基于双载波的物理层密钥高速生成方法研究

姜 禹[1,2,3], 王禹淳[1], 胡爱群[1,2,4]

[1] 东南大学网络空间安全学院 南京 中国 211189
[2] 网络通信与安全紫金山实验室 南京 中国 211189
[3] 江苏省计算机网络技术重点实验室 南京 中国 211189
[4] 东南大学移动通信国家重点实验室 南京 中国 210096

摘要 利用信道状态信息(Channel state information, CSI)的密钥生成技术已经发展多年, 常见的方法是利用 CSI 中全部子载波的信息作为一次密钥的数据源, 在此类方法中, 虽然合法通信双方所获 CSI 信息具有高度的相似性, 但密钥生成速率相对较低; 或者信道变化对所有子载波产生相似的影响, 整体趋势不变, 使得不同轮次生成的密钥重复率过高, 密钥更新率较低。为此, 本文首次提出了基于双载波商模型的密钥生成方法, 充分利用多载波的信息, 在保障合法通信双方数据互易性的基础上, 使得密钥生成速率明显提升; 针对序号邻近的子载波间存在相关性导致密钥重复段过多的问题, 提出了一种集合划分方法以及提取集合内具有代表性序列, 从而消除密钥重复段。最后, 搭建实验系统证明: 双载波商模型可以有效提升密钥生成速率; 基于集合划分的子载波相关性消除方法可以消除大量重复段, 提高密钥更新率; 基于双载波商模型和子载波相关性消除的密钥生成方法可以在保证双方不一致率小于给定阈值的情况下, 使得生成的密钥具有更良好的性能和安全性。实验结果表明, 本文提出的双载波商模型在静态及动态场景通信双方视距 1.5m 情况下, 将密钥生成速率提升 3～9 倍, 针对消除密钥重复段的子载波集合划分方法将密钥更新速率提升 2～3 倍, 由于子载波数只有 52, 可以忽略带来的额外计算开销。

关键词 物理层安全; 密钥分配; 信道状态信息; 双载波商模型; 密钥生成速率
中图法分类号 TN918.4 DOI 号 10.19363/J.cnki.cn10-1380/tn.2025.07.01

Research on High-Speed Generation Method of Physical Layer Key Based on Dual Carrier

JIANG Yu[1,2,3], WANG Yuchun[1], HU Aiqun[1,2,4]

[1] School of Cyberspace Security, Southeast University, Nanjing 211189, China
[2] Purple Mountain Laboratories, Nanjing 211189, China
[3] Key Laboratory of Computer Network Technology of Jiangsu Province, Nanjing 211189, China
[4] State Key Laboratory of Mobile Communication, Southeast University, Nanjing 210096, China

Abstract The key generation technology using channel state information (CSI) has been developed for many years, and the common method is to use all the subcarrier information in CSI as the data source of the primary key, in such methods, although the CSI information obtained by the legal communication parties has a high degree of similarity, but the key generation rate is relatively low. Or the channel change has a similar effect on all subcarriers, and the overall trend remains unchanged, making the key repetition rate generated in different rounds too high and the key update rate low. Therefore, this paper proposes a key generation method based on the dual-carrier quotient model for the first time, which makes full use of multi-carrier information and significantly improves the key generation rate on the basis of ensuring the data reciprocity of both parties to the legal communication. Aiming at the problem of excessive key duplication caused by the correlation between subcarriers adjacent to the serial number, a set division method and extracting representative sequences in the set are proposed to eliminate the key repeat segments. Finally, an experimental system is built to prove that the dual-carrier quotient model can effectively improve the key generation rate. The subcarrier correlation elimination method based on ensemble division can eliminate a large number of duplicate segments and improve the key update rate. The key generation method based on the dual-carrier quotient model and subcarrier correlation elimination can make the generated key have better performance and security under the condition that the inconsistency rate between the two parties is less than the legal threshold. The experimental results show that the dual-carrier quotient model proposed in this paper increases the key generation rate by 3～9 times under the condition of 1.5m between the communication parties in static and dynamic scenes, and increases the key update rate by 2～3 times for the subcarrier ensemble division method that elimi-

通讯作者: 姜禹, 博士, 东南大学副教授, Email: jiangyu@seu.edu.cn。
本课题得到国家重点研发计划项目(No. 2022YFB2902202)以及国家自然科学基金项目(No. 62171121 和 No. U22A2001)资助。
收稿日期: 2023-09-04; 修改日期: 2023-12-13; 定稿日期: 2024-06-13

nates the key repeat segment, because the number of subcarriers is only 52, which can ignore the additional computational overhead.

Key words physical layer security; key distribution; channel status information; dual carrier quotient model; key generation rate

1 引言

1.1 背景

无线通信系统的安全机制主要是从传统的有线通信安全机制中移植而来, 在链路层、网络层、传输层以及应用层设计相应的安全协议来保证信息的安全传输。传统的网络安全机制依赖于加密密钥来支持机密性和身份验证服务[1]。Wi-Fi 作为近年来最成功的无线通信技术之一, 目前已被广泛地应用于笔记本电脑、智能手机等各类设备。伴随物联网的发展许多具有 Wi-Fi 功能的智能终端也大量涌现, 这些设备间存在着密切频繁的数据交互, 通信的安全性受到了严峻的挑战。一方面, 设备身份可能未得到有效的认证管理, 另一方面, 无线媒介的广播性和网络的开放性使得信号暴露在外界环境中, 极易受到攻击者窃听和恶意攻击[2]。因此无线通信安全面临着两大关键安全问题, 一是接入时对用户身份的验证, 二是数据传输过程中的机密性[3]。前者是为了尽可能防止非法用户访问和读取数据, 后者目的是防范无线网络中的窃听攻击。

物理层密钥生成技术作为对现有对称密钥加密技术的补充, 利用无线电信道的固有随机性来提供动态共享密钥的一种很有前途的方法[4]。基于时变多输入多输出(Multiple-Input Multiple-Output, MIMO)信道的物理观测性质, 仔细考虑了利用时间和空间相关的信道样本来建立密钥[5]。因为无线信道的固有特性, 即互易性、空间去相关性和时间变化性, 所以由两个合法节点上的 CSI 生成的密钥是高度相关的和快速变化的。然而, 在信道相干时间较长的环境中, 信道的时变性质可能不够充分, 从而导致较低的密钥生成效率。

针对上述问题, Xiao S 等人[6], Qin D 等人[7]提出了部署更多通信设备的方法, 以获得更丰富 CSI 样本, 但增加中继、天线将会增加成本, 导致密钥生成过程不够轻量化。Zhang S 等人[8], Aldaghri N 等人[9]提出可以通过在单天线系统中引入人工随机性(比如: 随机信号)来改进密钥生成效率。Jin L 等人[10]指出, 上述方案限制了密钥生成速率的提升。因此, 如何以低成本提高密钥生成效率仍然是一个有待解决的问题。

1.2 目的

由于合法通信双方所获 CSI 信息具有高度的相似性, 大多数现有的基于 CSI 生成密钥的研究都依赖于此前提。但由于子载波数量有限, 信息利用率不高, 导致密钥生成速率较低; 或者信道变化对所有子载波产生相似的影响, 使得密钥重复率过高, 密钥更新率较低。

为了解决这类问题, 本文提出了基于双载波的物理层密钥高速生成方法。下面简单介绍本文的创作出发点:

虽然不同子载波的中心频率存在一定的差异, 但是它们产生于同一个物理设备, 都具有相同的物理性质和时钟频率。因此, 在经过一定的处理之后, 不同子载波的商可以一定程度上消除幅度噪声和相位偏移等误差, 合法通信双方能得到高信噪比且变化趋势高度相似的 CSI 商数。

现有方案大多需要引入额外的设备或装置, 通过提高随机性以满足密钥的生成速率。为了实现更加广泛的应用, 需要寻找一种低成本的解决方案同时满足控制设备成本以及提高密钥生成速率的要求。

本文希望能够充分利用多载波性质来建立双载波商数模型, 在保持设备原有发包率的基础上, 提升密钥生成效率; 提出一种消除子载波相关性的方法, 加快密钥更新速率, 保障通信安全。

1.3 主要贡献

本文的研究建立在无线信道多径效应[11]和双载波比率模型的理论基础上, 深入地探究了使用双载波之间的商数作为互易性数据源的可行性, 使其相较于传统 CSI 方法有更高的密钥生成速率, 更快的密钥更新率。同时, 以此为基础, 进行了基于双载波的高速密钥生成方法研究。

本文主要贡献如下:

针对密钥生成速率低的问题, 提出了双载波商模型, 利用单端多载波的性质, 将 CSI 中幅度和相位信息提取出来, 分别做商、做差, 得到的结果用于生成密钥。在保证合法节点间数据互易性的基础上, 尽可能提升多载波信息的利用率, 从而提升密钥生成速率。

针对双载波商模型中的邻近子载波相关性导致密钥重复段过多问题, 提出了一种子载波集合划分方法, 提取集合内具有代表性序列, 在保证提升密钥生成速率的基础上, 尽可能消除子载波间相关性, 从而提升密钥更新率。

利用两块 ESP32-S3 开发板和一台搭载 Ubuntu 20.04 系统的笔记本实现了一套基于双载波商模型和子载波相关性消除方法的密钥生成系统。实验证明, 本文所提出的方法可以在保证双方密钥一致率的基础上, 有效地提高 CSI 信息利用率、密钥生成速率和密钥更新速率。

2 相关工作

当设备处于信道相干时间较长的环境中, 信道的时变性质可能不够充分, 从而导致密钥更新速率慢、密钥生成速率慢等结果。针对上述问题, 本文对现有的相关研究进行了调研。

针对同一节点中的高相关连续样本产生的密钥随机性较差, da Cruz P I 等人[12]提出了一种离散差分预处理, 通过探索无线接收信号强度(Received Signal Strength Indicator, RSSI)中距离较远的样本之间的差异来增加加密密钥的随机性, 将基于累积分布函数(Cumulative distribution function, CDF)的量化算法扩展到基于有损 CDF(Lossy cumulative distribution function, LCDF)的量化算法, 该算法允许消除更有可能产生不一致位的样本, 增加了密钥的随机性。Hashem S 等人[13]以克服安全缺陷、提高 WEP 协议的安全性和保护水平为目的, 提出一种名为“提议的 RC4+S”的增强算法来改进有线等效协议(Wired Equivalent Protocol, WEP)的加密算法“标准 RC4”。RC4+S 算法将密钥的随机性增加了约 20%, 从而提高了修改后的 WEP 协议的输出密文随机性。Soni A 等人[14]提出利用移动窗口平均法(Moving window averaging, MWA)对 Alice 和 Bob 之间交换的信标的 RSSI 进行预处理, 在较低的信噪比范围内对性能有显著的改进。移动窗口的大小是根据 RSSI 模式的标准偏差来选择的。通过基于 Lloyd-Max 的量化器对预处理后的 RSSI 样本进行量化, 以最小化量化误差, 进一步降低了密钥不一致率。Zhang S 等人[8]提出 Alice 和 Bob 互相发送随机信号, 二者都不需进行信道估计, 直接将接收信号作为共享随机源去生成密钥, 由于增加了 Bob 向 Alice 发送随机信号的过程, 因此发双向随机信号方法比发单向随机信号方法的密钥生成速率更高。Li G 等人[15]提出了一种新的人工随机性(Artificial randomization, AR)辅助的慢衰落环境下快速密钥生成方法。将用户设计的随机性集成到信道探测中, 形成快速变化的组合信道, 实现信息论安全。当合法用户的信道条件比窃听者更好时, 通过引入 AR 可以提高密钥容量。仿真和实验结果表明, 通过精心设计探测数, AR 方法可以在慢衰落环境下有效地生成密钥。Zeng K 等人[16]提出并实现了一个基于多天线系统的共享密钥生成协议 MAKE。MAKE 允许发送方在发送每帧信号时动态改变发射天线。并在不同的场景中进行了实验。利用多天线分集和多级量化机制, MAKE 在性能上体现了优越性。针对无线体域网络(Wireless Body Area Networks, WBANs)容易受到主动窃听的问题, Xiao L 等人[17]提出了一种基于智能反射面(Intelligent reflection surface, IRS)辅助强化学习(Reinforcement learning, RL)的安全 WBAN 传输方案, 该方案使协调器能够优化传感器加密密钥和传输功率, 并针对主动窃听的 IRS 相移, 设计了一种 Dyna 架构, 通过模拟传输经验来提高学习效率, 采用安全探索来避免导致严重数据泄露的风险策略。

