

（1）接收者通常可以获得几条消息而不是仅获得一条消息，这样的信息泄露虽不严重但也导致它难以应用于一些对隐私保护要求较高的密码任务。

（2）发送者如果试图欺骗来推测“用户获得了哪条消息”，他的欺骗行为通常不能被实时发现，而只能在协议结束后当接收者发现自己没有得到正确的消息时才被发现（可能导致在构造其他密码协议时对某些隐私的保护也是非实时的）。

为解决上述问题，注意到 Liu 等人[27]于 2015 年基于 RRDPS QKD 提出的不经意密钥传输方案中 Alice 能从 Bob 传送的 N 个比特中精确地获得 1 个比特，无失败概率，魏春艳等人基于该方案设计了一个量子匿名认证密钥交换协议[50]。该协议能够实现用户和服务器的双向认证，且满足用户匿名性和会话密钥安全性。此外，若不诚实服务器方想要获取用户身份，其攻击行为要么无法奏效，要么能够与外部窃听区分开，从而被用户识别并认定为欺骗，因此服务器一般不会冒着名誉受损的风险来实施欺骗。

另外，所有的 QOBT 方案均不能容忍信道噪声，为解决该问题，借助 BB84 类的 QKD 协议设计了一个近似精准“N 传 1”的 QOBT 方案。该方案借助 Lagrange 插值来编码传递的信息，使得接收者以接近于 1 的概率仅获得 1 条消息，该方案能够实时检测发送者的欺骗，容忍轻度的信道噪声。最后，借助该 QOBT 方案也构造了一个匿名认证密钥交换协议，实现了双向认证性、用户匿名性和会话密钥安全性。

这一结果表明，尽管量子不经意传输不够理想，但仍可被用来构造其他密码协议，发挥基础构件的作用。该成果已投稿至 Physical Review A。

4 总结与展望

本文简要介绍了人们在 QPQ / QOT 协议构造和应用推广方面取得的研究进展。这些结果表明，QPQ 有望成为 QKD 外一类新的可实用化的量子密码协议。尽管如此，我们认为这类协议在真正走向实用的过程中，还存在如下需要进一步研究的问题。

第一，尽管现有方案分别解决了实用性、安全性方面的几个突出问题，包括对抗信道损失、容忍不完美光源、打破“高安全性总伴随着高失败概率”的僵局、抵抗用户的不诚实测量攻击、对发送方的欺骗进行“实时检测”、实现精准“多传 1”的不经意块传输等，但是能够同时解决这些问题的方案尚未被提出。如何同时、全面地解决这些问题是一个富有挑战性的课题。

第二，现有 QPQ 协议对信道噪声的容忍是有限的，能够容忍的噪声水平远低于 QKD 协议，对块传输协议来说这个问题尤为突出。因此，在现有技术条件下，这类协议难以实现远距离的可靠通信。因此，能否设计可容忍更大噪声的 QPQ/QOT 协议，或者发掘实用的量子纠错码来对抗信道噪声，是关系到 QPQ/QOT 乃至量子安全两方计算能否真正实用的关键问题。

第三，现有 QPQ 方案仅支持基于地址（用户提前获知自己要检索的条目所处的位置）的查询，尚未实现基于“关键词”的查询。鉴于现在人们在访问数据库或搜索引擎时多使用“关键词”检索的方式，设计出基于“关键词”查询的 QPQ 方案是量子保密“块查询”研究中十分有意义的课题。

第四，量子不经意传输目前还很少被用于实现具有公开验证需求的密码任务。这主要是因为量子公钥的发展还很不充分。现阶段的验证要么需要指定验证人，要么以泄露部分隐私（如认证密钥、口令）为代价，这对于某些应用来说是不被允许或十分不便的。因此，能否设计出可公开验证的量子公钥是关系到 QPQ/QOT 应用范围大小的关键问题。

参考文献

[1] SHOR P W. Algorithms for quantum computation: discrete logarithms and factoring [A]. Proceeding of the 35th Annual Symposium on the Foundations of Computer Science[C]. Santa Fe, New Mexico, U.S.A. Los Alamitos; IEEE Comp. Soc.Press,. 1994: 124-134.

[2] GROVER L K. A fast quantum mechanical algorithm for database search[A]. Proceeding of the 28th Annual ACM Symposium on Theory of Computing[C]. Philadephia, Pennsylvania, U.S.A.

New York; ACM Press. 1996: 212-219.

[3] BENNETT C H, BRASSARD G. Quantum cryptography: Public key distribution and coin tossing[A]. Proceeding of the IEEE International Conference on Computers Systems and Signal Processing[C]. Bangalore, India 1984: 175-179.

[4] JACOBI M, SIMON C, GISIN N, et al. Practical private database queries based on a quantum-key-distribution protocol[J]. Physical Review A, 2011, 83(2): 022301.

[5] GAO F, QIN S J, HUANG W, et al. Quantum private query: A new kind of practical quantum cryptographic protocol[J]. Science China-Physics Mechanics & Astronomy, 2019, 62(7): 070301.

