超晶格密钥分发后处理技术研究

作者姓名 解建国

指导教师姓名、职称 陈小明 教授

申请学位类别 工学硕士

学校代码 10018 学 号 20199101
分 类 号 TN82 密 级 公开

北京电子科技学院

硕士学位论文

超晶格密钥分发后处理技术研究

作者姓名：解建国

一级学科：网络空间安全

二级学科（研究方向）：网络空间安全

学位类别：工学硕士

指导教师姓名、职称：陈小明 教授

提交日期：2022 年 5 月

Research on Post-processing Technology in Semiconductor Superlattice Secure Key Distribution

A thesis submitted to

Beijing Electronic Science and Technology Institute

in partial fulfillment of the requirements

for the degree of Master

in Cyber Science and Engineering

By

Jianguo Xie

Supervisor: Xiaoming Chen Title:Professor

May 2022

北京电子科技学院
学位论文独创性（或创新性）声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得北京电子科技学院或其它教育机构的学位或证书而使用过的材料。与我一同工作的同事对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文若有不实之处，本人承担一切法律责任。

本人签名：_____　　日　　期：_____

北京电子科技学院
关于论文使用授权的说明

本人完全了解北京电子科技学院有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权属于北京电子科技学院。学校有权保留送交论文的复印件，允许查阅、借阅论文；学校可以公布论文的全部或部分内容，允许采用影印、缩印或其它复制手段保存论文。同时本人保证，结合学位论文研究成果完成的论文、发明专利等成果，署名单位为北京电子科技学院。

保密的学位论文在_____年解密后适用本授权书。

本人签名：_____　　导师签名：_____

日　　期：_____　　日　　期：_____

摘要

密码是国之重器，是国家安全三大支撑技术之一，是网络空间安全体系的基石，对个人乃至国家网络信息安全意义重大。

本文的主要研究内容和创新性如下：

（1）

（2）

关键词：密钥协商， 后量子， 物理不可克隆函数， 纠错码， 极小熵

ABSTRACT

As one of the three supporting technologies for national security，cryptography is the most important weapon of the country, and the cornerstone of the cyberspace security system, which is of great significance to personal and even national network information security.

The main research contents and innovations of this paper are as follows:

(1)

(2)

Keywords: Key aggrement, Post-quantum, PUF, Error correction codes, Min-entropy

插图索引

表格索引

符号对照表

符号	符号名称
$\in$	属于
n	纠错码码长
k	纠错码信息位长度
t	纠错码纠错能力
ϵ	安全参数
Ext	提取器
$\mathbf{SD}(X,Y)$	统计距离
H	校验矩阵
G	生成矩阵
G	生成矩阵
W_B	全周期波形
$ErrorIndex$	错误组索引
$\mathscr{M}$	度量空间
P	状态转移矩阵
$\{X_n\}_{n\in N}$	随机过程
H_∞	极小熵
U_l	均匀分布
$p(x)$	反馈多项式

缩略语对照表

缩略语	英文全称	中文对照
SSL-SKD	Semiconductor superlattice secure key distribution	超晶格密钥分发
SSL	Semiconductor Superlattice	半导体超晶格
MBE	Molecular Beam Epitaxy	分子束外延
PQC	Post-Quantum Cryptography	后量子密码
OTP	One time pad	一次一密
QKD	Quantum key distribution	量子密钥分发
PUF	Physical unclonable functions	物理不可克隆函数
ECC	Error correction codes	纠错码
LLR	Log-Likelihood Ratio	对数似然比
FFT	Fast Fourier Transformation	快速傅里叶变换
LDPC	Low-density parity-check	低密度奇偶校验
ADC	Analog to Digital Converter	模数转换器
DFT	Discrete Fourier Transform	离散傅里叶变换
DAC	Digital to Analog Converter	数模转换器
PA	Privacy amplification	保密增强
FER	Frame Error Rate	误帧率
SC	Successive Cancellation	串行抵消
UHF	Universal Hash functions	一致哈希函数簇
LFSR	Linear Feedback Shift Register	线性反馈移位寄存器
DMA	Direct Memory Access	直接存储器读取
SG	Scatter Gather	分散收集
SKR	Secure Key Rate	安全成码率
BSC	Binary symmetric channel	二元对称信道
BEC	Binary erasure channel	二元擦除信道
BI-AWGN	Binary input additive white Gaussian noise channel	二元高斯加性白噪声信道
MCD	Multi-Codeword decoding	多码字并行译码

目录

第一章 绪 论

沉舟侧畔千帆过，病树前头万木春。

—白居易

密码技术从来都不是现代科技的产物，它伴随着信息、数据的发展以多种形态应运而生。信任是数据共享、开放的基础，密码助力打造数据共享信任价值链。现代密码学通常将密码系统的整体安全性归结于密钥体系的安全性，因而衍生出以计算理论为基础的计算安全和以物理密码为代表的信息论安全。超晶格密钥分发（Semiconductor superlattice secure key distribution, SSL-SKD）是一种基于超晶格器件的物理不可克隆孪生理论衍生出的信息论安全的密钥分发方式，为密钥管理的痛点难题提供了一种新的实用化解决方案。

1.1 密码学：历史，现在，未来

密码技术的雏形最早可追溯到古埃及的墓碑铭文，发展至今大致可分为古典密码、近代密码和现代密码三个阶段[1,2]。在古代，密码技术主要用于军事目的，古希腊著名的凯撒（Caesar）密码、史巴达（Syta）密码以及中国古代的虎符，都是军事密码的典型例子[3]。为了进行加密，Caesar 密码将消息中的每个字母（明文）替换为移动一定位数（密钥）的另一个字母（密文）。解密就是加密的反向操作以此恢复出明文。很显然，该密码的可能密钥数量为 26，因此，一个未经授权的人可以在暴力攻击中测试所有可能的密钥，以恢复明文。甚至，通过对密文的频率分析，可以进行更加快速有效的攻击。近代密码的特点是替代传统手工密码，利用机械或机电进行加解密，因此也被称为机电密码时代，以美国的 M209 密码机和德国的 ENIGMA 密码机为代表[4,5]。现代密码的学术研究始于计算机的出现，如今，密码学已经成为日常生活中不可或缺的一部分，所有基于互联网的活动都有密码学的参与，如简单的 Web 浏览，或银行系统中的汇款信息。计算机密码时代根据加解密使用的密钥是否相同将密码技术主要分为两类：对称密码学和非对称密码学[6]。