3 双载波商模型

3.1 研究基础

在无线通信领域, 信道状态信息即通信链路的信道属性。它描述了信号在每条传输路径上的衰弱因子, 即信道增益矩阵(信道衰落矩阵)中每个元素的值, 如信号散射、环境衰弱、距离衰减等信息。在采用正交频分复用(Orthogonal frequency-division multiplexing, OFDM)调制的 Wi-Fi 系统中, 每个子载波都是一个复数[18], 其表达式如公式(3.1)所示:

$$H(k)=a+bj \tag{3.1}$$

其中, $H(k)$ 为第 k 个子载波的信道状态信息。每一个子载波的 CSI 也可以表示为公式(3.2):

$$H(k)=\|H(k)\|^{j\angle H(k)} \tag{3.2}$$

其中, $\|H(k)\|$ 为第 k 个子载波的幅度, $\angle H(k)$ 为第 k 个子载波的相位。因此, CSI 是信道频率响应(Channel frequency response, CFR)不同子载波上的离散采样, CSI 反映的是无线信道随时间变化的情况, 它包含了无线信号从发射端到接收端所经历的直射、衍射、反射、折射、散射、偏振等变化信息。子载波的 CFR 可以表示为公式(3.3)[15]:

$$H(f,t)=\sum_{n}^{N}a_n(t)\cdot e^{j\cdot 2\pi f\cdot \tau_n(t)} \tag{3.3}$$

其中, $a_n(t)$ 为子载波幅度衰减因子, $\tau_n(t)$ 为子载波传播时延, f 为子载波频率, N 代表所有路径。

3.2 双载波商模型的提出

3.2.1 方法推导

目前常见的通信体制都采用 OFDM 调制方式, OFDM 利用不同频点的子载波发送信息, 可获取不同子载波的 CFR。CSI 实际上就是 CFR 在不同子载

波上的离散采样, 所以双载波商模型的核心思想是在时域上提取任意两个子载波的 CFR 离散采样序列, 将两个序列做商, 得到的 CFR 商数作为生成密钥的数据源。在采用 OFDM 调制方式时, 每个子载波上的幅度和相位信息可以视为对 CFR 频谱的一组采样数据, 而基于传统 CSI 的密钥生成方法是将同时测量的多个子载波的 CFR 作为共享随机源, 且互易性良好, 这说明两端对应的单一子载波的 CFR 具有相近的数值。在此基础之上, 本文从提升 CSI 信息利用率和密钥生成速率的角度出发, 提出了基于双载波商模型的密钥生成方法, 利用任意两个子载波的 CFR 在时间维度上的离散序列做商, 通过推导, 我们证明了其在理论上具备可行性, 具体推导过程如下:

首先, 接收端子载波 CSI 商数可以表现为公式(3.4):

$$CSI_r = \frac{CSI(k_1)}{CSI(k_2)} = \frac{H(f_1,t)}{H(f_2,t)} \tag{3.4}$$

其中, $CSI(k_1)$、$CSI(k_2)$分别表示序号为 k_1 和 k_2 的子载波的 CSI, f_1 和 f_2 为两条被选择的子载波的中心频率。上述公式可以进一步展开为公式(3.5):

$$\begin{aligned} CSI_r &= \frac{CSI(k_1)}{CSI(k_2)} \\ &= \frac{A(f_1,t)\cdot e^{-j\cdot 2\pi\Delta\theta(f_1,t)}\cdot H_{ideal}(f_1,t)}{A(f_2,t)\cdot e^{-j\cdot 2\pi\Delta\theta(f_2,t)}\cdot H_{ideal}(f_2,t)} \end{aligned} \tag{3.5}$$

其中, H_{ideal} 为理想状态的 CSI。由于一根天线共享相同的射频链和时钟, 因此同一接收器中 CSI 的振幅噪声和相位偏移可以近似看作相同, 即信道脉冲噪声 $A(f_1,t)$ 和 $A(f_2,t)$、CSI 相位偏移 $e(f_1,t)$ 和 $e(f_2,t)$ 近似相等, $H(f_1,t)$ 和 $H(f_2,t)$ 的商实际上就是消除了这些偏差。因此, $\frac{CSI(k_1)}{CSI(k_2)}$ 可以改写为公式(3.6):

$$\begin{aligned} CSI_r &= \frac{CSI(k_1)}{CSI(k_2)} \\ &= \frac{A(f_1,t)}{A(f_2,t)}\cdot e^{-j2\pi(\Delta\theta(f_1,t)-\Delta\theta(f_2,t))}\cdot\frac{H_{ideal}(f_1,t)}{H_{ideal}(f_2,t)} \\ &\approx \frac{H_{ideal}(f_1,t)}{H_{ideal}(f_2,t)} \end{aligned} \tag{3.6}$$

其次, 通信双方的信号在多径传输过程中, 假设发送信号为多载波信号, 如公式(3.7)所示:

$$x(t) = \sum_{k=0}^{K-1} X(k)e^{j\omega_k t} \tag{3.7}$$

其中, t 为子载波传播时延, ω_k 为载波频率。该信号经过无线信道的多径传输后, 接收信号如公式(3.8)所示:

$$y(t) = \sum_{m=0}^{M-1} a_m x(t-\tau_m) \tag{3.8}$$

其中 a_m、τ_m 分别为第 m 条路径对应的衰减和延时, 将(3.7)代入(3.8)中, 有公式(3.9):

$$\begin{aligned} y(t) &= \sum_{m=0}^{M-1} a_m \sum_{K=0}^{k-1} X(k)e^{j\omega_k(t-\tau_m)} \\ &= \sum_{K=0}^{k-1} X(k) \sum_{m=0}^{M-1} a_m e^{j\omega_k(t-\tau_m)} \\ &= \sum_{K=0}^{k-1} X(k)e^{j\omega_k t} \sum_{m=0}^{M-1} a_m e^{-j\omega_k\tau_m} \end{aligned} \tag{3.9}$$

令

$$\epsilon(\omega_k) = \sum_{m=0}^{M-1} a_m e^{-j\omega_k\tau_m} \tag{3.10}$$

$$y(t) = \sum_{K=0}^{k-1} \epsilon(\omega_k) X(k)e^{j\omega_k t} \tag{3.11}$$

由上式可见, 产生了衰落 $\epsilon(\omega_k)$, 它不仅与多径的衰减和延时有关, 还与载波频率 ω_k 有关。由(3.11)式可见, 我们在基带观测到的频谱实际上是公式(3.12):

$$\hat{Y}(t) = \epsilon(\omega_k) X(k) \tag{3.12}$$

即如公式(3.13)所示:

$$CSI(k) = \epsilon(\omega_k) = \sum_{m=1}^{M-1} a_m e^{-j\omega_k\tau_m} \tag{3.13}$$

将其写成矩阵形式如(3.14)所示:

$$\begin{aligned} &\begin{bmatrix} CSI(0) \\ CSI(1) \\ \vdots \\ CSI(K-1) \end{bmatrix} \\ &= \begin{bmatrix} e^{-j\omega_0\tau_0} & \cdots & e^{-j\omega_0\tau_{M-1}} \\ \vdots & \ddots & \vdots \\ e^{-j\omega_{K-1}\tau_0} & \cdots & e^{-j\omega_{K-1}\tau_{M-1}} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{bmatrix} \end{aligned} \tag{3.14}$$

上式中, 等号左边为测量值, 如 Wi-Fi 的 52 个子载波对应的 CSI。我们将时延 τ_m 以采样间隔进行量化, 这个采样值是对载波高频信号的采样, 也是对多径的最大时间分辨, 即 $\tau_m = m(m+1)\cdot T_s$。对于频率为 f_0 的载频信号, $\omega_k = \omega_0 + k\cdot\Delta\omega$, 其中 $\omega_k = 2\pi f_0 + k\cdot 2\pi\cdot\Delta f$。

把上述时间参数代入(3.14)式, 则可解出多径衰减系数为公式(3.15)所示,

$$A_M = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{bmatrix} = \boldsymbol{E}^{-1}\cdot \begin{bmatrix} CSI(0) \\ CSI(1) \\ \vdots \\ CSI(K-1) \end{bmatrix} \tag{3.15}$$

其中,

$$\boldsymbol{E}=\begin{bmatrix} e^{-j\omega_0\tau_0} & e^{-j\omega_0\tau_1} & \cdots & e^{-j\omega_0\tau_{M-1}} \\ e^{-j\omega_1\tau_0} & e^{-j\omega_1\tau_1} & \cdots & e^{-j\omega_1\tau_{M-1}} \\ \vdots & \vdots & \cdots & \vdots \\ e^{-j\omega_{K-1}\tau_0} & e^{-j\omega_{K-1}\tau_1} & \cdots & e^{-j\omega_{K-1}\tau_{M-1}} \end{bmatrix} \tag{3.16}$$

由公式(3.16)可以看出, 衰减系数与载波频率和时延相关, 且具有多径独立性。因此, 双方在信道环境近似相同的情况下, 理论上可以获得具有互易性的信息。

所以, 结合信号传输过程的多径效应和接收端信号的双载波商数推导过程, 可以说明, 在双载波商模型中, 合法通信双方大致可以抵消振幅噪声和相位偏移, 使得双方得到的商数 CSI_r 高度相似, 为生成一致性密钥提供了可用的共享随机源, 并且, 子载波个数只有 52 个, 计算量非常小, 所以带来的额外计算开销可以忽略不计。

