

面向移动即时通信的量子随机数增强密钥方案研究

Research on Quantum Random Number Enhanced Key Scheme for Mobile Instant Messaging

孟小燕, 马伟, 陈颀, 侯玉华, 齐霄, 李长连(中讯邮电咨询设计院有限公司, 北京 100048)

Meng Xiaoyan, Ma Wei, Chen Si, Hou Yuhua, Qi Xiao, Li Changlian (China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China)

摘要:

针对移动即时通信端到端加密会话密钥依赖伪随机数发生器(PRNG)、熵源不足或实现缺陷导致可预测的隐患,研究量子随机数发生器(QRNG)密钥增强方案。构建终端 QRNG 芯片、云端密钥服务与加密通道 3 层架构,设计量子熵种与确定性随机比特发生器(DRBG)混合派生流程,克服速率受限问题,并给出降级策略。分析与实验表明,主密钥可预测性降至量子熵主导的可忽略量级,抗预测、抗泄露能力优于纯 PRNG 方案,可支撑政务、金融等高安全即时通信应用。

Abstract:

To address the predictability risk of end-to-end encryption session keys in mobile instant messaging (IM) arising from the entropy shortage or implementation flaws of pseudo-random number generators (PRNGs), a quantum random number generator (QRNG) key enhancement scheme is proposed. A three-layer architecture is constructed, comprising terminal QRNG chips, a cloud key service, and an encrypted channel, and a hybrid key derivation process combining quantum entropy seeds with a deterministic random bit generator (DRBG) is designed to overcome rate limitations, with a fallback strategy provided. Analysis and experiments show that the scheme reduces master key predictability to a negligible level dominated by quantum entropy, and outperforms pure PRNG solutions in predictability resistance and leakage resilience, which can support high-security IM applications in government and finance.

Keywords:

Quantum random number; Mobile instant messaging; Key derivation; End-to-end encryption; True randomness

引用格式: 孟小燕, 马伟, 陈颀, 等. 面向移动即时通信的量子随机数增强密钥方案研究[J]. 邮电设计技术, 2026(9): 42-46.

1 概述

1.1 移动即时通信加密的随机性隐患

即时通信端到端加密普遍采用 Signal 协议, 会话密钥质量取决于随机数的质量。现有终端依赖 PRNG, 其输出为确定性运算, 熵源不足或实现存在缺陷时密钥可被预测, 已有 RSA 与椭圆曲线因弱密钥导致泄露的事件^[1]。移动终端熵池易供给不足, 政务与金融等高安全场景亟需原理上不可预测的真随机源。

1.2 QRNG 研究现状与产业化进展

量子随机数发生器(QRNG)以量子态测量塌缩为物理熵源, 输出统计独立、不可预测^[2]。国盾量子、IDQ 等已推出 QRNG 板卡与芯片^[3-4], 三星与中国联通手机已集成 QRNG 用于加密即时消息^[5-7], 国家标准《器件无关量子随机数发生器通用要求》(器件无关属高端形态, 与终端芯片路线互补)已于 2025 年发布^[8]。QRNG 正沿板卡—模块—芯片路径实现小型化, 可低侵入集成。

1.3 本文研究目标与内容

针对会话密钥依赖伪随机数、存在被预测的隐患, 本文利用 QRNG 真随机特性构建可工程落地的密钥增强方案, 提出终端 QRNG 芯片、云端密钥服务与加

关键词:

量子随机数; 移动即时通信; 密钥派生; 端到端加密; 真随机

doi: 10.12045/j.issn.1007-3043.2026.09.008

文章编号: 1007-3043(2026)09-0042-05

中图分类号: TN918

文献标识码: A

开放科学(资源服务)标识码(OSID):



收稿日期: 2026-07-29

密通道的3层架构,设计量子熵种与DRBG结合的混合派生流程及降级策略,并通过安全性量化分析、方案对比与数值实验验证其有效性。

2 技术基础

2.1 即时通信端到端加密的密钥派生与随机数需求

Signal协议的密钥协商(X3DH)与消息加密(双棘轮)均高度依赖随机数^[9]。当随机数质量不足时,当前与后续会话密钥均可被预测,前向保密被削弱。

2.2 伪随机数发生器及其局限

伪随机数发生器采用DRBG结构,以种子经确定性算法扩展输出,由NIST SP 800-90A标准进行规范^[10]。其安全性依赖种子熵量,熵不足则密钥可被推断,状态泄露则后续输出全被复现^[11]。