在对称密码中，合法双方使用相同的密钥进行加解密。数据加密标准（DES）是 IBM 于 20 世纪 70 年代初设计的一种对称加密算法，并于 1977 年由美国国家安全局（NSA）修改，作为商业应用的加密标准[7]。典型的对称密码还包括 2000 年美国国家标准与技术研究院公布的高级加密标准 AES[8]、国际数据加密算法 IDEA[9] 等。对称密码体系本质上是以复杂但确定的方式混合密钥和明文。加密过程通常由相同

的操作进行多轮运算，在每轮中，子密钥和明文通过由一系列替换和置换算法混合在一起。在这种情况下，安全性基于以下假设：由于加密的复杂性，最佳攻击是对密钥的全搜索（密钥的长度决定安全性）。而在对称密码体系中由于加密和解密需要共用密钥，高速高质量的密钥分发是痛点难题。在图 1.1 中，我们显示了 IDEA 的一轮加密。

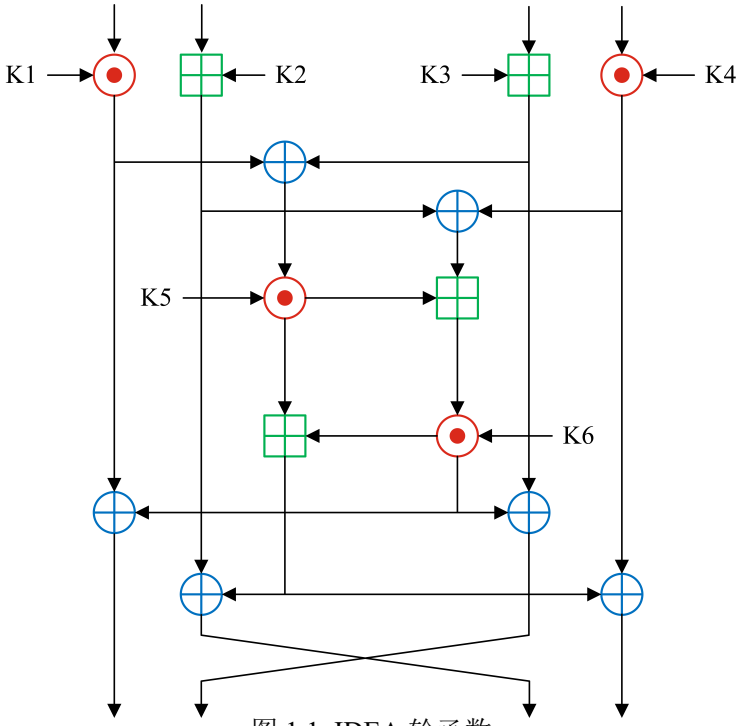

图 1.1 IDEA 轮函数

对于非对称密码（公钥密码）体系，加密和解密使用不同的密钥，公钥用于加密，私钥用于解密[10]。只有拥有私钥的人才能解密使用公钥加密的消息，而对于不拥有私钥的人来说，解密任务是非常复杂的（计算上不可行的）。原则上 Rivest Shamir Adleman（RSA）[11]，Elgamal[12] 和 ECDSA[13] 等非对称加密算法都是基于数学难解问题（整数分解，离散对数和椭圆曲线）的，这意味着这些算法的安全性取决于窃听者的计算能力。和对称加密相比，公钥密码计算复杂，在加密和解密过程都会浪费大量时间。图 1.2 显示了一个典型的公钥密码系统。

在实际使用中，为了平衡对称密码和非对称密码的优缺点，双方之间进行通信通常使用混合加密系统[14]。首先使用非对称加密在彼此之间达成共享会话密钥（如 Diffe-Hellman 密钥交换协议），然后使用对称加密传输实际数据。混合加密体制已经安全高效的运行了许多年，但是量子计算的快速发展给以数学问题复杂性为基础的公钥密码体制和以密钥长度为安全基础的对称密码体制带来了灾难性的影响，将使得目前主流的密码体制变得不堪一击[15]。

针对当前密码体制面临的量子计算威胁，近年来美国国家标准与技术研究院（Na-

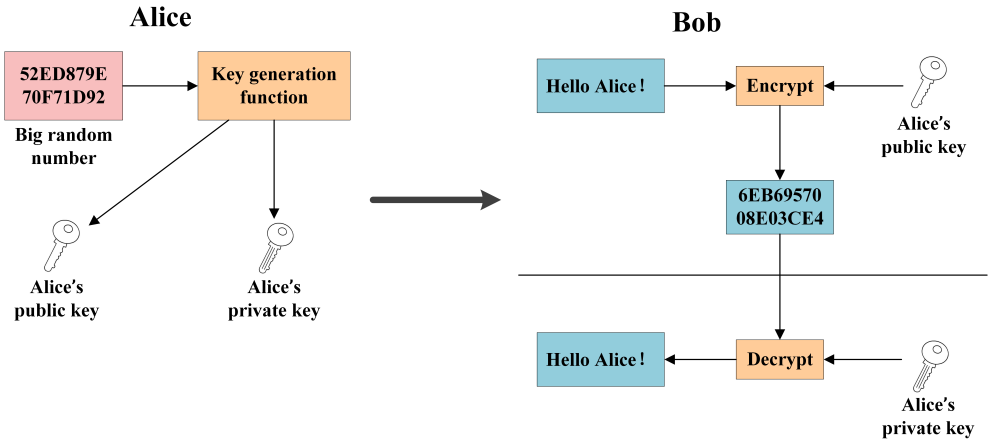

图 1.2 典型公钥密码系统

tional Institute of Standards and Technology, NIST）已经开始组织设计后量子密码技术（Post-Quantum Cryptography, PQC）[16]。后量子密码技术即探索在量子计算条件下的非线性多项式时间可解的 NP 数学问题，典型例子如格基加密算法和基于哈希的加密算法等。无论是经典密码体制，抑或是 PQC，其安全性基础都建立在计算资源有限条件下的数学难解问题，无法满足香农提出的信息论安全（Information-theoretical security）级别的信息加密服务。信息论安全通信由信息论安全的密钥分发和信息论安全的加密算法来保证，1949 年，C. E. Shannon 首次系统阐述了消息使用一次一密（One time pad, OTP）算法可以满足无条件安全通信的要求，并指出该算法所使用密钥的熵必须大于等于明文的熵[17]。