由于在 CSI 矩阵中, 每个子载波均以复数形式存在。根据复数相关运算法则, 可以求得该复数的模为 $\sqrt{a^2+b^2}$, 辐角值为 $arctan\dfrac{b}{a}$。因此可将该值作为 $H(f_k)=H(f_k)e^{j\sin(\angle H(f_k))}$ 的振幅和相位。因此, CSI_r 可以写为幅度比值 A_r 和相位差值 P_r 的形式, 如公式(3.17)、公式(3.18)所示:

$$A_r=\frac{Amplitude(k_1)}{Amplitude(k_2)} \tag{3.17}$$

$$P_r=Phase(k_1)-Phase(k_2) \tag{3.18}$$

3.2.2 计算步骤

双载波商模型的具体计算步骤如图 1 所示。

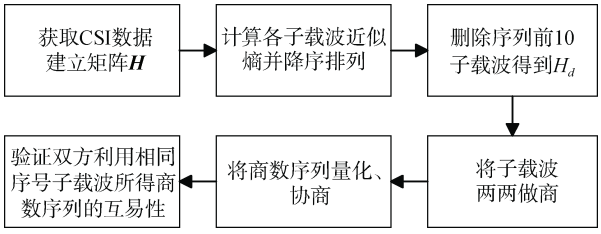


图 1 计算步骤示意图

Figure 1 Schematic diagram of calculation steps

我们使用矩阵表示探测到的CSI数据, 其中有 m 帧 CSI 数据, 每帧有 52 个有效子载波, 子载波序号范围为[–26,–1]∪[1,26], 可以表示为 m×52 的矩阵, 如公式(3.19)所示:

$$H=\begin{bmatrix} h_{(1,-26)} & \cdots & h_{(1,26)} \\ \vdots & \ddots & \vdots \\ h_{(m,-26)} & \cdots & h_{(m,26)} \end{bmatrix} \tag{3.19}$$

其中, $h_{(m,n)}$ 表示第 m 帧数据中第 n 号子载波的 CSI。为了保证双载波商模型始终高效运行, 需要根据每个子载波的近似熵进行筛选, 删除互易性差的子载波。近似熵可以定量描述时间序列的复杂程度, 它是一种用于量化时间序列波动的规律性和不可预测性的非线性动力学参数, 用一个非负数来表示一个时间序列的复杂性, 反映了时间序列中新信息发生的可能性, 越复杂的时间序列, 对应的近似熵越大。

设有长度为 N 的子载波采样序列 X, 其近似熵的计算步骤如下:

1) 将序列 X 的元素按时间顺序排序为具有 m ($m<N$)维数的向量, m 可以被称之为模式维数, 它是用来计算近似熵的时序列的窗口长度。在一般情况下, 选择 $m=2$ 或 $m=1$。只有当 $N>1000$ 时, 才会选择 $m>2$。因此我们选择 $m=2$。如公式(3.20)所示:

$$X_i=[x(i),x(i+1),\ldots,x(i+m-1)] \tag{3.20}$$

其中 $i=1,2,\ldots,N-m+1$。

2) 定义 $d\left[X_i,X_j\right]$ 为向量 X_i 与 X_j 的距离, 如公式(3.21)所示:

$$d\left[X_i,X_j\right]=\max|x(i+k)-x(j+k)| \tag{3.21}$$

其中, $k\in(0,m-1)$, $i,j\in[1,N-m+1]$。

3) 记 B_i 为 $d\left[X_i,X_j\right]\leqslant r$ 的个数, 其中 $i,j\in[1,N-m+1]$, r 为相似容限, 当 r 的值较大时, 会丢失较多的信息, 而当 r 的值较小时, 也不能理想的估计出系统的统计特性, 因此 r 在 0.1 和 0.5 之间时, 近似熵有比较合理的统计特性[19]。计算 B_i 与 $N-m+1$ 比值, 如公式(3.22)所示:

$$B_i^m(r)=\frac{B_i}{N-m+1} \tag{3.22}$$

4) 对 $B_i^m(r)$ 进行取对数运算, 再求其对所有 i 的平均值, 记作 $B^m(r)$, 如公式(3.23):

$$B^m(r)=\frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\ln B_i^m(r) \tag{3.23}$$

5) 令 $m=m+1$, 并重复步骤 1)到 4), 即可得到 $B^{m+1}(r)$;

6) 得到此序列的近似熵 $ApEn(m,r)$, 如公式(3.24)所示:

$$ApEn(m,r)=\lim\left[B^m(r)-B^{m+1}(r)\right] \tag{3.24}$$

其中, $N\to\infty$。对于本文的离散采样序列, N 不会趋于无穷, 因此近似熵可以表示为公式(3.25):

$$ApEn(m,r,N) = B^m(r) - B^{m+1}(r) \quad (3.25)$$

通过大量实验数据验证, 近似熵过高的子载波, 其 CSI 商数的互易性大概率很差, 无法生成一致性密钥。因此, 需要将各个子载波的近似熵降序排列, 将序列前 10 的子载波删除, 称为不可用子载波。

经过筛选后的矩阵变为 $m \times n$ 的矩阵 H_d (n <52), 如公式(3.26)所示:

$$H_d = \begin{bmatrix} h_{(1,x_1)} & \cdots & h_{(1,x_n)} \\ \vdots & \ddots & \vdots \\ h_{(m,x_1)} & \cdots & h_{(m,x_n)} \end{bmatrix} \quad (3.26)$$

其中, $x_1 \sim x_n$ 为剩余的子载波序号, 按序号的升序排列。用 $\mathrm{H}_{(n_i)}$ 表示第 n_i 号子载波的离散采样序列, 然后将任意两个子载波进行商运算, 得到商数 $R_{n_1 n_2}$, n_1、n_2 代表子载波序号, 其表达式如公式(3.27)所示:

$$CSI_{n_1 n_2} = \frac{\mathrm{H}_{(n_1)}}{\mathrm{H}_{(n_2)}} \quad (3.27)$$

其中 $\mathrm{H}_{(n_i)}$ 如公式(3.28)所示:

$$\mathrm{H}_{(n_i)} = \begin{bmatrix} h_{(1,n_i)} \\ \vdots \\ h_{(m,n_i)} \end{bmatrix} \quad (3.28)$$

合法节点用相同的子载波序号 n_1、 n_2 进行商运算, 各自得到 $CSI_{n_1 n_2}$。

然后通过分块格雷量化法分别将两端的 $CSI_{n_1 n_2}$ 进行量化, 协商阶段使用 BCH(Bose-Chaudhuri-Hocquenghem codes, BCH)[20]纠错算法, 利用不同的本原多项式实现多位纠错。我们设想不一致率不能超过 20%, 以免因为频繁协商增加信息泄漏风险, 因而将纠错位数上限设置为 22 位。最终根据能否协商成功, 验证其互易性。

3.3 双载波商模型的结果与验证

为了对比验证双载波商模型的理论效果和现实表现, 本节设计了对比分析实验。实验环境为空旷的办公室, 实验设备为 2 个 ESP32-S3 开发板, 分别作为 Alice 和 Bob。Alice 工作在 Access Point 模式, 提供无线接入服务, 允许其他无线设备接入, 提供数据访问, Bob 工作在 Station 模式, 可以连接到 AP。设备工作在 2.4GHz 频段, 带宽为 20MHz。一台搭载 Ubuntu 20.04 系统的笔记本用于实时监视捕获的 CSI 数据并分析。根据 802.11n 协议相关说明[21], 最终含有效数据的子载波为 52 个, 序号为–26 至–1 和 1 至 26。Alice 和 Bob 间的距离约 1.5m, 如图 2 所示。

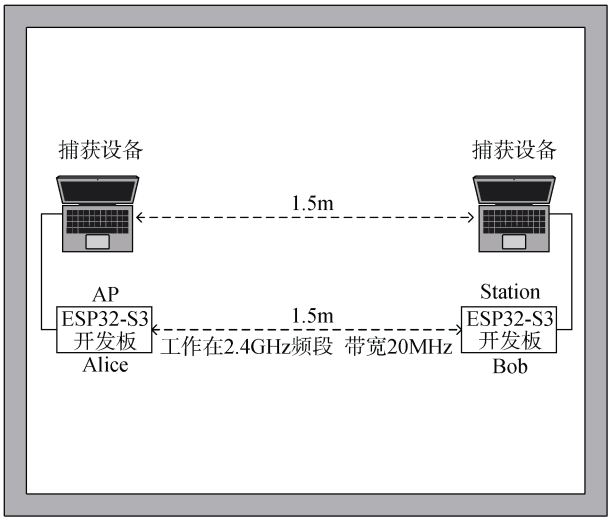


图 2 实验环境示意图

Figure 2 Schematic diagram of experimental environment

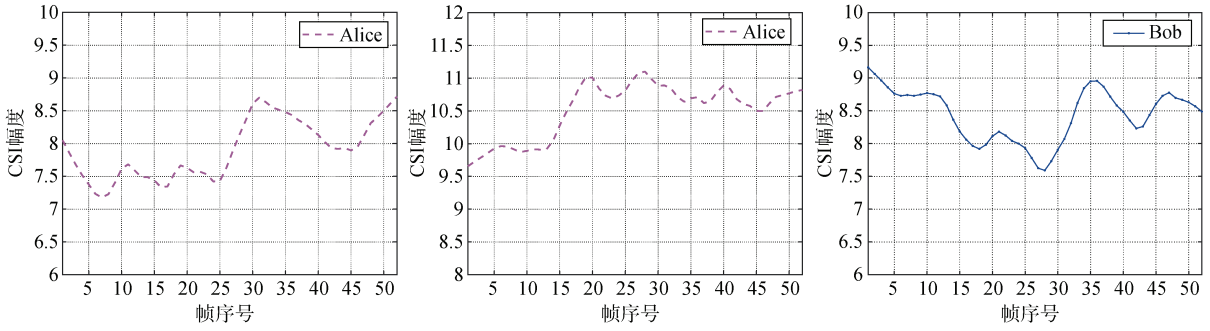
在 Alice 和 Bob 捕获的 CSI 数据中, 分别提取了连续 52 帧的 CSI 幅度和相位, 两端的帧序号一一对应。如图 3、图 4 所示为 Alice 的–26 号、12 号子载波的幅度图和 Bob 端的–26 号、12 号子载波的幅度图, Alice 端–17 号、–3 号子载波的相位图和 Bob 端的–17 号、–3 号的相位图, 可以看出两端相同序号子载波幅度随时间变化的轨迹完全不重合, 几乎不存在互易性, 无法作为生成一致性密钥的数据源。所以我们用双载波商模型将数据进行处理, 将同一端不同序号(此处为–26 号和 12 号)的子载波幅度相除、相位做差, 结果如图所示, Alice 端和 Bob 端的 CSI 幅度比值和相位差值的变化趋势高度相似, 具备生成密钥所需的互易性条件。因此, 验证了双载波商模型的有效性和可用性, 为后续密钥生成奠定了基础。

4 子载波集合划分方法

4.1 密钥重复段问题

虽然 OFDM 信道中的子载波信号频率正交, 但相邻的子载波具有非常接近的频率, 导致频域产生相似的信道响应, 来自它们测量的 CSI 可能具有很强的相关性。因此与邻近的子载波做商生成的密钥可能有很大比例的重复段, 容易被攻击。在这种情况下, 生成的部分密钥是不可用的。如图 5 所示为–26 号和 12 号、–25 号和 12 号子载波的幅度比值示意图。