2.3 量子随机数发生器原理与优势

量子随机数发生器以量子态测量塌缩为熵源,根据Born法则,叠加态测量结果的概率由概率幅模平方确定,单次结果原理上不可预测,其随机性可用最小熵下界量化^[2],区别于确定性算法递推的伪随机序列。QRNG已形成终端芯片化与服务端板卡化方案,可集成于密钥派生链路^[3-4]。

QRNG按物理实现分为光子分束、激光相位噪声、真空涨落等路线,其中CMOS工艺芯片化方案因体积小、成本低、易集成,成为面向移动终端的主流方向。研究表明此类芯片方案能显著提升会话密钥的不可预测性^[11],与本文思路一致。

2.4 QRNG与伪随机数发生器的对比

QRNG与PRNG在随机性来源、可预测性和工程代价等方面差异显著,其特性对比如表1所示。

3 QRNG增强密钥方案与关键流程

3.1 总体架构

本文提出面向移动即时通信的QRNG增强密钥总

体架构(见图1),该架构由终端QRNG芯片层、云端QRNG密钥服务层和即时通信加密通道层构成。其中,终端层在本地产生量子真随机数注入量子熵;云端层集中部署QRNG板卡供给密钥材料;加密通道层将量子熵与现有派生流程融合,完成密钥生成、更新与加解密。

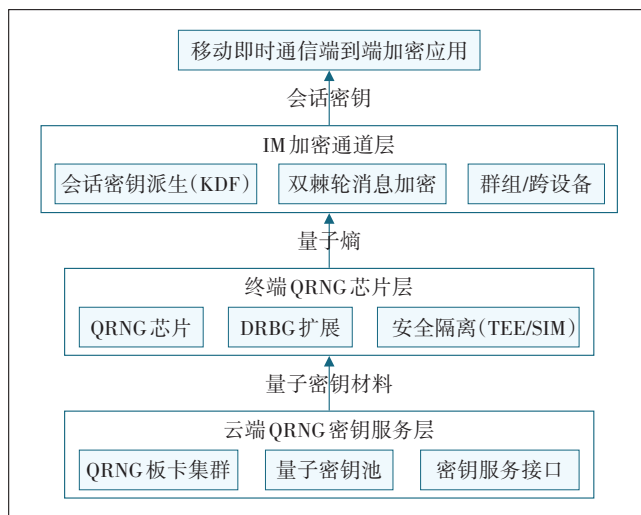


图1 面向移动即时通信的QRNG增强密钥总体架构

3.2 终端侧:QRNG芯片注入量子熵

终端集成QRNG芯片(如国盾QRNG300C或IDQ IDQ20MC1),产生量子真随机数作为高熵种子注入DRBG,以QRNG熵种+DRBG混合模式兼顾真随机质量与速率。芯片经安全隔离区(TEE[Trusted Execution Environment]或安全SIM卡)与客户端进行交互,以防止被窃取,并实时对输出进行随机性检验,剔除异常样本。

3.3 云端:QRNG密钥服务

云端QRNG密钥服务由板卡集群、密钥池与服务接口构成,批量生成量子随机数并集中供给高熵随机材料(作为终端熵种子的补充,非会话密钥本身)。材料经终端公钥加密下发,仅作为熵源参与终端侧派生,服务端不持有最终主密钥。在跨设备、群组等终端量子熵不足时,该服务可提供补充;其集中部署可摊薄成本,弥补单终端速率受限的短板^[12]。云端材料须与终端本地熵在TEE内混合后参与KDF,单点泄露不危及主密钥。

3.4 混合密钥派生与采集流程

由于单终端QRNG速率有限,该方案采用量子熵种+DRBG混合派生的方式,以QRNG周期性高熵种子

表1 QRNG与伪随机数发生器特性对比

比较维度	QRNG	PRNG/DRBG
随机性来源	量子测量	确定性算法
可预测性	原理不可预测	熵不足时可被预测
状态泄露影响	不依赖状态	泄露致全暴露
熵的可证明性	可证明	难严格证明
生成速率	受限	高
终端集成成本	需专用芯片	纯软件
成熟度	产业化	成熟