物理密码技术的出现打破了传统密码技术的发展瓶颈，并迅速成为国际上信息安全技术的研究热点，以达到信息论防护级别的量子密钥分发（Quantum key distribution, QKD）和物理不可克隆函数（Physical unclonable functions, PUF）为典型代表。量子密钥分发于 1984 年 C. H. Bennett 和 G. Brassard 首次提出，基于量子力学的基本特征（不可克隆定理，量子纠缠，测不准原理）为 Alice 和 Bob 提供了信息论安全的密钥交换[18,19]。因此，Alice 和 Bob 可以基于量子态的传输和测量来共享密钥，而 Eve 在不引入扰动的情况下无法提取有关密钥的任何信息[20,21]。量子密钥分发由于其中继转发和经典认证等问题，目前还难以大面积实用化，但是其理论研究和工程实践的成果不仅推动了物理密码的发展，还带来了量子信息领域的变革[22–25]。物理不可克隆函数最早可追溯至 2001 年 R. Pappu 等人提出的物理单向函数（Physical one-way functions, POF）概念，是一种基于物理、生物微观结构构建 POF 的思想[26]，并由 B. Gassend 等人实现第一个硅基集成电路 PUF[27]。PUF 是一种架构在物理实体上的函数关系，挑战响应由其特殊的物理结构决定，并且特定输入对应唯一输出，具备唯一性、稳定性且不可预测，PUF 的不可克隆特性来源于制造（成长）过程中的不可避免

的微小偏差[24,28]。

半导体超晶格密钥分发是基于孪生超晶格 PUF 驱动混沌同步的一种新型信息论安全密钥分发方案，超晶格密钥分发技术是一种物理随机数的异地共享技术拓展。使用孪生超晶格 PUF 物理实体的预分配代替密钥信息的预分配，由器件不可克隆避免了密钥管理中的敏感信息泄露风险；并且，部署完成的孪生超晶格密钥分发设备，能在公开信道中建立密钥分发协议机制，降低密钥部署和更新的成本，以高速、便捷、随遇组网等特点逐渐成为国内研究热点[29,30]。

1.2 超晶格密钥分发

1.2.1 超晶格研究历史

XXXXXXX

1.2.2 超晶格物理不可克隆孪生机理

XXXXXXX

1.2.3 基于孪生超晶格的密钥分发方案

基于超晶格 PUF 的物理不可克隆孪生特性，陈小明等人提出了利用该特性实现点对点密钥分发的构想[31]。超晶格孪生器件在密码学上可以认为是在异地各自运行的同一密钥同一单向函数，且攻击者无法克隆器件也无法建模器件的任意运行过程及结果。因此利用孪生超晶格器件可以安全实现物理随机数的异地排他性共享，即密钥分发[30,32]。密钥协商的原理架构如图 1.3 所示，可简述如下：

(1) 定义 Alice 为密钥产生（重建）端，Bob 为密钥重建（产生）端，Alice 和 Bob 在常规通信信道上协商激励信号 C（可以由第三方负责发送，也可由其中一方传给另一方）；

(2) Alice 和 Bob 分别在各自的超晶格 PUF 中输入 C 得到原始输出数据，对原始输出进行高精准离线序列同步得到响应 W 和 W'，W 和 W' 汉明距离大约为 10%；

(3) Alice 计算辅助数据 H，并将 H 通过公开信道传送给 Bob，Bob 在 H 的帮助下和 W' 计算出 W，此过程即为信息调同（Information reconciliation）[33]；

(4) Alice、Bob 双方根据 H 泄露的信息和器件输出的极小熵，从 W 中提取出相同密钥 R，此过程即为保密增强（Privacy amplification, PA）[34]。

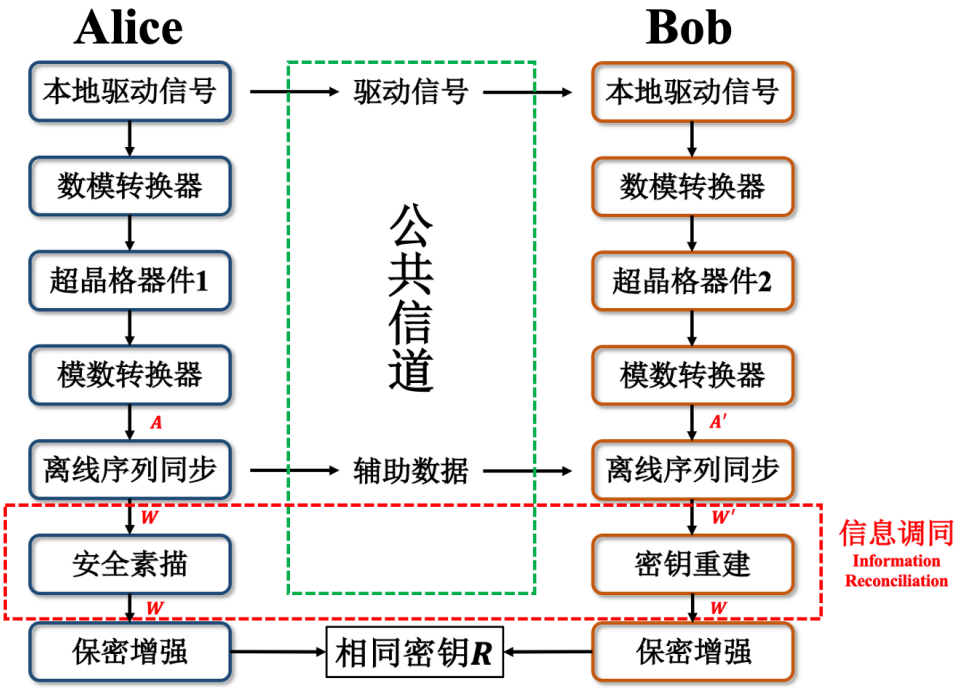

图 1.3 基于超晶格 PUF 的密钥分发协议

信息调同用来描述从 W' 恢复出 W 的过程，此过程不会泄漏很多关于 W 的信息，一般用安全素描实现。保密增强的一般实现方案为一致哈希函数（Universal hash functions, UHF）提取器实现[35]，该函数的输入为非满熵的待提取的随机数和公开的满熵随机数，输出保密的满熵随机数。信息调同和保密增强统称模糊提取器（Fuzzy extractor），可以从物理、生物等唯一特征数据中提取出可用作加密、认证的密钥。