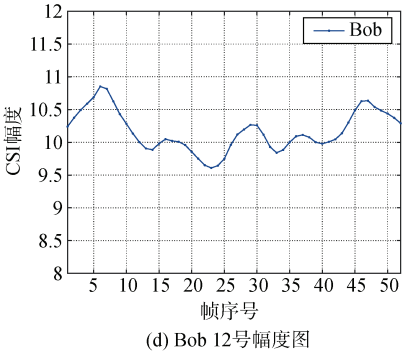
从图中可以看出两对子载波商数的整体变化趋势相似, 可以推测其量化结果存在一定数量的重复段(连续 5 位及以上重复)。经过验证, 在其量化结果的 0, 1 序列中, 确实存在大量重复段, 累计重复位数



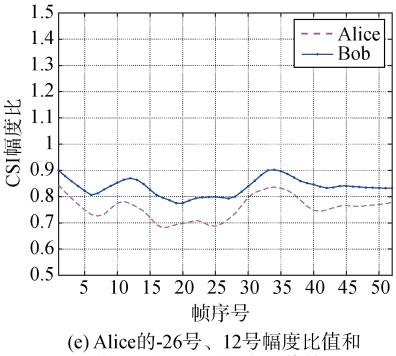
(a) Alice -26号幅度图
(a) Alice's -26 amplitude plot

(b) Alice 12号幅度图
(b) Alice's 12 amplitude plot

(c) Bob -26号幅度图
(c) Bob's -26 amplitude plot



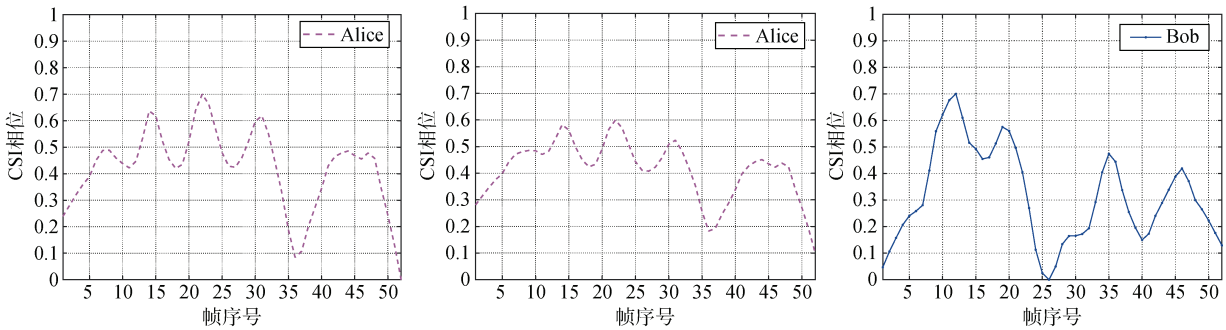
(d) Bob 12号幅度图
(d) Bob's 12 amplitude plot



(e) Alice的-26号、12号幅度比值和
Bob的-26号、12号幅度比值效果图
(e) Alice's ratio of -26 and 12 amplitudes and
Bob's -26, 12 amplitude ratio comparison chart

图 3　幅度比值效果图

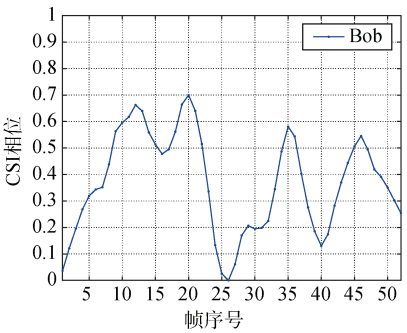
Figure 3　Amplitude ratio effect diagram



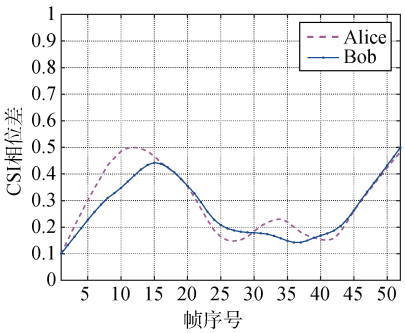
(a) Alice -17号相位图
(a) Alice's -17 phase map

(b) Alice -3号相位图
(b) Alice's -3 phase map

(c) Bob -17号相位图
(c) Bob's -17 phase map



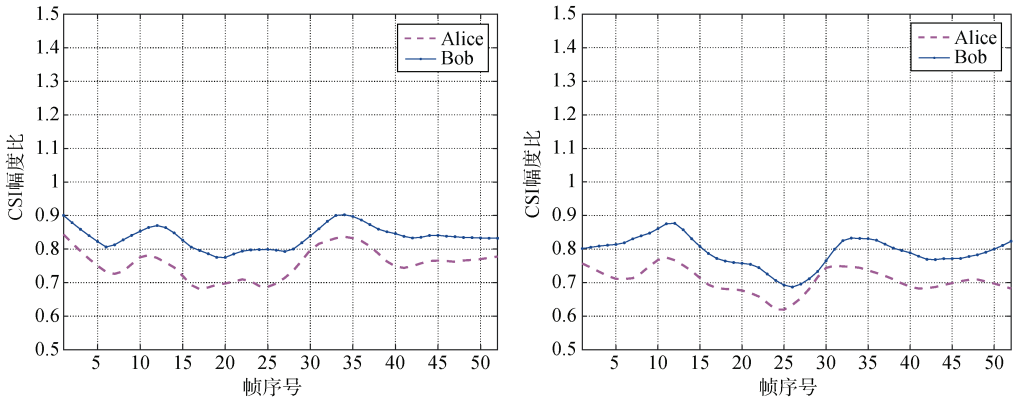
(d) Bob -3号相位图
(d) Bob's -3 phase map



(e) Alice的-17号、-3号相位差值和
Bob的-17号、-3号相位差值效果图
(e) Alice's -17 and -3 phase difference values
Bob's contrast diagram of phase difference values -17 and -3

图 4　相位差值效果图

Figure 4　Phase difference effect diagram



(a) -26号和12号子载波幅度比值图
(a) Subcarrier amplitude ratio plot of No. - 26 and No. 12

(b) -25号和12号子载波幅度比值图
(b) Subcarrier amplitude ratio plot of No. -25 and No. 12

图 5　两个相邻子载波与同一子载波比值效果图

Figure 5　Effect diagram of the ratio between two adjacent subcarriers and the same subcarrier

占总密钥长度的 38.46%。因此, 为了在一定程度上消除密钥中的重复段, 需要进一步处理子载波集合。

4.2　子载波集合划分

由于存在多个相似度较高的子载波, 只能取其中一个来生成密钥。因此, 通过对子载波进行集合划分, 将相似度大于设定阈值的子载波划分到同一个集合中, 使得在同一集合中的子载波, 每对都具有较强的相关性。本文用相关系数衡量子载波间相关性, 相关系数是用以反映变量之间相关关系密切程度的统计指标[22], 其公式如(4.1)所示:

$$p=\frac{\operatorname{cov}(X,Y)}{\sigma_x\sigma_y} \tag{4.1}$$

其中, X, Y 分别表示不同子载波序列, $\operatorname{cov}(X,Y)$ 代表它们之间的协方差, σ_x、σ_y 分别表示它们的标准差, p 称为相关系数。

首先, 大量实验数据显示, 相关系数介于 0～0.85 之间的子载波生成的密钥重复率较低。因此, 计算 52 个子载波中每对子载波的相关系数, 结果如图 6 所示。其中, 横纵坐标都为子载波序号, 红色代表相关系数大于 0.9, 黄色代表相关系数介于 0.85～0.9 之间, 绿色代表相关系数介于 0～0.85 之间, 即为可用的子载波对。根据结果观测, 相关系数介于 0～0.85 之间的子载波不会产生大量重复段。为了方便进行集合划分, 把图 6 转化为了 0, 1 矩阵, 如图 7 所示, 将红色和黄色部分变为 1, 其余为 0。

由此可以得到一个 52×52 的代表子载波间相关性的 0,1 矩阵 C, 如公式(4.2)所示:

$$\mathrm{C}=\begin{bmatrix} corr_{(-26,-26)} & corr_{(-26,-25)} & \cdots & corr_{(-26,26)} \\ \vdots & \vdots & \ddots & \vdots \\ corr_{(26,-26)} & corr_{(26,-25)} & \cdots & corr_{(26,26)} \end{bmatrix} \tag{4.2}$$

其中 $\mathrm{C}_{(n)}=\begin{bmatrix} corr_{(n,-26)} & corr_{(n,-25)}\cdots corr_{(n,26)} \end{bmatrix}$, 表示第 n 号子载波与其他子载波的相关性。将其称为 n 号子载波的相关集。

然后通过对每对子载波的相关集的异或运算, 得到具体的集合划分方案。运算规则如下:

1) 若两个相关集 $C_{(n_1)}$,$C_{(n_2)}$ 的按位异或的结果 $\mathrm{C}_{(n_1)}\oplus\mathrm{C}_{(n_2)}$ 为全 0 集, 则 n_1、n_2 两个子载波可以划分到同一集合;

2) 若 $\mathrm{C}_{(n_1)}\oplus\mathrm{C}_{(n_2)}$ 不为全 0 集, 则 n_1、n_2 两个子载波不属于同一个集合, 互不相关;

3) 将所有相关集进行两两异或运算, 划分子载波集合, 此时每个集合至多有两个子载波;

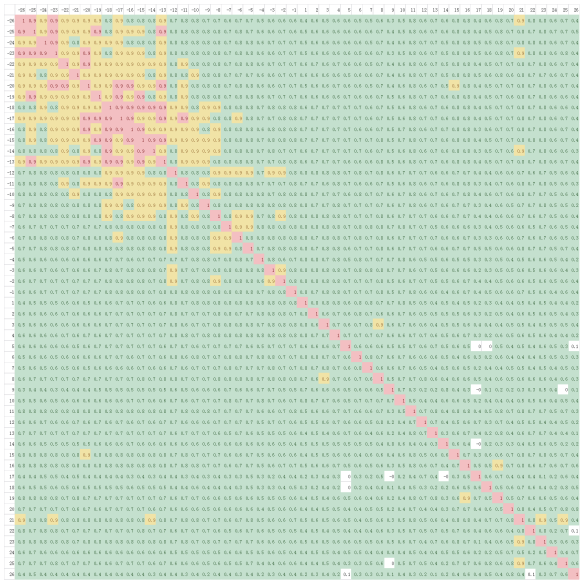


图 6　子载波相关系数矩阵示意图

Figure 6　Schematic diagram of subcarrier correlation coefficient matrix

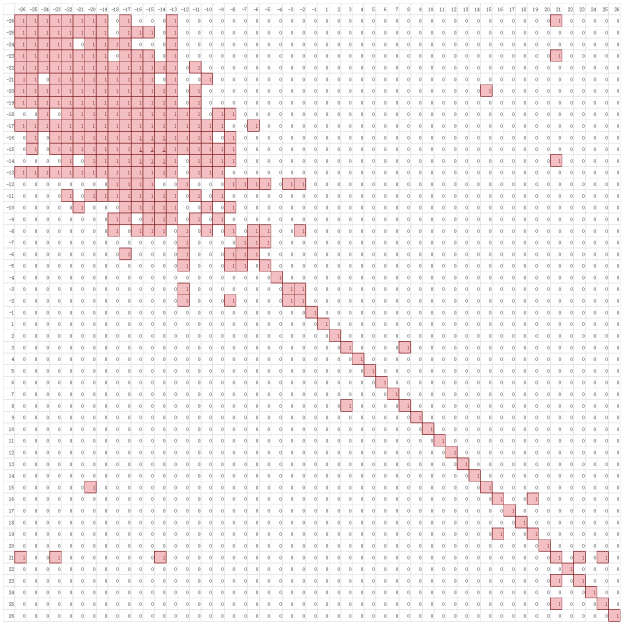


图 7　相关系数矩阵转换为 0, 1 矩阵示意图
Figure 7　Schematic diagram of converting the correlation coefficient matrix into a 0,1 matrix