重置 DRBG, 两次重置间高速派生消息密钥。

周期 T 内 DRBG 可派生约 1 Gbit 密钥流, 足以覆盖高频消息消耗 (详见第 4.3 节)。 T 可根据速率与频次进行动态调整。

量子熵采集与混合密钥派生流程如图 2 所示。在会话建立时, 终端 QRNG 采集量子熵, 与云端经终端公钥加密下发的量子随机材料在终端 TEE 内混合后, 送入密钥派生函数 (KDF) 生成主密钥, 再由双棘轮逐条派生消息密钥。

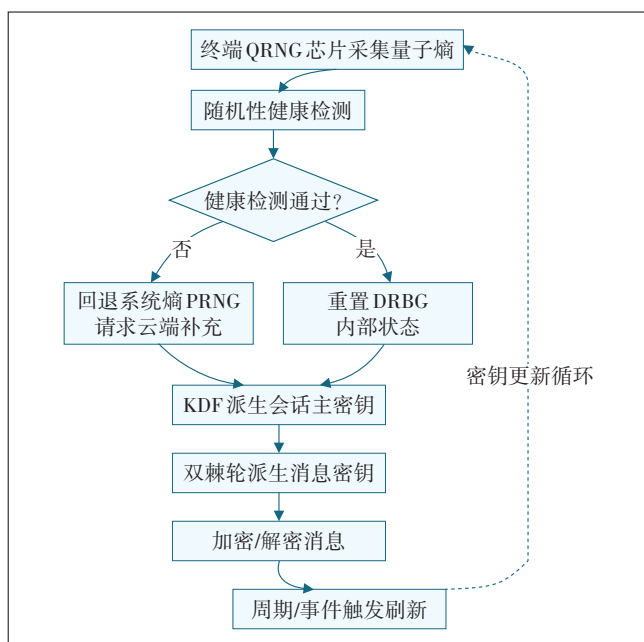


图2 量子熵采集与混合密钥派生流程

3.5 会话密钥的 QRNG 增强更新

在会话中, 终端按周期以 QRNG 新量子熵刷新 DRBG, 使消息密钥受量子真随机支撑并限制状态泄露的影响。此外, 也可以由新成员入群、设备切换等事件触发刷新。

3.6 群组消息与跨设备密钥同步

群组 IM 由群主基于主密钥派生成员子密钥进行加密分发, 成员变动时经量子熵刷新轮换密钥; 跨设备登录时由云端下发经设备公钥加密的量子随机材料, 在新设备 TEE 内与本地熵混合后派生设备子密钥。

3.7 QRNG 故障或速率不足时的降级策略

当 QRNG 芯片出现故障或速率不足时, 本方案设计了降级策略以保障通信的连续性。终端持续对 QRNG 输出执行 NIST SP 800-90B 健康检测 (重复计数

检测 RCT、自适应比例检测 APH), 连续 N 次未通过或最小熵低于阈值 $H_{\min}$ 时触发降级, 回退为系统熵池驱动的 DRBG 派生并向云端请求后量子密码 (PQC) 协商补充, 待恢复后切回。降级期间终端运行于安全等级 L2 (传统计算安全), 低于正常态 L1 (量子增强), 并在消息元数据中携带等级标记, 敏感业务可据此拒绝降级态消息。

降级态下系统熵池 (如 Linux getrandom) 仍须满足最小熵下界评估以保证基本密码强度; QRNG 恢复后连续通过 M 轮检测且熵估计达标后, 方可切回 L1。

由于每次量子熵刷新都将 DRBG 状态的不确定区间约束在一个刷新周期内, 降级期间 DRBG 状态虽不再由量子源刷新, 其影响仍被限制于降级持续时长内; 加之双棘轮前向保密机制不依赖单一随机源, 历史消息密钥安全性得以保持。