1.3 论文主要工作与结构安排

1.3.1 主要工作

本文的主要研究内容是超晶格密钥分发系统的后处理部分，研究面向实际应用的高精准序列同步算法、高吞吐率的信息调同算法、无条件安全保密增强算法，并开展了相关实验验证。本文对超晶格密钥分发系统的高速高效后处理算法进行了深入研究，取得的主要创新工作如下：

(1) 研究超晶格密钥分发过程中通信双方的高精准序列同步技术。XXXXXXX

(2) 研究高吞吐率的信息调同技术。XXXXXXX

(3) 研究超晶格密钥分发无条件安全保密增强技术。XXXXXXX

(4) 超晶格密钥分发系统安全密钥分发实验。XXXXXXX

1.3.2 结构安排

本文结构组织如下：

第一章，绪论，即本章。从现代密码学中的密钥管理问题介绍了现今物理密码体系的发展和突破，进而引出超晶格密钥分发技术，介绍了超晶格密钥分发系统后处理技术的研究背景、意义以及本文的主要研究内容。

第二章，超晶格密钥分发后处理技术基础，主要介绍 XX。

第三章，高精准序列同步技术研究。XXXXXXX。

第四章，高吞吐率的信息调同算法研究。XXXXXXX

第五章，信息论安全的保密增强算法研究。XXXXXXX

第六章，异地实时超晶格密钥分发实验。XXXXXXX

第七章，总结与展望。对全文工作进行总结，并对未来研究作出展望。

第二章　超晶格密钥分发后处理算法基础

合抱之木，生于毫末；九层之台，起于累土；千里之行，始于足下。

—老子

本章介绍了一些基本定义、示例等用于描述基本概念和基础知识。其中包括信息论基础、信道编码基础以及一致哈希函数提取器等相关内容。

2.1　熵

香农于 1948 年在“通信的数学原理”一文中首次将物理热力学中熵的概念引入信息论中，因而迎来信息论学科研究的热点。信息是指某一过程中包含的不确定性，而熵就是用来度量这种不确定性的指标，它反映了在观察一个值之前预测该值的不确定性（熵值）[36]。

定义 2.1. ***(香农熵，Shannon Entropy)***：设随机变量 X 的取值为 $x_1, x_2, ..., x_n$，与之对应的分布概率为 $p_1, p_2, ..., p_n$。则 X 的香农熵为

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i. \tag{2-1}$$

相比较于上述香农熵的讨论，本文更关注 Rényi 熵。Rényi 熵是 Alfred Rényi 于 1976 年提出的对香农熵、碰撞熵、极小熵的推广。在后文的保密增强算法中，二阶 Rényi 熵是能从理论上被证明可以提取出无条件安全的密钥，而相关文献也报道了关于利用香农熵提取密钥的安全风险问题。

定义 2.2. ***(Rényi 熵，Rényi Entropy)***：设随机变量 X 的取值为 $x_1, x_2, ..., x_n$，与之对应的概率为 $p_1, p_2, ..., p_n$。则 X 的 *Rényi* 熵为

$$H_\alpha(X) = \frac{1}{1-\alpha} \log(\sum_{i=1}^{n} p_i^\alpha). \tag{2-2}$$

当 $\alpha = 1$ 时，上式表示 Shannon 熵，当 $\alpha = 2$ 时，即二阶 Rényi 熵，也称作碰撞熵，此时式中的概率表示为碰撞概率 $p_c(i)$。但是很多物理系统计算其碰撞熵是相当复杂的，因此目前许多文献都表明计算极小熵用于提取无条件安全的密钥方案是切实可行的。当 $\alpha = \infty$ 时，定义极小熵如下：

定义 2.3. ***(极小熵，Min-Entropy)***：设随机变量 X 的可预测性用 $\max_x P[X=x]$ 表示，那么随机变量 X 的极小熵定义为

$$H_\infty(X) = -\log(\max_x P[X=x]). \tag{2-3}$$

通常情况下，平均极小熵更为准确，因为实际系统中攻击者有各种因素在平均情况下获得与随机变量 X 不独立的事件 Y。

定义 2.4. ***(平均极小熵，Average Min-Entropy)***：给一对随机变量 X 和 Y，假如对手知道随机变量 Y 中的 y，则变量 X 的可预测性用 $\max_x P[X=x|Y=y]$ 表示, 所以对手以 $E_{y\leftarrow Y}(\max_x P[X=x|Y=y])$ 预测 X，那么变量 X 的平均极小熵定义为

$$\tilde{H}_\infty(X|Y) = -\log(E_{y\leftarrow Y}(\max_x P[X=x|Y=y])) = -\log(E_{y\leftarrow Y}(2^{-H_\infty(X|Y=y)})). \tag{2-4}$$

2.2 信道编码基础

2.2.1 纠错码介绍

基于上述互信息量的介绍，我们可以分析有噪信道下的信息保真传输问题，即香农第二定理，有噪信道编码定理，设信道的输入为 X，输出为 Y，首先定义信道容量为 $C = \max_{P(X)} I(X;Y)$。

定理 2.5. ***(有噪信道编码定理，Noisy channel coding theorem)***：每个信道都有自己对应的信道容量 C，当信息编码码率 $R \leq C$ 且码长足够长的情况下，总可以找到一个码字，可以使用最大似然译码，使得误码率随着码长 N 的增加而减小至趋于 *0*。

香农提出有噪信道编码定理之后，各种纠错码实现方案如雨后竹笋般出现[37]，追求香农为我们设定的天花板。

XXXX。

2.2.2 信道模型

用 $\mathscr{X}$ 表示信道输入，$\mathscr{Y}$ 表示信道输出。对于离散信道模型[38]，我们用离散概率 $Pr_{Y|X}(y|x)$ 描述随机变量 X 和 Y 的信道模型。对于连续信道模型，我们用条件密度 $f_{Y|X}(y|x)$ 表示。如果 $|\mathscr{X}|=2$ 表示二元输入信道，即 $\mathscr{X}=\{-1,+1\}$ 或 $\mathscr{X}=\{0,1\}$。

定义 2.6. ***(二元擦除信道，Binary erasure channel, BEC)***：参数为 ϵ 的二元擦除信道表示为 $BEC(\epsilon)$。信道输入端的随机变量 X 可以取值 $x \in X = \{-1,+1\}$，信道输出端

的随机变量 Y 可以取值 $y \in Y = \{-1, ?, +1\}$。转移概率表示如下：

$$Pr_{Y|X}(y|x) = \begin{cases} 1-\epsilon, & y = x, \\ \epsilon, & y = ?, \\ 0, & otherwise \end{cases} \tag{2-5}$$