4) 对存在相同子载波序号的集合进行合并, 最终确定原 CSI 数据被划分成了多少个集合, 以及各集合内的子载波序号。

4.3　集合内主成分提取

设想在每个集合中, 基于主成分分析法[23], 得到可以代表整个集合数据特征的序列。假设集合中有 n 个子载波, 表示为 $\{H_1, H_2, \ldots, H_n\}$。用主成分分析法可以得到前 p 个主成分以及其对应的特征值 α 和特征向量 h。则每个主成分中对应指标的系数为公式(4.3)所示:

$$\sigma = \frac{h}{\sqrt{\alpha}} = \frac{\text{每个主成分中各子载波对应元素}}{\sqrt{\text{每个主成分对应的特征值}}} \tag{4.3}$$

$$\sigma_i = \frac{h^i}{\sqrt{\alpha_i}} = \begin{bmatrix} \frac{h_1^i}{\sqrt{\alpha_i}} \\ \frac{h_2^i}{\sqrt{\alpha_i}} \\ \frac{h_3^i}{\sqrt{\alpha_i}} \\ \vdots \\ \vdots \\ \frac{h_i^i}{\sqrt{\alpha_i}} \end{bmatrix} \tag{4.4}$$

$$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_p) = \begin{bmatrix} \frac{h_1^i}{\sqrt{\alpha_1}} & \cdots & \frac{h_1^p}{\sqrt{\alpha_i}} \\ \vdots & \ddots & \vdots \\ \frac{h_n^i}{\sqrt{\alpha_i}} & \cdots & \frac{h_n^p}{\sqrt{\alpha_i}} \end{bmatrix} \tag{4.5}$$

每个主成分 F_i 都可以用如下的线性组合表示为公式(4.6):

$$F_i\varphi = \sigma_{i1}H_1 + \sigma_{i2}H_2 + \cdots + \sigma_{in}H_n \tag{4.6}$$

然后利用主成分的方差贡献率确定综合得分模型系数, 记前 p 个主成分特征值的方差贡献率为 φ, 综合得分模型系数为 γ (γ 对应每个子载波的综合系数), 则有公式(4.7):

$$\gamma_i = \frac{\sum_{j=1}^{n} \phi_i \sigma_{ij}}{\sum_{k=1}^{p} \Phi_k} \tag{4.7}$$

则得到综合得分模型为公式(4.8)所示:

$$Y = \gamma_1 H_1 + \gamma_2 H_1 + \cdots + \gamma_n H_1 \tag{4.8}$$

以集合{–16, –15, –14}为例对此方法进行验证, 主成分分析结果如表 1 所示:

表 1　集合{-16, -15, -14}分析结果表
Table 1　Analysis results table of set {-16, -15, -14}

名称	主成分 1		
特征根	2.721	综合得分系数	权重
方差解释率	90.69%		
–16 号	0.5797	0.5797	33.47%
–15 号	0.5730	0.5730	33.08%
–14 号	0.5794	0.5794	33.45%

其中, 方差解释率表示提取的主成分对原子载波组合特征的解释能力, 方差解释率越大, 解释能力越强, 提取的主成分或因子越有效。它可以由多元回归模型的结果进行计算, 其计算方式为: 方差解释率=1–残差平方和/样本总体变量方差。由于集合内子载波之间相关性都很强, 所以一般情况下, 只会得到一个可以代表整个集合数据特征的主成分, 如集合{–16, –15, –14}, 通过上述分析和计算, 最终得到一个主成分。将其与 7 号子载波做商, 两端的结果如图 8 所示:

4.4　密钥更新率提升

我们将密钥更新率定义为单位时间内密钥更新次数, 单位: 次/秒。如公式(4.9)所示:

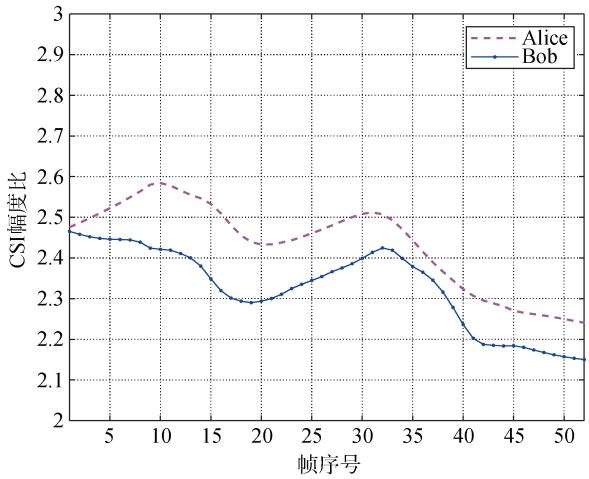


图 8　集合主成分与 7 号子载波的比值效果图

Figure 8　Ratio effect diagram of aggregate principal components and subcarrier No. 7

$$密钥更新率=\frac{密钥更新次数}{消耗时间} \tag{4.9}$$

随机选取一组 CSI 数据, 考虑到生成密钥以及通信过程中耗费的时间成本, 每隔 5 组提取 1 组密钥, 每组 104 位, 计算与上一次提取密钥相比的更新情况。当前密钥与上一次密钥对比, 重复段(连续 5 位及以上重复)小于两段的视为密钥更新一次。

首先, 利用传统 CSI 方法生成了 52 组密钥, 每组密钥按 CSI 的帧序号升序进行排列。每隔 5 组计算一次密钥重复段, 结果显示密钥只更新了 4 次。

然后, 利用双载波商模型, 共生成了 85 组密钥, 每组密钥按照 CSI 商数中分母的子载波序号升序排列, 分母序号相同的情况下, 按分子的子载波序号升序排列。理想情况下密钥更新次数应该为 17 次, 但实际情况是只更新了 8 次。

经过集合划分的 CSI 数据中, 共生成 71 组密钥, 同样, 每隔 5 组计算一次密钥重复段, 理想更新次数为 14 次, 实际上更新了 11 次。

由此可见, 原始 CSI 数据经过子载波集合划分之后, 虽然牺牲掉了一部分密钥生成速率, 但是提升了密钥更新率, 提高了密钥可用性和安全性。对于不同场景的密钥更新率提升情况, 将在后续章节中详细展示。

5　密钥生成方案设计

结合上述提出的双载波商模型和子载波集合划分方法, 在本章中, 提出了完整的密钥生成方案, 包括信道探测、预处理模块、量化模块和协商模块, 整体流程如图 9 所示。

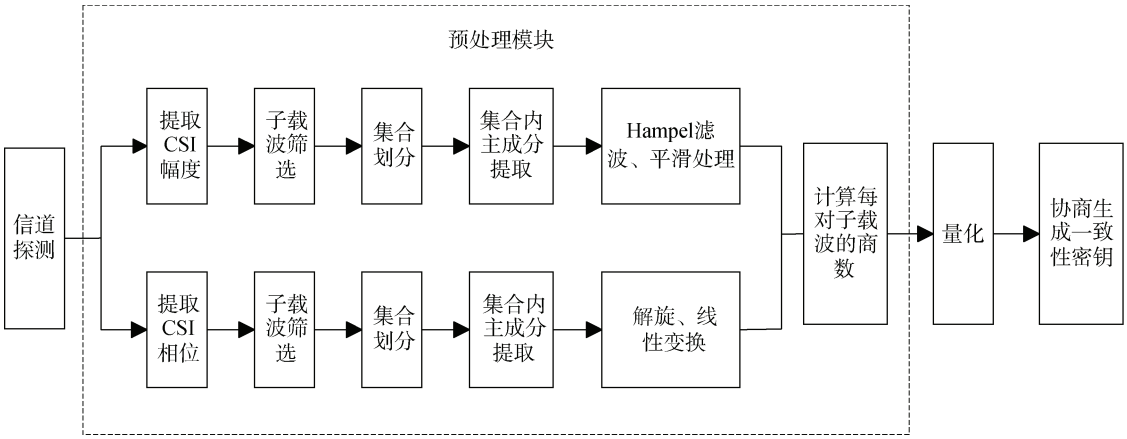


图 9　密钥生成方案流程图

Figure 9　Key generation scheme flow chart

5.1　CSI 信息处理模块

通过对原始数据结构分析可知, CSI 信息主要包括两部分, 幅度信息和相位信息[24]。其中, CSI 的原始幅度信息因受环境及物理干扰的影响, 在时域范围内含有较多的噪声, 会对通信双发互易性造成一定影响, 不能直接用于密钥生成。采集的 CSI 原始相位信息因时钟同步误差和载波频率误差的影响, 在频域范围内呈现随机分布。因此, 幅度和相位的处理方式存在一些差异。

5.1.1　幅度信息处理

由于现实情况的不可预测, 所以 CSI 在产生和提取过程中会无法避免地产生一些异常的离群点。离群点的数据是无效数据, 会对后续数据处理和密钥一致性造成一定的偏差。因此, 可以使用 Hampel 滤波器来检测和消除这些离群点。Hampel 滤波器是一种常用的高效消除离群点的方法, 其基本原理是

利用绝对中位差对估值区间外的离群值进行检测和过滤。

假设第 k 条子载波表现为公式(5.1):

$$H(k)=\{H_1,\ H_2,\ H_3,\ H_4,\ \ldots,\ H_n\} \tag{5.1}$$

其中, n 表示样本数量, H_n 表示第 k 条子载波第 n 个样本点的 CSI 幅度或者相位。Hampel 滤波器围绕 $H(k)$ 每个元素生成观测窗口, 假设半个窗口宽度为 x, x 默认值为 5, 整个窗口的宽度为 $2x+1$ (包含中心元素), 计算该窗口中所有元素的中值, 并利用中位数的绝对值估计各样本对中值的标准差。如果某个样本与中值相差超过三个标准差, 则用中值替换该样本。同时对数据进行平滑处理, 以启发方式确定的固定窗长度返回向量元素的移动平均值。窗口向下滑动向量的长度, 计算每个窗口中的元素的平均值。最后, 对实验中的一组 CSI 幅度数据应用了该方法, 其结果如图 10 所示, 处理后的幅度总体变化趋势更加明晰, 消除了一定噪声, 保留了幅度的原有特征, 有利于提高合法节点间的互易性。