4 安全性与性能分析

4.1 安全性分析

设敌手 A 为概率多项式时间对手, 可被动观测公开信道并于某一时刻获取 DRBG 当前的内部状态; 假设云端诚实但好奇 (或可被攻陷, 即已知其下发的密钥材料), 但不得攻破量子测量的内禀随机性与底层密码原语; 终端 TEE 为信任根, TEE 不被攻破时本地量子熵与 DRBG 状态不可获取。记终端 QRNG 单次注入种子 Q 的最小熵为 $H_{\infty}(Q)$, 取 $\ell=256$ bit、 $\rho=0.9$, 则 $H_{\infty}(Q) \geq 230$ bit。

a) 性质 1: 抗密钥预测界。考虑会话主密钥 K 在一次量子熵重注入后立即派生, 且敌手所获状态不晚于该次重注入, 其猜测优势满足 $\text{Adv}^{\text{guess}}(K) \leq 2^{-H_{\infty}(Q)} + \varepsilon_{\text{DRBG}}$, 首项由最小熵定义 (最优猜测优势不超过 $2^{-H_{\infty}(\cdot)}$), $\varepsilon_{\text{DRBG}}$ 为 DRBG 伪随机性假设下的可忽略边界; 由 $H_{\infty}(Q) \geq 230$ 得 $\text{Adv}^{\text{guess}}(K) \leq 2^{-230} + \text{negl}(n)$, 可预测性由量子种子熵主导。

b) 性质 2: 前向保密与状态泄露约束。方案以 T_{refresh} 为周期向 DRBG 注入新量子熵。每次刷新后, 敌手即便获取了前一时刻的 DRBG 状态, 对新状态的预测优势仍受新量子熵的约束。单次状态泄露的可预测区间限制为 $[t_{\text{leak}}, t_{\text{leak}} + T_{\text{refresh}}]$; 此后历史消息密钥能够保持前向保密性。

4.2 与纯伪随机数方案的对比

本方案与纯伪随机数方案的对比如表 2 所示。

表2 本方案与纯伪随机数方案对比

维度	纯PRNG方案	QRNG增强方案
熵质量	依赖熵池,波动大	量子熵,可证明高熵
抗密钥预测	熵不足时可被预测	原理不可预测
抗状态泄露	泄露致全暴露	刷新,影响受限
密钥空间	受熵源限制	近理想随机
终端开销	纯软件,低	需QRNG芯片

4.3 性能分析与数值实验

首先,建立参数化性能模型。云端单板可支撑并发会话数正比于 $R_{\text{cloud}} \times T/L_K$,终端采集种子耗时为 $L_{\text{seed}}/R_{\text{QRNG}}$,参数取值如表3所示(公开指标与典型区间值)。由于量子密钥集中用于主密钥周期性派生,QRNG速率主要约束并发会话容量而非消息吞吐。

表3 性能模型参数取值

参数	取值	取值依据
终端QRNG芯片速率	典型取值为5 Mbit/s (区间为0.5~10)	手机QRNG芯片 ^[5-6] ,实测速率多在0.5~10 Mbit/s
云端板卡速率(单板)	40 Mbit/s	IDQ Quantis PCIe-40M指标 ^[4]
DRBG派生速率	1 Gbit/s	NIST SP 800-90A与CPU benchmark ^[10]
密钥长度	256 bit	AES-256会话密钥
单会话消息频次	10(区间为5~20)条/s	即时通信流量统计
刷新种子长度	256 bit	NIST SP 800-90A ^[10]
主密钥刷新周期	$T = 1 \text{ s}$	根据消息频次与种子长度进行权衡

基于该模型进行数值实验,以QRNG速率与刷新周期为扫描变量,逐一取值运行模型,以泊松过程模拟主密钥派生请求的随机到达,对密钥池水位进行蒙特卡洛仿真(固定随机种子可复现)。理论容量上界估算如表4所示,终端QRNG速率对种子采集时延的影响如图3所示,过载场景下密钥池水位演化(降级策略对比)如图4所示。

典型参数下单板可支撑约 1.56×10^5 个并发会话(集群线性扩展),终端采集256 bit的种子约需51 μs ,

表4 理论容量上界估算

测算指标	典型值
单DRBG周期可派生消息密钥数	$\approx 3.9 \times 10^6$ 条
云端单板可支撑并发会话数	$\approx 1.56 \times 10^5$
终端采集256 bit种子耗时	$\approx 51 \mu\text{s}$
方案端到端附加时延	$\approx 60 \mu\text{s}$,占网络往返时延(RTT)不足0.2%
单板容量上界 N^* (蒙特卡洛验证)	$\approx 1.56 \times 10^5$