[6] GERTNER Y, ISHAI Y, KUSHILEVITZ E, et al. Protecting data privacy in private information retrieval schemes[J]. Journal of Computer and System Sciences, 2000, 60(3): 592-629.

[7] LO H K, CHAU H F. Is quantum bit commitment really possible?[J]. Physical Review Letters, 1997, 78(17): 3410-3413.

[8] MAYERS D. Unconditionally secure quantum bit commitment is impossible[J]. Physical Review Letters, 1997, 78(17): 3414-3417.

[9] D'ARIANO G M, KRETSCHMANN D, SCHLINGEMANN D, et al. Reexamination of quantum bit commitment: The possible and the impossible[J]. Physical Review A, 2007, 76(3): 032328.

[10] LO H K. Insecurity of quantum secure computations[J]. Physical Review A, 1997, 56(2): 1154-1162.

[11] GIOVANNETTI V, LLOYDS S, MACCONE L. Quantum private queries[J]. Physical Review Letters, 2008, 100(23).

[12] GIOVANNETTI V, LLOYD S, MACCONE L. Quantum Private Queries: Security Analysis[J]. IEEE Transactions on Information Theory, 2010, 56(7): 3465-3477.

[13] OLEJNIK L. Secure quantum private information retrieval using phase-encoded queries [J]. Physical Review A, 2011, 84(2).

[14] WANG C, HAO L, ZHAO L J. Implementation of Quantum Private Queries Using Nuclear Magnetic Resonance [J]. Chinese Physics Letters, 2011, 28(8): 080302.

[15] YU F, QIU D W. Coding-Based Quantum Private Database Query Using Entangle ment[J]. Quantum Information & Computation, 2014, 14(1-2): 91-106.

[16] SCARANI V, ACÍN A, RIBORDY G, et al. Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations[J]. Physical Review Letters, 2004, 92(5): 057901.

[17] GAO F, LIU B, WEN Q Y, et al. Flexible quantum private queries based on quantum key

distribution[J]. Optics Express, 2012, 20(16): 17411-17420.

[18] YANG Y G, SUN S J, XU P, et al. Flexible protocol for quantum private query based on B92 protocol[J]. Quantum Information Processing, 2014, 13(3): 805-813.

[19] ZHANG J L, GUO F Z, GAO F, et al. Private database queries based on counterfactual quantum key distribution[J]. Physical Review A, 2013, 88(2): 022334.

[20] YANG Y G, ZHANG M O, YANG R. Private database queries using one quantum state[J]. Quantum Information Processing, 2015, 14(3): 1017-1024.

[21] XU S W, SUN Y, LIN S. Quantum private query based on single-photon interference[J]. Quantum Information Processing, 2016, 15(8): 3301-3310.

[22] WEI C Y, GAO F, WEN Q Y, et al. Practical quantum private query of blocks based on unbalanced-state Bennett-Brassard-1984 quantum-key-distribution protocol[J]. Scientific Reports, 2014, 4(7537).

[23] SHI W X, LIU X T, WANG J, et al. Multi-Bit Quantum Private Query[J]. Com munications in Theoretical Physics, 2015, 64(3): 299-304.

[24] PEI T R, MENG X L, WEI C Y, et al. Practical quantum private query of blocks based on the two-dimensional QKD system [J]. Quantum Information Processing, 2019, 18(8): 240.

[25] YANG Y G, SUN S J, TIAN J, et al. Secure quantum private query with real-time security check[J]. Optik, 2014, 125(19): 5538-5541.

[26] YU F, QIU D W, SITU H Z, et al. Enhancing user privacy in SARG04-based private database query protocols[J]. Quantum Information Processing, 2015, 14(11): 4201-4210.

[27] LIU B, GAO F, HUANG W, et al. QKD-based quantum private query without a failure probability[J]. Science China-Physics Mechanics & Astronomy, 2015, 58(10): 100301.

[28] LI J, YANG Y G, CHEN X B, et al. Practical Quantum Private Database Queries Based on Passive Round-Robin Differential Phase-shift Quantum Key Distribution[J]. Scientific Reports, 2016(6)31738.

[29] WEI C Y, WANG T Y, GAO F. Practical quantum private query with better performance in resisting joint-measurement attack[J]. Physical Review A, 2016, 93(4): 042318.

[30] LIU B, GAO Z F, XIAO D, et al. QKD-Based Quantum Private Query Protocol in the Single-Photon Interference Communication System[J]. IEEE Access, 2019, (7):104749-104758.

[31] YAN L L, LIU D M, ZHANG S B, et al. Practical Quantum Database Private Query Protocol with Classical Database Owner[J]. International Journal of Theoretical Physics, 2020, 59(9): 3002-3008.

[32] WANG T Y, WANG S Y, MA J F. Robust Quantum Private Queries[J]. International Journal

of Theoretical Physics, 2016, 55(7): 3309-3317.