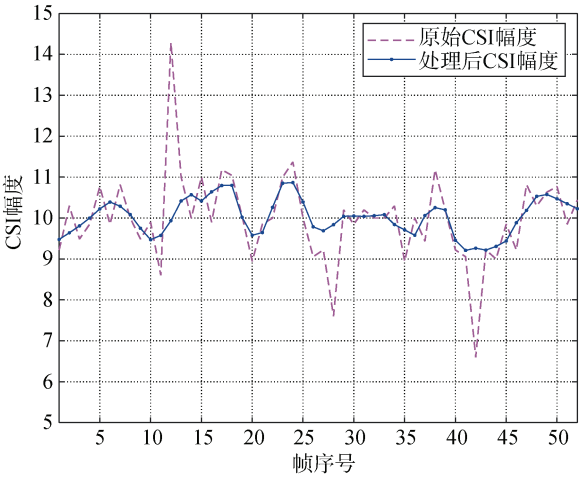


图 10 经离群点检测和平滑处理后的幅度效果对比图

Figure 10 Comparison of amplitude effects after outlier detection and smoothing processing

5.1.2 相位信息处理

与 CSI 载波幅度不同, CSI 载波原始相位周期范围为 $[-\pi,\ \pi]$, 临界点 $-\pi$ 和 π 处会发生反相, 使得相邻载波间的相位跳变大于 π, 使得双方的相位信息不一致位数过高, 无法进行纠错, 生成密钥。所以需要解卷绕全部载波的相位, 将在 $-\pi$ 到 π 之间跳变的相位信号, 变为不跳变的形态。经过解卷绕之后, 双方 CSI 相位信息变化趋势高度相似, 具备生成一致性密钥的条件。

MIMO-OFDM 系统中收发端之间可能存在的同步误差, 会导致采集到的 CSI 数据含有误差成分, 包括时钟同步误差和载波频率误差, 这会引起 CSI 振幅和相位的变化[25]。其中, 时钟同步误差对 CSI 振幅影响较小, 一般可以忽略不计, 而载波频率误差对 CSI 相位影响较大, 因此需要对因载波频率误差引起的相位偏移进行处理, 以得到不含相位偏移成分的真实相位信息。

这些误差因素会导致 CSI 相位偏移, 一部分来自与子载波序号和时钟同步误差成正比的相位误差, 另一部分来自与子载波频偏成正比的未知常数相位误差。若要获取可靠的有规律的相位信息, 需要通过一定的数学方法对相位信息进行预处理。Qian K 等人[26]通过对原始相位进行线性变换来移除随机相位偏移量, 得到有效相位, 并在人体检测领域中得到应用, 取得了很好的定位效果。

所以, 本文为了改善原始 CSI 相位在合法节点之间的互易性, 尝试先将原始相位解卷绕, 再利用数学上的线性变换预处理, 最终得到的有效相位作为生成密钥的数据源。具体步骤如下, 假设测量得到的第 i 个子载波相位为 $\hat{\phi}_i$, 如公式(5.2)所示:

$$\hat{\phi}_i=\phi_i+\beta-2\pi\frac{k_i}{N}\delta+Z \tag{5.2}$$

其中, $\hat{\phi}_i$ 表示测量相位, ϕ_i 表示真实相位, δ 表示接收端相对于发射端的时钟同步误差, 其对应产生的相位偏移表示为 $2\pi\frac{k_i}{N}\delta$, k_i 表示第 i 个子载波的编号值($-26\leqslant k_i\leqslant 26$), N 表示傅里叶变换 FFT 的大小(基于 IEEE 802.11n 标准中的 FFT 大小值为 64)。β 表示未知常数的相位偏移, Z 表示测量噪声, 也就是随机相位误差。

从 CSI 的测量相位偏移误差可以看出, 该相位偏移误差 $\beta-2\pi\frac{k_i}{N}\delta$ 是一个子载波编号的线性函数 ak_i+b, 其误差主要来自时钟同步误差 δ 和未知常数的相位偏移 β。因此, 为消除 δ 和 β 的影响, 可以考虑引入整个频段所有子载波的相位来对测量相位进行变换。

变量 a 和 b 的定义, 如公式(5.3)和(5.4)所示:

$$a=\frac{\hat{\phi}_n-\hat{\phi}_1}{k_n-k_1}=\frac{\phi_n-\phi_1}{k_n-k_1}-\frac{2\pi}{N}\delta \tag{5.3}$$

$$b=\frac{1}{n}\sum_{j=1}^{n}\hat{\phi}_j=\frac{1}{n}\sum_{j=1}^{n}\phi_j-\frac{2\pi\delta}{nN}\sum_{j=1}^{n}k_j+\beta \tag{5.4}$$

此处的 a 表示相位相减的值, 也就是计算累积相位差的结果, b 表示 n 个测量相位的平均值。在子载波的编码取值范围–26 到 26 内, 子载波的中心频

率是对称的, 因此整个频段内所有子载波相加之和为 0, 即有 $\sum_{j=1}^{n}k_j=0$, 进一步可以得到 b 的化简形式为 $b=\frac{1}{n}\sum_{j=1}^{n}\varphi_j+\beta$。从测量相位 $\hat{\phi}_i$ 中减去估计线性误差项 ak_i+b, 即可消除来自时钟同步误差 δ 和未知常数 β 的随机相位偏移, 进一步得到最终较为可靠的真实相位, 如公式(5.5)所示:

$$\tilde{\phi}_i=\hat{\phi}_i-(ak_i+b)=\phi_i-\frac{\phi_n-\phi_1}{k_n-k_1}k_i-\phi_n\frac{1}{n}\sum_{j=1}^{n}\phi_j \quad (5.5)$$

由此可见, 原始 CSI 相位信息经线性变换处理后已经消除了误差项 δ 和 β 的影响, 此时测量相位只包含真实相位信息的线性组合形式。

如图 11 所示为原始 CSI 的相位信息经解卷绕和线性变换后的对比图, 可以看出经解旋、线性变换后得到平稳的相位信息, 且 Alice 和 Bob 的拟合效果更好。

5.2 信道特征量化模块

本文采用了综合性能较好的分块格雷量化算法[27]。最终选择的 20MHz 带宽能捕获到 64 个子载波的信道特征, 根据 802.11n 协议相关说明, 最终含有效数据的子载波序号为–26 至–1 和 1 至 26, 共 52 个。分块格雷量化算法的流程为:

1) 设备对所获一轮信道探测的一组信道测量值按升序或降序排列;

2) 设备将排序后的信道测量值尽可能均分为 $2k$ 个数据块, k 为格雷编码位数。每个数据块按顺序标上序号, 数据块里的元素标号与其所属块序号一致;

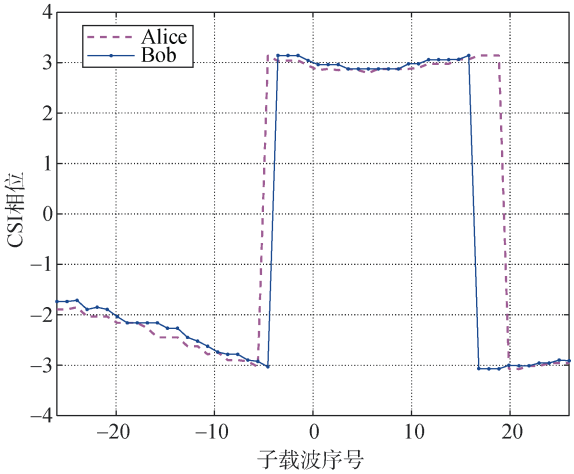
3) 按原有的信道测量值分布顺序使用 k 位的格雷编码将每个数据元素量化为 0,1 序列。

5.3 信息协商模块

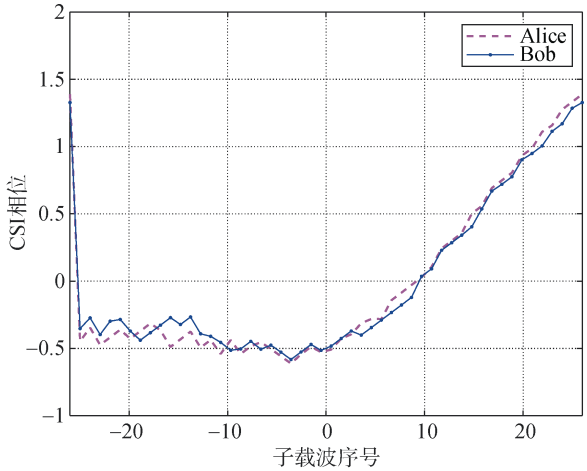
受硬件不对称等影响, 双方信道测量值经过量化后仍可能有不相同的密钥位, 如果双方不同的密钥位数目较少, 则可以等同于噪声对一般性通信过程的影响, 可以使用纠错算法来纠正。密钥协商通常使用纠错码来实现。BCH 纠错码能够利用不同的本原多项式实现多位纠错。根据其原理, 以及每组数据量化所得原始密钥长度, 设想不一致率不能超过 20%, 以免因为频繁协商增加信息泄漏风险, 因而将纠错位数上限设置为 22 位。

6 系统实验及分析

前文已经从理论上证明了本文提出方法的有效性, 本章将从实验角度进行实际效果验证。以评估不同场景下的密钥一致率、生成速率、随机性和安全性为目标, 设计了三种实验场景。实验环境均为空旷的办公室, 实验设备为 2 个 ESP32-S3 开发板, 分别作为 Alice 和 Bob, 设备工作在 2.4GHz 频段, 带宽为 20MHz。一台搭载 Ubuntu 20.04 系统的笔记本用于实时监视捕获的 CSI 数据并分析。发包率均设为每秒 5 个, 统计时间为 100 秒。为了保证实验结果的准确性和有效性, 使得数据具有普遍性意义, 我们提取了 5 组数据进行结果分析。首先从第 1 号数据帧开始, 提取连续 52 帧作为第 1 组, 从第 100 号开始连取连续 52 帧数据作为第 2 组, 从第 200 号开始连取连续 52 帧数据作为第 3 组, 以此类推, 共选取第 5 组数据进行实验。然后选取密钥不一致率、随机性测试、密钥生成速率、信道信息泄露率和窃听方密



(a) 原始相位效果图
(a) Original phase rendering



(b) 处理后相位效果图
(b) Phase rendering after processing

图 11 原始 CSI 相位和预处理后相位对比图
Figure 11 Comparison diagram of original CSI phase and preprocessed phase

钥错误率作为密钥评估指标, 最后在不同场景下对实验结果进行分析。

6.1 场景设计

本节以分析不同场景下的密钥随机性、提高密钥的可用性为目标, 设置了一个静态和两个动态场景。如图 12 所示, 静态场景为无人走动干扰; 第一种动态场景为人员在垂直两个设备连接线的方向上, 按 1 字行走; 第二种动态场景为人员环绕两个设备进行 8 字行走, 交叉点在两个设备的正中间。在后续实验中分别称为动态场景 1、动态场景 2, 所有场景中 Alice 和 Bob 均相距 1.5m。

6.2 密钥评估指标

6.2.1 密钥不一致率

本文将密钥不一致率(Key disagreement rate, KDR)定义为: 合法通信双方 Alice 和 Bob 在进行密钥生成时, 相同索引位置上的比特不一致的数量占密钥总位数的百分比。如公式(6.1)所示:

$$KDR = \frac{L_{diff}}{L} \tag{6.1}$$

其中 L_{diff} 为密钥位不一致的数量, L 为密钥总位数。密钥不一致率的高低表明了信道的互易性, 互易性不够高会导致密钥协商不成功, 使得通信双方无法生成一致密钥, 通信进程被拖长。