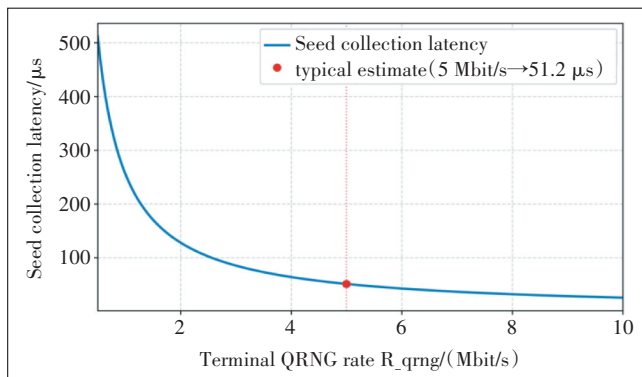


图3 终端QRNG速率对种子采集时延的影响

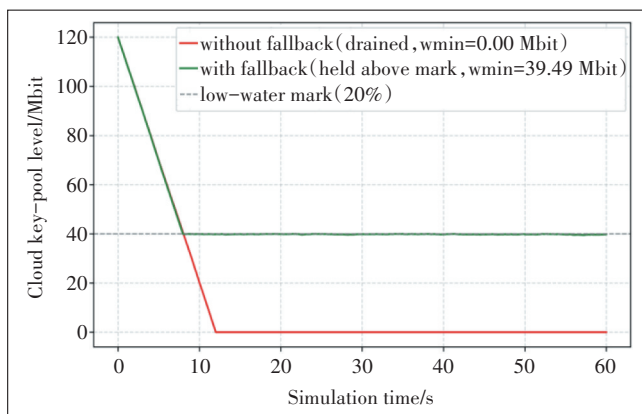


图4 过载场景下密钥池水位演化(降级策略对比)

附加时延约为60 μs ,占30 ms网络RTT不足0.2%。蒙特卡洛仿真显示,过载时降级策略可稳定密钥池水位。并发容量与 $R_{\text{cloud}} \cdot T_{\text{refresh}}$ 成正比,与 L_K 成反比(当速率从5 Mbit/s减低至0.5 Mbit/s时,时延增长10倍,增长后仍约为0.5 ms)。以上为基于公开指标的理论上界,实际性能待真机验证。

5 应用场景、产业实践与展望

5.1 典型应用场景

先存储后解密(Harvest Now Decrypt Later, HNDL)攻击对长保密数据的威胁主要由后量子密码(PQC)协商应对,而会话密钥可预测与状态泄露风险由量子真随机应对。本方案聚焦后者,并与PQC互补,适用场景如表5所示。

政务与金融移动展业即时通信承载涉密公文等高价值内容,是本方案的典型场景。方案部署采取“云端先行、终端渐进”的策略,先在数据中心部署QRNG板卡集群供给量子密钥,终端侧随芯片普及逐步下沉。

表5 典型应用场景与密钥供给模式

应用场景	安全等级	时延要求	密钥供给模式
政务办公即时信	高	中	终端QRNG+云端实时
金融移动展业/交易	高	低(高频)	终端QRNG+云端熵材料(TEE内混合)
军工/国防通信	极高	中	终端QRNG+离线充注

5.2 产业实践印证与标准化生态

本方案各环节已有产业基础。三星 Galaxy 系列^[5-6]与中国联通抗量子安全手机^[7]已集成 QRNG 用于加密即时消息,国盾量子^[3]、IDQ^[4]芯片与板卡均规模出货,硅臻、循态^[13-14]与启科量子^[15]也提供相关产品。

量子安全融合已规模化落地。国内京沪干线量子保密通信骨干网已运行,中电信量子集团发布全球首个融合量子密钥分发(QKD)与PQC的分布式密码体系;国际上Signal提出PQXDH协议,Chrome默认启用X25519Kyber768混合密钥交换,覆盖数十亿终端。

在标准化方面,NIST于2024年8月发布FIPS 203/204/205,国内发布了YD/T 4301—2023等规范。中国联通研究院指出,密码敏捷性应成为系统架构的永久性设计原则,本方案可与QKD、PQC互补协同。

5.3 演进方向与展望

QRNG芯片的小型化、低功耗化与SoC集成是其规模落地的关键。清华大学等已实现无需后处理即具有良好随机性的QRNG^[16];随着《器件无关量子随机数发生器通用要求》^[8]《移动金融业务量子安全应用规范》^[17]等规范的完善,合规部署将有据可依。QRNG还可与QKD、PQC互补,共同构建移动即时通信量子安全体系。