XXX。

2.2.3 BCH 码

XXXX。

2.2.4 LDPC 码

低密度奇偶校验（Low-density parity-check, LDPC）码是由 Gallager 于 1963 年提出[39]，并且证明了 LDPC 码的性能接近信道容量且非常易于实现。但是 LDPC 码在当时由于计算机计算能力受限和存储较小等问题导致当时的人们认为该码是不符合实际的（impractical）。在上世纪 90 年代由于计算机存储和计算技术的蓬勃发展，LDPC 码被 MacKay，Luby 等人重新提出，从而进入研究的热点，并很快成为深空通信，移动通信，卫星通信等信道编码标准[40]。

XXXX。

2.2.5 Polar 码

Polar 码一种基于信道极化理论的线性分组码，由土耳其 Bilkent 大学教授 Erdal Arikan 于 2007 年提出[41]，应用 Polar 码时，首先选定码长 n 和信息位长度 k，然后需要对所选的信道进行极化，此时可以确定 Polar 码放入编译码结构。Polar 码的编码过程和 BCH 码等代数编码思想一致，译码过程采用串行抵消译码算法，时间复杂度为 $O(n \log n)$。Polar 码是目前唯一被证明在 $n \to \infty$ 时、利用串行抵消译码算法性能达到香农限的纠错码。

XXXX。

2.3 提取器基础知识

2.3.1 强提取器和统计距离

随机性提取也叫做提取器，是指从弱安全的原始随机序列提取出与均匀分布统计距离可忽略的无条件安全随机数的函数。提取器自提出之后便得到广泛研究，无论是在密码学上还是计算机中的分布式计算都实现了大规模应用，尤其在物理随机数发生器的后处理过程。由于弱提取器的密钥不可公开导致无法开展实际应用，本文叙

述的提取器若不具体说明均为强提取器。首先定义两个分布的统计距离如下：

定义 2.7. ***(统计距离，Statistical Distance)***：设 X 和 Y 是集合 $\mathscr{M}$ 上的两个随机变量，X 和 Y 之间的统计距离定义为：

$$\mathbf{SD}(X,Y)=\frac{1}{2}\sum_{\omega}|Pr[X=\omega]-Pr[Y=\omega]|. \tag{2-6}$$

如果 X 和 Y 之间的统计距离 $\mathbf{SD}(X,Y)\leq\epsilon$，则称 X 和 Y 是 ϵ 统计接近的，如果 ϵ 是可忽略的，则称 X 和 Y 是统计不可区分的。如果对于提取出的随机数与均匀分布的统计距离 ϵ 是可忽略的，那么即可认为提取器输出的随机数是可用的。

定义 2.8. ***(强提取器，Strong Extractor)***：对于式子 $Ext:\{0,1\}^n\times\{0,1\}^r\ \rightarrow\{0,1\}^l$，如果对于所有的最小熵为 m 的分布 W, 有

$$\mathbf{SD}((Ext(W;X),X),(U_l,X))\leq\epsilon, \tag{2-7}$$

其中 X 是 $\{0,1\}^r$ 上的均匀分布，U_l 表示 l 比特的均匀分布的随机数，则 Ext 是一个 (n,m,l,ϵ) 强提取器

2.3.2 一致哈希函数

经典密码学中的哈希函数通常具有单向性、雪崩性等特点，可作为提取器使用，但是单个哈希函数对于所有的输入无法满足提取器的要求。一致哈希函数簇是一族哈希函数的统称，每次均匀的从一族哈希函数中选择一个应用于待提取样本。

2.4 本章小结

本章介绍了超晶格密钥分发后处理技术的相关基础知识，包括信息论基础，信道编码基础，提取器相关内容介绍，这部分内容为信息调同、保密增强等后处理技术打下了坚实的基础。XX。

第三章　后处理方案改进及高精准序列同步算法研究

昨夜西风凋碧树，独上高楼，望尽天涯路。

—晏殊

超晶格密钥分发过程中通信双方得到一致的驱动序列后，激励孪生超晶格产生相似的原始数字序列。

3.1　在线序列同步技术

3.1.1　原理方案

XXXX。

3.1.2　实验结果与分析

XXXXXXXX。

3.2　高精准离线序列同步技术

3.2.1　原理方案

由于在线同步算法的计算复杂度和通信负载在实际应用中难以接受，因此本文设计了一种追峰算法来进行超晶格输出序列的高精准同步，无需进行通信交互。

XXXX。

3.2.2　实验结果与分析

XXXX。

3.3　超晶格密钥分发协议实现方案改进

XXXXX。

3.4　本章小结

本章首先研究了在线同步算法在超晶格 PUF 激励响应的序列同步问题的应用，并进行了相关实验评估了在线同步算法的准确性、效率、资源消耗等。

第四章　高吞吐率的信息调同算法研究

衣带渐宽终不悔，为伊消得人憔悴。

—柳永

经过高精准离线序列同步后，Alice 和 Bob 间的比特串由于模拟系统不可避免的差异仍然有不一致的地方，所以需要通过信息调同技术使两边的比特序列达成一致。

4.1　信息调同问题模型

信息调同问题最早由 Bennett 提出并应用于量子密钥分发[25]，随后 Brassard 提出信息调同的公开讨论模型使其正式步入实用化[42]。

定义 4.1. ***(信息调同的信道编码等效模型，Equivalent model of information reconciliation as channel coding)***：$\mathscr{M}$ 是具有距离函数 dis 的度量空间（*Metric Space*），一个参数为 $(\mathscr{M}, m, \tilde{m}, t)$ 的信息调同算法分为“产生”（*GEN*）和“重建”（*REC*）两个阶段，具有如下三个性质：

- 在产生阶段，输入 $w \in \mathscr{M}$，输出辅助数据 $GEN(w) = h \in \{0,1\}^*$；
- 在重建阶段，输入 $w' \in \mathscr{M}$ 和产生阶段输出的辅助数据 h，在保证 $dis(w, w') \leq t$ 的前提下，能够得到 $REC(w', h) = w$；
- 安全性保障：对 $\mathscr{M}$ 上的任意分布 W 都有极小熵 m，W 的值可以被第三方通过观察辅助数据 s 获得的信息不超过 $2^{-\tilde{m}}$，即 $\tilde{H}_\infty(W|h) \geq \tilde{m}$。