6.2.2 密钥随机性

随机性是一项检验密钥生成系统性能的非常重要的标准, 它反映了密钥序列中“1”, “0”比特值分布的均匀程度。目前大多数研究人员使用美国国家标准与技术研究所(National Institute of Standards and Technology, NIST)提供的统计随机性测试套件来对密钥的随机性进行检测。该套件用来检测随机数生成器(Random number generators, RNGs)和伪随机数生成器(Pseudorandom number generators, PRNGs)所生成的序列的随机性, 而密钥生成系统以信道特征作为随机源生成密钥, 本质也是随机数的生成, 因此本文也使用该套件进行检测。

6.2.3 密钥生成速率

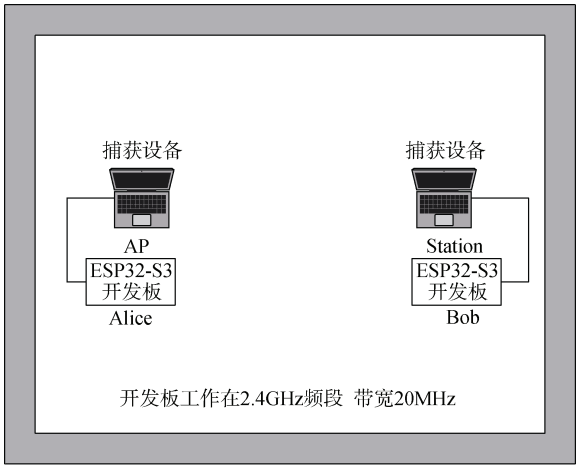
密钥生成速率是评价密钥安全性能的重要指标, 定义为单位时间内生成的密钥比特数, 密钥生成速率的表达式如公式(6.2)所示[28]:

$$KGR = \frac{L}{t} \tag{6.2}$$

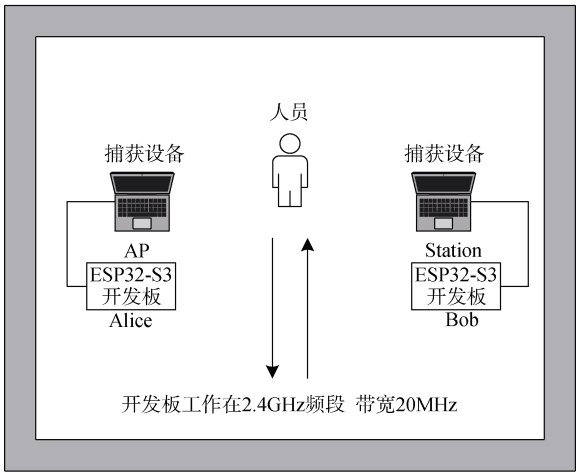
其中, L 为生成的密钥长度, t 为消耗的时间。

6.2.4 信道信息泄露率

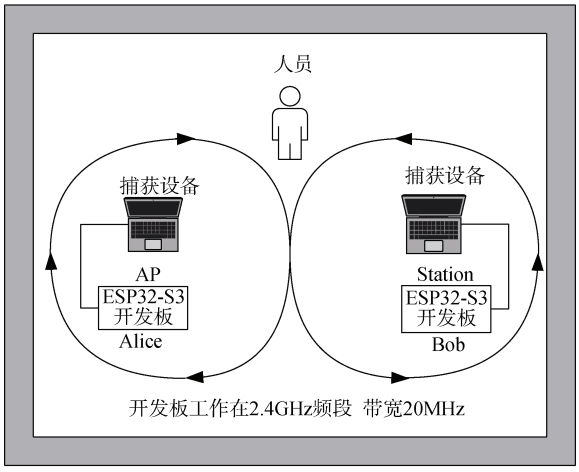
在信道探测过程中, 为了分析各种通信场景下 Eve 在不同位置能够获取到的合法信道相关的信息量, 我们在窃听者存在和不存在两种场景中分别计算密钥容量, 并计算由这两种密钥容量定义的信道



(a) 静态场景示意图
(a) Schematic diagram of a static scene



(b) 动态场景1示意图
(b) Schematic diagram of dynamic scenario 1



(c) 动态场景2示意图
(c) Schematic diagram of dynamic scene 2

图 12 实验场景示意图

Figure 12 Schematic diagrams of experimental scene

信息泄露率。密钥容量是合法通信双方提取安全密钥的速率上界, 可以从信息论的角度定量描述密钥提取性能。$\hat{H}_A$ 和 $\hat{H}_B$ 分别表示 Alice 和 Bob 对合法信道的估计, $\hat{H}_E$ 表示 Eve 通过监听合法节点的信道探测帧获取的信道估计。在不考虑窃听 Eve 的情况下, 合法通信双方可以达到的密钥容量, 即为双方信道估计之间的互信息, 如公式(6.3)所示:

$$C = I\left(\hat{H}_A;\hat{H}_B\right) = \log_2 \frac{\left|R_A\right|\left|R_B\right|}{\left|R_{AB}\right|} \tag{6.3}$$

其中, $|\cdot|$ 表示行列式, $R_{X_1X_2\ldots X_n} = Cov\left(\hat{H}_{X_1},\hat{H}_{X_2},\ldots,\hat{H}_{X_n}\right)$ 为信道协方差矩阵, $X_1,X_2,\ldots,X_n \in \{A,B,E\}$。

进一步考虑窃听者 Eve 存在的情况, 此时 Alice 和 Bob 之间的安全密钥容量如公式(6.4)所示:

$$C_s = \min\left\{I\left(\hat{H}_A;\hat{H}_B\right), I\left(\hat{H}_A;\hat{H}_B\middle|\hat{H}_E\right)\right\} \tag{6.4}$$

其中, $I\left(\hat{H}_A;\hat{H}_B\middle|\hat{H}_E\right) = \log_2 \frac{\left|R_{AE}\right|\left|R_{BE}\right|}{\left|R_E\right|\left|R_{ABE}\right|}$。

因此, 信道信息泄露率如(6.5)所示[29]:

$$\gamma_n = 1 - \frac{C_s}{C} \tag{6.5}$$

其中, C 是无窃听情况下密钥容量, C_s 是有窃听情况下密钥容量。

6.2.5 窃听方密钥错误率

在信息协商阶段, Alice 和 Bob 为了消除量化后所得初始密钥中的不一致比特, 需要在公共信道上交换密钥校验值, 这一过程也会向 Eve 泄露一定量的信息。为了评估经过信息协商后所得密钥的安全性, 本文定义窃听方密钥错误率为 Eve 仿造的密钥与合法双方提取密钥中不一致的比特数占密钥长度的比例, 公式如(6.6)所示:

$$\gamma_e = \frac{\sum\left(K_{AB} \oplus K_E\right)}{L} \tag{6.6}$$

其中, K_{AB} 和 K_E 分别为合法双方和 Eve 在信息协商后所得到的密钥, L 为密钥长度, $\oplus$ 表示逐位异或运算。

6.3 实验结果分析

在双载波商模型中, 我们选取的收发设备为 ESP32-S3 芯片, 其在传输数据时使用 UDP 协议。为了论证此方法的正确性和有效性, 分别在 6.1 中所设计的三种实验场景进行 CSI 采集, 生成一致性密钥, 最终对各项密钥性能指标进行评估。

6.3.1 互易性数据量对比

本节将基于传统 CSI 方法和基于双载波幅度比值、相位差值方法在相同时间内的所产生的互易性数据量进行了对比, 如图 13、图 14 所示, 通过对不同子载波采样序列的商数运算, 使得相同时间内的互易性数据量成倍数增长, CSI 信息利用率有显著提升。

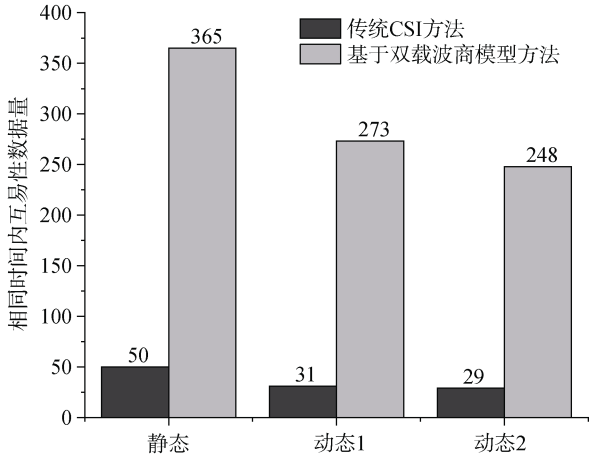


图 13 幅度比值互易数据量对比图

Figure 13 Comparison chart of amplitude ratio reciprocity data volume

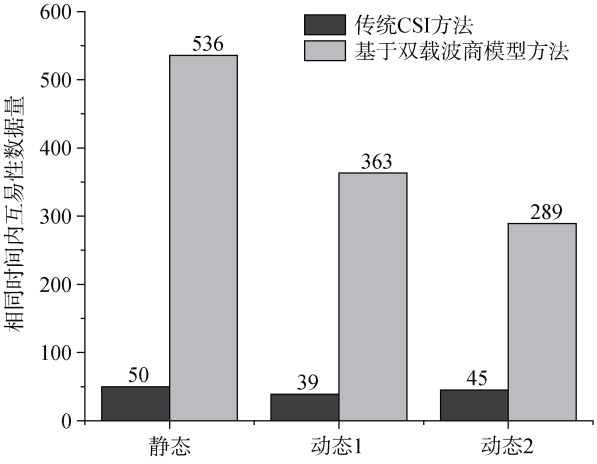


图 14 相位差值互易数据量对比图

Figure 14 Comparison chart of phase difference reciprocity data volume

6.3.2 密钥性能分析

1) 密钥不一致率

计算两种方法在上述的每个场景下的平均密钥不一致率, 结果如图 15 所示。

实验结果显示, 这两种方法在三种场景下的不一致率均低于 20%, 在可纠错的范围内。

2) 密钥随机性

选取 10 组 104 位密钥拼接成 1040 位密钥进行验证。由于数据量限制, 选取了其中 7 种测试项目, P 值>0.01 即通过测试。测试结果如表 2、表 3 所示,

7 个项目全部通过测试, 相比于静态环境, 动态场景中有 4 项 P 值升高, 代表密钥随机性提升。

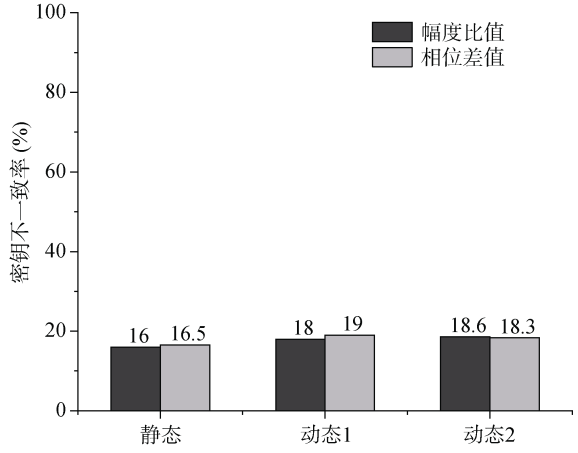


图 15　两种方法不一致率统计图

Figure 15　Statistical chart of inconsistency rate between two methods

3) 密钥生成速率

在三种测试场景下的平均密钥提取速率如图 16、图 17 所示, 可以看出基于双载波幅度比值的密钥生成方法与传统 CSI 方法相比, 可以将密钥生成速率提升 3～7 倍, 基于双载波相位差值密钥生成方法可以将生成速率提升 6～9 倍。此结果说明本文提出的密钥生成方案能够保证密钥的持续更新, 从而提供良好的前向安全性。此外, 动态场景下的密钥生成速率比静态时略低, 原因是周围有人员运动时信道互易性有所下降, 合法双方的信道估计相似性变差, 经过对信道估计的量化后得到的一致密钥比特数减少, 使得密钥提取速率有一定程度的降低。