6 结论

本文构建终端QRNG芯片、云端密钥服务与加密通道3层协同架构,将密钥随机性源头转向量子物理,并以量子熵种与DRBG结合的混合派生兼顾真随机质量与高频消息供给。分析表明,主密钥可预测性被约束至由量子熵主导的可忽略量级,状态泄露影响仅限于单个刷新周期,契合政务、金融等高保密场景。该方案以可接受的终端硬件代价换取安全性提升,具备低侵入集成条件;后续将沿芯片小型化、评测标准化及与QKD、PQC协同的方向进行深入研究。

参考文献:

[1] Heninger N, Durumeric Z, Wustrow E, et al. Mining your Ps and Qs:

detection of widespread weak keys in network devices [C]//Proceedings of the 21st USENIX conference on Security symposium. Bellevue:USENIX Association,2012:35.

- [2] 刘宇轩,白玉明,李俊林.量子随机性与量子随机数发生器的发展[J].物理,2025,54(10):679-689.
- [3] 国盾量子.量子随机数发生器[EB/OL].[2026-06-18].<https://www.quantum-info.com/product/coredevice/2017/0504/112.html>.
- [4] ID Quantique. Quantis QRNG PCIe Board[EB/OL].[2026-06-25].<https://www.idquantique.com/random-number-generation/>.
- [5] 刘于苇.全球首款量子加密手机要来了,什么是QRNG和高熵随机数?[EB/OL].[2026-07-28].<https://www.eet-china.com/news/202004241033.html>.
- [6] 科学网.“量子加密智能手机”真的来了[EB/OL].[2026-07-28].<https://news.sciencenet.cn/sbhtmlnews/2020/4/354932.shtm>.
- [7] 中国联通.中国联通发布抗量子安全手机[EB/OL].[2026-07-28].<https://www.chinaunicom.com.cn/>.
- [8] GB/T 45240-2025 器件无关量子随机数发生器通用要求[S].
- [9] Perrin T, Marlinpike M. The double ratchet algorithm [EB/OL].[2026-07-03].https://www.cs.wm.edu/~smherwig/readings/papers/16-misc-double_ratchet.pdf.
- [10] NIST SP 800-90A Rev. 1 Recommendation for random number generation using deterministic random bit generators[S].
- [11] 韩宇,费洋洋.面向典型密码学场景的量子随机数预测应用研究[J].计算机科学与应用,2026,16(2):395-404.
- [12] 安全内参.量子保密通信应用与技术探讨[EB/OL].[2026-07-03].<https://www.secrss.com/articles/46135>.
- [13] 硅臻量子.量子随机数发生器产品手册[EB/OL].[2026-07-10].<https://www.gz-ichip.com/src/pages/production-center/static/1-%E7%A1%85%E8%87%BB%E9%87%8F%E5%AD%90%E9%87%8F%E5%AD%90%E9%9A%8F%E6%9C%BA%E6%95%B0%E5%8F%91%E7%94%9F%E5%99%A8%E4%BA%A7%E5%93%81%E6%89%8B%E5%86%8C.pdf>.
- [14] 循态量子.QRNG系列产品白皮书[EB/OL].[2026-07-14].<https://www.xtquantech.com/>.
- [15] 启科量子.量子安全云QSCloud[EB/OL].[2026-07-19].<https://www.qudoor.com/>.
- [16] 清华大学.物理系李俊林研究组在基于电子体系的无后处理量子随机数发生器研究方面取得进展[EB/OL].[2026-07-07].<https://www.tsinghua.edu.cn/info/1175/105205.htm>.
- [17] T/CQJR 005-2023 移动金融业务量子安全应用规范[S].

作者简介:

孟小燕,工程师,主要从事移动终端信息安全研究工作;马伟,工程师,主要从事移动安全应用服务研究工作;陈颢,高级工程师,硕士,主要从事移动终端信息安全、云手机安全研究工作;侯玉华,高级工程师,主要研究方向为移动信息安全、终端操作系统;齐霄,高级工程师,主要从事移动安全应用与密码技术应用工作;李长连,高级工程师,主要从事网络安全技术研究工作。