XXXXX。

4.2　基于 BCH 码的高速纠错方案

4.2.1　编译码方案

BCH 码是一种典型的线性分组码，在码长较短时具有出色的性能表现。他的结构简单，易于实现，在资源受限的场景下很受欢迎。因此目前常见的 PUF 认证系统以及大多数 SSD 控制器大多采用 BCH 码作为其纠错方案。BCH 码在码长 n，信息位长度 k，和纠错能力 t 之间存在严格的代数关系。对于任何的正整数 $m \geq 3$ 并且

$t < 2^{(m-1)}$，都会存在一个如式 (4-1) 所示参数的一个二元 BCH 码。

$$\begin{cases} n = 2^m - 1 \\ n - k \leq mt \\ d_{min} \geq 2t + 1 \end{cases} \tag{4-1}$$

算法 4.1: 优化后的 BCH 码译码器

Data: $R(x)$, LT, ALT, $blockNum$
Result: 译码码字 C
Initialize parameters, $\sigma^{(0)}(x) = 1$, $D(0) = 0$, $d_0 = s_1$, $\sigma^{(-\frac{1}{2})}(x) = 1$, $D(-\frac{1}{2}) = 0$, $d_{-\frac{1}{2}} = 1$;
#pragma omp parallel for shared(LT, ALT) firstprivate($R(x)$);
threadNum = 0;
while *threadNum* < $blockNum$ **do**
    计算 syndrome polynomial $S = \{s_1, s_2, \cdots, s_{2t}\} = \{R(\alpha), R(\alpha^2), \cdots, R(\alpha^{2t-1})\}$;
    从 BM 算法计算错误位置多项式;
    while $j < 2t$ **do**
        if $d_j == 0$ **then**
            $\sigma^{j+1}(x) = \sigma^j(x)$;
            $D(j+1) = D(j)$;
        end
        else
            计算 d_{j+1};
            $\sigma^{j+1}(x) = \sigma^j(x) - d_j d_i^{-1} x^{j-i} \sigma^i(x)$;
        end
        一直迭代直到计算出 $\sigma^t(x)$;
    end
    从 Chien 搜索算法计算 $\sigma(x)$ 的根，根取反即错误位置;
    threadNum = threadNum + 1;
end
The output is codeword C;

XXXXXX。

4.2.2 实验结果与分析

本文在室温实验室环境下采集孪生超晶格 PUF 的输出，通过高精准序列同步算法后量化为二元序列，并按照码长 $n = 4095$ 进行分组总共得到 120,000 组数据。

XXXXX。

4.3　基于 LDPC 码的高速纠错方案

在上一节中我们提出了基于 BCH 码的超晶格信息调同高速纠错方案，但是 BCH 码在码长较长时性能表现一般，速度差强人意，无法满足更高速率的超晶格密钥分发系统的需求，因此需要更高效的纠错方案来加快系统吞吐率。

4.3.1　编译码方案

1963 年，Robert G. Gallager 的博士论文中论述了校验矩阵为稀疏矩阵的纠错码，进而创新性的提出一种新的线性分组码，LDPC 码。刚开始提出由于当时计算机的计算性能的限制和理论研究的现状导致当时被认为是不符合实际的，使其无法得到长足发展。MacKay，Luby 在上世纪 90 年代重新将 LDPC 码进一步发展[43]，随着计算机存储和计算性能的逐渐发展 LDPC 码逐渐成为研究热点方向，在目前计算机的高速发展的今天，LDPC 码以其在各种信道上提供接近香农限的性能立足在各大传输和存储设备的纠错编码的应用上。

XXXXXX。

4.3.2　实验结果与分析

同上一节 BCH 码优化方案测试吞吐率使用相同的超晶格数据，对于设计的这四种 LDPC 码，按照码长进行分组我们同样测试了 120,000 组码字数据。

XXXXX。

4.4　基于 Polar 码的高速纠错方案

Polar 码是一种由信道极化理论演化而来的新型线性分组纠错码，于 2007 年被 Arikan 提出。Polar 码由于其优异的性能一经提出便受到广泛关注，对于超晶格密钥分发系统也是如此，Polar 码的诸多特性使其对超晶格密钥分发系统的纠错方案极其匹配。首先，Polar 码是第一个在理论上被证明可以达到香农限的纠错码，是一类具有较低的编码和译码复杂度的纠错码，尤其在码长较长时表现优异。再者，和 LDPC 码不同的是，Polar 码和 BCH 码一样容易构造，只要确定了码长和码率其结构就可以确定。此外，Polar 码的另外一个重要特点是其译码器的规则递归结构，使得译码器的软件实现速度大大高于其他纠错码[?]。

4.4.1　编译码方案

XXXXX。

4.4.2 实验结果与分析

XXXX。

4.5 分析比较

XXXX。

4.6 本章小结

在本章中，我们提出了一种针对超晶格密钥分发系统的多线程高吞吐率的信息调同纠错方案，本文提出的纠错优化方案可以高效地应用于不同的场景。XXXX。

第五章　信息论安全的保密增强算法研究

众里寻他千百度，蓦然回首，那人却在，灯火阑珊处。

—辛弃疾

保密增强是超晶格密钥分发系统实现安全密钥提取的至关重要的步骤，其目标是剔除超晶格原始输出不足熵以及后处理过程中攻击者可能获取的部分密钥串信息，并生成相对于攻击者而言信息论安全的密钥。

5.1　保密增强问题模型

XXXXXX。

5.2　极小熵评估

5.2.1　定义和假设

XXXXXX。

5.2.2　基于超晶格 PUF 的随机模型

XXXXXXXXX。

5.2.3　极小熵结果与分析

在本节中，我们在室温实验室环境下采集多组超晶格 PUF 的原始输出数据，对提出的极小熵估计算法进行了多组实验以得到准确的评估结果。

$$P=\begin{bmatrix}0.088235 & 0.058824 & 0.147059 & 0.147059 & 0.235294 & 0.117647 & 0.029412 & 0.176471\\ 0.205882 & 0.088235 & 0.235294 & 0.000000 & 0.058824 & 0.205882 & 0.088235 & 0.117647\\ 0.073171 & 0.170732 & 0.170732 & 0.146341 & 0.048780 & 0.073171 & 0.170732 & 0.146341\\ 0.147059 & 0.088235 & 0.147059 & 0.117647 & 0.117647 & 0.147059 & 0.117647 & 0.117647\\ 0.113636 & 0.090909 & 0.113636 & 0.113636 & 0.295455 & 0.068182 & 0.136364 & 0.068182\\ 0.114286 & 0.085714 & 0.142857 & 0.171429 & 0.057143 & 0.085714 & 0.257143 & 0.085714\\ 0.078947 & 0.184211 & 0.078947 & 0.131579 & 0.157895 & 0.052632 & 0.131579 & 0.184211\\ 0.102564 & 0.128205 & 0.102564 & 0.076923 & 0.179487 & 0.205128 & 0.076923 & 0.128205\end{bmatrix}$$