[33] YANG Y G, LIU Z C, CHEN X B, et al. Robust QKD-based private database queries based on alternative sequences of single-qubit measurements[J]. Science China-Physics Mechanics & Astronomy, 2017, 60(12): 120311.

[34] LI N, LI J, CHEN X B, et al. Quantum Private Query With Perfect Performance Universally Applicable Against Collective-Noise[J]. IEEE Access, 2019(7):29313 -29319.

[35] ZHAO Y, YIN Z Q, CHEN W, et al. Loss-tolerant measurement-device-independent quantum private queries[J]. Scientific Reports, 2017, (7):39733.

[36] MAITRA A, PAUL G, ROY S. Device-independent quantum private query[J]. Physical Review A, 2017, 95(4): 042344.

[37] BASAK J, MAITRA S. Clauser-Horne-Shimony-Holt versus three-party pseudo- telepathy: on the optimal number of samples in device-independent quantum private query[J]. Quantum Information Processing, 2018, 17(4): 77.

[38] ROY S, MAITRA A, Mukhopadhyay S. Measurement-device-independent quantum private query with qutrits[J]. International Journal of Quantum Information, 2018, 16(5): 1850045.

[39] SHEN D S, ZHU X C, MA W P, et al. Improvement on private database queries based on the quantum key distribution[J]. Journal of Optoelectronics and Advanced Materials, 2012, 14(5-6): 504-510.

[40] PANDURANGA RAO M V, JAKOBI M. Towards communication-efficient quantum oblivious key distribution[J]. Physical Review A, 2013, 87(1): 012331.

[41] YANG Y G, LIU Z C, CHEN X B, et al. Novel classical post-processing for quantum key distribution-based quantum private query[J]. Quantum Information Processing, 2016, 15(9): 3833-3840.

[42] GAO F, LIU B, HUANG W, et al. Postprocessing of the Oblivious Key in Quantum Private Query[J]. IEEE Journal of Selected Topics in Quantum Electronics, 2015, 21(3): 98-108.

[43] WEI C Y, CAI X Q, LIU B, et al. A Generic Construction of Quantum-Oblivious -Key-Transfer-Based Private Query with Ideal Database Security and Zero Failure [J]. IEEE Transactions on Computers, 2018, 67(1): 2-8.

[44] CHAN P, LUCIO-MARTINEZ I, MO X F, et al. Performing private database queries in a real-world environment using a quantum protocol [J]. Scientific Reports, 2014, 4(5233).

[45] WEI C Y, CAI X Q, WANG T Y, et al. Error Tolerance Bound in QKD-Based Quantum Private Query [J]. IEEE Journal on Selected Areas in Communications, 2020, 38(3): 517-527.

[46] LI N, LI J, CHEN X B, et al. Quantum Wireless Network Private Query With Multiple

Third Parties[J]. IEEE Access, 2019(7): 33964-33969.

[47] KON W Y, LIM C C W. Provably Secure Symmetric Private Information Retrieval with Quantum Cryptography[J]. Entropy, 2021, 23(1): 54.

[48] XU M, SHI R H, LUO Z Y, et al. Nearest private query based on quantum oblivious key distribution[J]. Quantum Information Processing, 2017, 16(12): 286.

[49] LUO Z Y, SHI R H, XU M, et al. A Novel Quantum Solution to Privacy-Preserving Nearest Neighbor Query in Location-Based Services [J]. International Journal of Theoretical Physics, 2018, 57(4): 1049-1059.

[50] 魏春艳, 蔡晓秋, 王天银, 等. 基于量子不经意密钥传输的量子匿名认证密钥交换协议[J]. 电子与信息学报, 2020, 42(2): 341-347.

中国密码学会

中国密码学会是由密码学及相关领域的科技工作者和单位自愿结成并依法登记的全国性、学术性、非营利性法人社会团体，是中国科协组成部分。经民政部批准于2007年3月成立，业务主管单位中国科协，挂靠单位国家密码管理局。

中国密码学会积极参加国家创新体系建设，大力推动密码研究和学术交流，促进密码科技人才成长和进步，推动产学研用结合，传播密码科技知识；团结联系全国密码科技工作者，为密码学科和我国密码事业发展贡献力量。

截至2022年年底，学会设有17个分支机构，包括5个工作委员会：组织工作委员会、学术工作委员会、教育与科普工作委员会、青年工作委员会、密码应用工作委员会；12个专业委员会：量子密码专业委员会、密码数学理论专业委员会、密码芯片专业委员会、密码算法专业委员会、电子认证专业委员会、安全协议专业委员会、混沌保密通信专业委员会、密码测评专业委员会、区块链专业委员会、商用密码应用安全性评估联委会、大数据与人工智能安全专业委员会、物联网密码专业委员会。拥有个人会员4710人，单位会员295家。

在社会各界的关心支持下，中国密码学会已成为我国密码领域最具影响力的全国性科技社团。