5.3 保密增强算法实现方案

5.3.1 输出密钥长度估算

XXXXXX。

定理 5.1. 假设函数簇 $\{H_x : \{0,1\}^n \to \{0,1\}^l\}_{x\in X}$ 是一致哈希函数簇，信息调同过程的纠错码为 (n,k,t)，任意随机输入 W 的极小熵为 m，那么 $l \le m-(n-k)-2\log\frac{1}{\epsilon}+2$。

证明. 首先由剩余哈希引理可知，对任意的随机输入 W，有

$$SD((H_x(W),X),(U_l,X)) \le \frac{1}{2\sqrt{2^{-H_\infty(W)+l}}}, \tag{5-1}$$

其中，U_l 表示 $\{0,1\}^l$ 上的均匀分布。又 $\frac{1}{2\sqrt{2^{-H_\infty(W)+l}}} \le \epsilon$ 可知

$$l \le m - 2\log\frac{1}{\epsilon} + 2, \tag{5-2}$$

其中 m 表示平均条件下的 W 的极小熵，上述定理说明保密增强的熵损失为 $2\log\frac{1}{\epsilon}-2$ 的比特数，即最多能够提取出 $m-2\log\frac{1}{\epsilon}+2$ 长度的满熵比特。

下面再证明信息调同过程中的熵损失为 $(n-k)$ 即可。首先给出一个引理以便后续证明需要。

引理 5.2. 假设 A,B,C 是随机变量，此时如果 B 至多有 2^λ 种可能的值，那么

$$\hat{H}_\infty(A|(B,C)) \ge \hat{H}_\infty((A,B)|C) - \lambda \ge \hat{H}_\infty(A|C) - \lambda, \tag{5-3}$$

特别的，

$$\hat{H}_\infty(A|B) \ge \hat{H}_\infty(A|B) - \lambda \ge H_\infty(A) - \lambda. \tag{5-4}$$

因此，如果公开的 helper data 至多有 2^λ 种可能的值，对于任意的信息调同的输入 W，都有

$$\hat{H}_\infty(W|hd) \ge H_\infty(W) - \lambda. \tag{5-5}$$

即表明熵损失最多为 λ。已知 n 表示纠错码码长，k 表示纠错码的信息位的长度，helper data 的长度也为 n。

那么由上述引理可知，

$$\hat{H}_\infty(W|(hd,I)) \ge \hat{H}_\infty((W,hd)|I) - n, \tag{5-6}$$

又因，

$$P(W=w,hd=s|I=i) \le \frac{1}{2^k}P(W=w|I=i). \tag{5-7}$$

利用式 (5-7) 可得，

$$\hat{H}_\infty((W, hd)|I) \geq \hat{H}_\infty(W|I) + k, \tag{5-8}$$

结合式 (5-6) 和式 (5-8) 可知，

$$\hat{H}_\infty(W|(hd, I)) \geq \hat{H}_\infty(W|I) + k - n. \tag{5-9}$$

所以，信息调同过程的熵损失为 $(n-k)$。 □

综上所述，超晶格密钥分发系统最终可提取出的密钥长度为 $l \leq m - (n-k) - 2\log\frac{1}{\epsilon} + 2$，提取出的密钥是满熵的，且该协议过程无条件安全。

5.3.2　LFSR-Toeplitz 提取器

XXXXXX。

5.3.3　基于 FFT 的加速方案

XXXXXX。

5.3.4　实验结果

XXXXXX。

5.4　本章小结

本章给出了

第六章　异地实时超晶格密钥分发实验

Stay Hungry, Stay Foolish.

—Steve Jobs

XXX

6.1　系统架构

XXXXXXXXXX。

6.2　分模块

6.2.1　驱动层

XXXXXX。

6.2.2　功能组件层

XXXXXX。

6.2.3　业务逻辑层

XXXXXXXXX。

6.3　北京–长沙异地实时超晶格密钥分发实验

XXXXXXXXXXXXX。

6.4　本章小结

本章首先叙述了

第七章　全文总结与展望

每个人都会经过这个阶段，见到一座山，就想知道山后面是什么。我很想告诉他，可能翻过去，你会发觉没什么特别，再翻过来，会觉得这边更好。但我知道他不会听，自己不走一走，又怎么会甘心。

—《东邪西毒》

7.1　全文总结

XXXX。本文首先建立了超晶格密钥分发后处理整体架构，分别对序列同步、信息调同、保密增强三个部分的内容展开了研究。本文专注于核心问题展开理论研究和应用拓展，取得的主要研究成果包括：

(1) XXX

(2) XXX

(3) XXX

(4) XXX

7.2　后续工作展望

随着超晶格密钥分发技术实用化的稳步推进，后处理阶段要研究的东西还有很多，尽管作者尽了最大努力进行了本课题的研究，但所做的工作仍然有许多不足，一些研究内容仍需要完善，并且还有如下方向值得进一步探索：

(1) XXX

(2) XXX

参考文献

[1] KATZ J, LINDELL Y. Introduction to modern cryptography[M]. London : CRC press, 2020.

[2] KAHN D. The Codebreakers[J]. New York: Scribner's, 1967.

[3] STINSON D R, PATERSON M. Cryptography: theory and practice[M]. London : CRC press, 2018.

[4] BUCHMANN J. Introduction to cryptography[M]. Berlin/Heidelberg : Springer Science & Business Media, 2013.

[5] 白炎林. 穿越时空的秘密——二战中的美军密码机 (下)[J]. 兵器, 2005(5) : 24 – 28.

[6] 施奈尔, 吴世忠. 应用密码学: 协议、算法与 C 源程序 [M]. 北京 : 机械工业出版社, 2000.

[7] DIFFIE W, HELLMAN M. Special Feature Exhaustive Cryptanalysis of the NBS Data Encryption Standard[J]. Computer, 1977 : 74 – 84.

[8] STANDARD N-F. Announcing the advanced encryption standard (AES)[J]. Federal Information Processing Standards Publication, 2001 : 1 – 51.

[9] LAI X, MASSEY J L. A proposal for a new block encryption standard[C] // Advances in Cryptology EUROCRYPT 1990. 1990 : 389 – 404.

[10] DIFFIE W, HELLMAN M. New directions in cryptography[J]. IEEE transactions on Information Theory, 1976, 22(6) : 644 – 654.

[11] RIVEST R L, SHAMIR A, ADLEMAN L. A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of the ACM, 1978 : 120 – 126.

[12] ELGAMAL T. A public key cryptosystem and a signature scheme based on discrete logarithms[J]. IEEE transactions on information theory, 1985 : 469 – 472.

[13] JOHNSON D, MENEZES A, VANSTONE S. The elliptic curve digital signature algorithm (ECDSA)[J]. International journal of information security, 2001 : 36 – 63.

[14] MENEZES A J, VAN OORSCHOT P C, VANSTONE S A. Handbook of applied cryptography[M]. London : CRC press, 2018.

[15] STEANE A. Quantum computing[J]. Reports on Progress in Physics, 1998 : 117.

[16] BERNSTEIN D J. Introduction to post-quantum cryptography[G] // Post-quantum cryptography. Berlin : Springer, 2009 : 1 – 14.

[17] SHANNON C E. Communication theory of secrecy systems[J]. The Bell system technical journal, 1949 : 656 – 715.

[18] BENNETT C H, GRASSARD G. Quantum Cryptography: Public Key Distribution and Coin Tossing[J]. Proceedings of IEEE International Conference on Computers, Systems, and Signal Process-

ing, 1984 : 175 – 179.

[19] SHOR P W, PRESKILL J. Simple proof of security of the BB84 quantum key distribution protocol[J]. Physical review letters, 2000 : 441.

[20] GISIN N, RIBORDY G, TITTEL W, et al. Quantum cryptography[J]. Reviews of modern physics, 2002 : 145.

[21] 刘博. 星地量子保密通信安全密钥提取关键技术研究 [D]. 长沙 : 国防科学技术大学, 2017.

[22] XU F, MA X, ZHANG Q, et al. Secure quantum key distribution with realistic devices[J]. Reviews of Modern Physics, 2020 : 025002.

[23] PIRANDOLA S, ANDERSEN U L, BANCHI L, et al. Advances in quantum cryptography[J]. Advances in Optics and Photonics, 2020 : 1012 – 1236.

[24] MAES R. Physically unclonable functions: Properties[G] // Physically Unclonable Functions. Berlin : Springer, 2013 : 49 – 80.

[25] BENNETT C H, BESSETTE F, BRASSARD G, et al. Experimental quantum cryptography[J]. Journal of cryptology, 1992 : 3 – 28.

[26] PAPPU, RAVIKANTH, RECHT, et al. Physical One-Way Functions[J]. Science, 2002.

[27] GASSEND B, CLARKE D, van DIJK M, et al. Silicon Physical Random Functions[J]. Association for Computing Machinery, 2002 : 148 – 160.

[28] HALAK B. Physically Unclonable Functions: From Basic Design Principles to Advanced Hardware Security Applications[M]. Berlin : Springer, 2018.

[29] LIU W, YIN Z, CHEN X, et al. A secret key distribution technique based on semiconductor superlattice chaos devices[J]. Science Bulletin, 2018 : 1034 – 1036.

[30] WU H, YIN Z, TONG X, et al. An experimental demonstration of long-haul public-channel key distribution using matched superlattice physical unclonable function pairs[J]. Science Bulletin, 2020 : 879 – 882.

[31] 童新海, 陈小明, 徐述. 超晶格密码的研究进展 [J]. 科学通报, 2020 : 10 – 18.

[32] XIE J, WU H, XIA C, et al. High throughput error correction in information reconciliation for semiconductor superlattice secure key distribution[J]. Scientific Reports, 2021, 11(1) : 1 – 9.

[33] DODIS Y, REYZIN L, SMITH A. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data[C] // International conference on the theory and applications of cryptographic techniques. 2004 : 523 – 540.

[34] BENNETT C H, BRASSARD G, CRÉPEAU C, et al. Generalized privacy amplification[J]. IEEE Transactions on Information theory, 1995, 41(6) : 1915 – 1923.

[35] CARTER J L, WEGMAN M N. Universal classes of hash functions[J]. Journal of computer and system sciences, 1979, 18(2) : 143 – 154.

[36] GRAY R M. Entropy and information theory[M]. Berlin/Heidelberg: Springer Science & Business Media, 2011.

[37] 朱雪龙. 应用信息论基础 [M]. 北京: 清华大学出版社, 2001.

[38] LIN S, COSTELLO D J. Error control coding: Vol 2[M]. New Jersey: Prentice hall New York, 2001.

[39] GALLAGER R. Low-density parity-check codes[J]. IRE Transactions on information theory, 1962, 8(1): 21–28.

[40] RYAN W, LIN S. Channel codes: classical and modern[M]. London: Cambridge university press, 2009.

[41] ARIKAN E. Systematic polar coding[J]. IEEE communications letters, 2011, 15(8): 860–862.

[42] BRASSARD G, SALVAIL L. Secret-key reconciliation by public discussion[J], 1993: 410–423.

[43] MACKAY D J, NEAL R M. Near Shannon limit performance of low density parity check codes[J]. Electronics letters, 1996, 32(18): 1645.

致谢

全情投入，守正出奇，愿等花开。

—佚名

当论文写作接近尾声时，暑天已过去大半，可南方的燥热依旧不减。所幸住在湖边不远，偶尔湖风吹过，又颇有一点“明月别枝惊鹊，清风半夜鸣蝉”之感。回想起这近二十载求学生涯，如梦一场，恍如昨日。时光里有少年的不羁和浪荡，有青春的颓废和迷茫，也有成熟之后的坦然和温暖，白驹过隙，事实上想来，只是寥寥数语。

此刻最为感激的当数硕士生涯为我执灯的恩师陈小明先生。

作者简介

1. 基本情况

男，安徽 XX 人， XXXX 年 X 月出生，北京电子科技学院密码系网络空间安全专业 2019 级硕士研究生。

2. 教育背景

2015.09～ 2019.06，XXXXXX，本科，专业：XXXXXX

2019.09～ ，北京电子科技学院，硕士研究生，专业：网络空间安全

3. 攻读硕士学位期间的研究成果

3.1 发表学术论文

[1] **Zhang san.** et al. XXXXXX. PKC, (2021).

3.2 申请（授权）专利

[1] 无

3.3 参与科研项目及获奖

[1] 无