4) 密钥更新率

在 4.4 节中, 将密钥更新率定义为单位时间内的密钥更新次数。如图 18、图 19 所示为三种场景中, 基于双载波商模型方法、基于双载波商模型与集合划分相结合方法的密钥更新率, 可以看出, 与集合划分方法相结合可以在一定程度上提高密钥更新率。并且在动态场景中, 因为有人员走动, 影响了环境复杂性, 所以在动态场景中的密钥更新速率更快。

6.3.3　密钥安全性分析

使用 2.4GHz 的频段时, 对应半波长约为 12 厘米, 将 Eve 放置在距离 Bob 约 20 厘米处, 超过半波长。对比三种场景中 Eve 捕获到的 MAC 地址为 Bob 的 CSI 幅度数据以及 Alice、Bob 端的 CSI 幅度、相位数据。

1) 三种场景下窃听对比图

Eve 在三种场景下的窃听实验结果如图 20 所示, 为保证实验结果的准确性, Alice、Bob 选取编号为 1

表 2　基于幅度比值方法生成密钥随机性测试表

Table 2　Key randomness test table generated based on amplitude ratio method

检测项目	静态		动态 1		动态 2	
频数检验	0.012555		0.026500		0.026500	
块内频数检验	0.992613		1.000000		0.994051	
游程检验	0.637417		0.637417		0.363370	
累加和检验	0.053001	0.025110	0.053001	0.025110	0.053001	0.025110
近似熵检验	0.983165		1.000000		0.981762	
离散傅里叶变换检测	0.373037		0.030513		0.030513	
序列检测	0.290153	0.156310	0.999787	0.000290	0.980389	0.561915

表 3　基于相位差值方法生成密钥随机性测试表

Table 3　Key randomness test table generated based on phase difference method

检测项目	静态		动态 1		动态 2	
频数检验	0.018362		0.037551		0.026500	
块内频数检验	1.000000		1.000000		1.000000	
游程检验	0.583677		0.583677		0.608921	
累加和检验	0.053001	0.025110	0.053001	0.025110	0.053001	0.025110
近似熵检验	0.983165		0.981762		0.990437	
离散傅里叶变换检测	0.036791		0.378842		0.378842	
序列检测	0.317909	0.110042	0.990598	0.001035	0.983067	0.763535

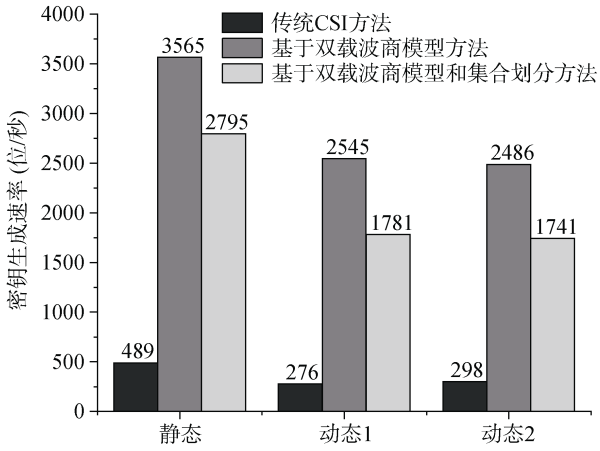


图 16 幅度比值密钥生成速率对比图

Figure 16 Amplitude ratio key generation rate comparison chart

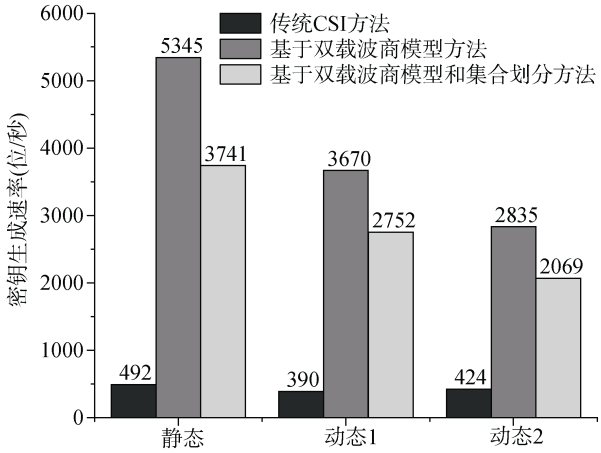


图 17 相位差值密钥生成速率对比图

Figure 17 Phase difference key generation rate comparison chart

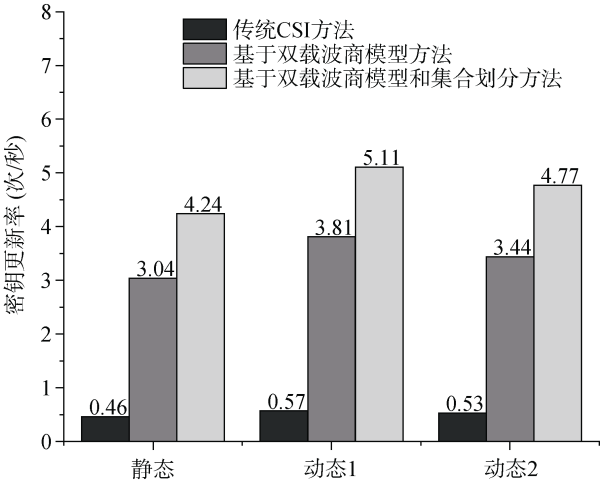


图 18 幅度比值密钥更新率对比图

Figure 18 Amplitude ratio key update rate comparison chart

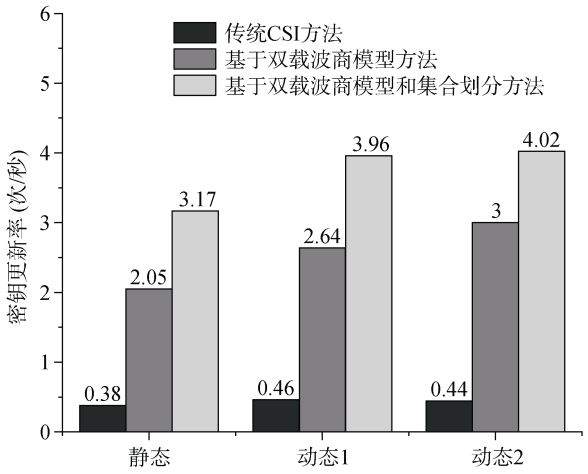


图 19 相位差值密钥更新率对比图

Figure 19 Phase difference key update rate comparison chart

的数据帧为首帧, Eve 选取第一次捕获到的 Bob 端的数据帧为首帧, 用编号相同的子载波做幅度比值和相位差值。

可以看出, Alice 和 Bob 具有高度相似的信道状态信息, 而 Eve 获取的信道特征与 Alice 和 Bob 的差别很大, 基本不具备互易性, 即 Eve 与 Bob 密钥不一致位数远超过能够纠错的 22 位, 无法获得与 Alice 和 Bob 一致的密钥。

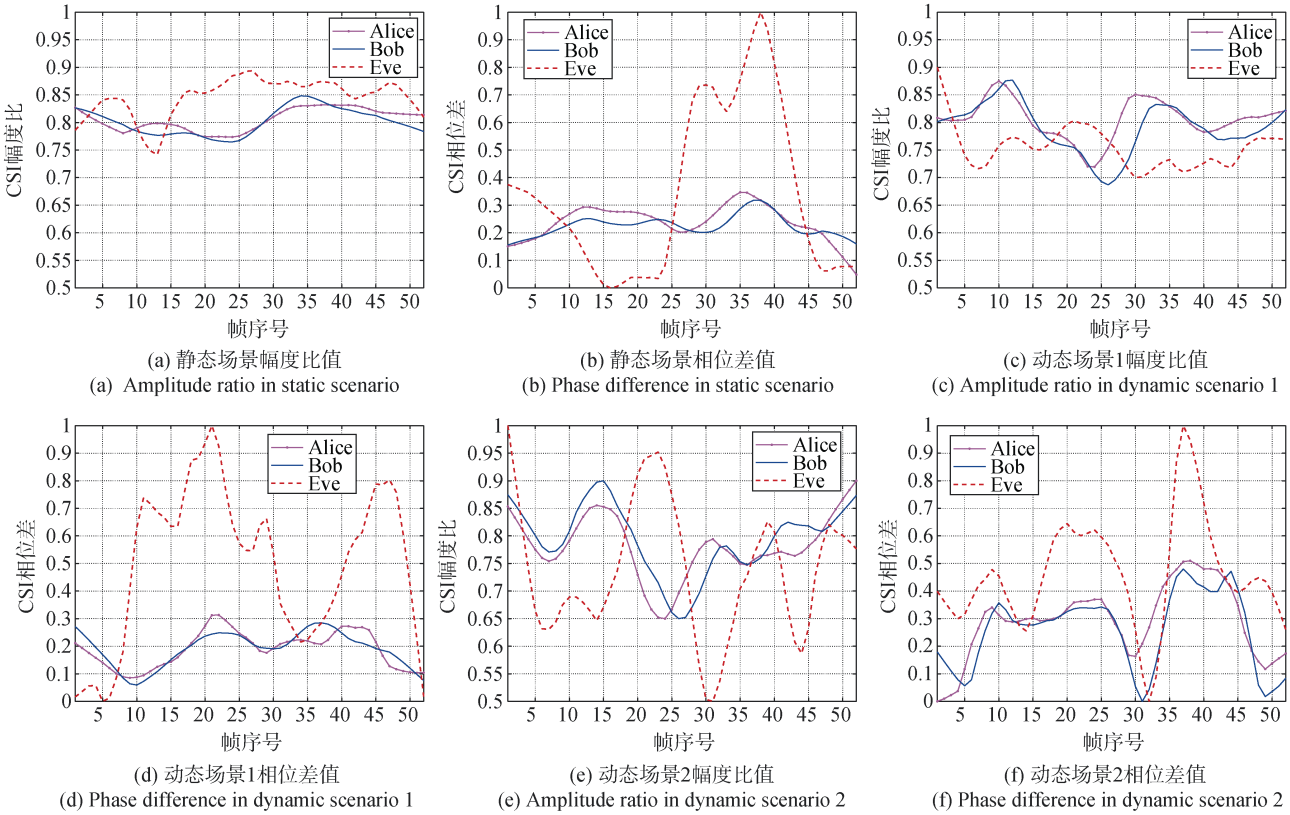
2) 信道信息泄露率

根据 6.2.4 中所述公式, 计算在三种场景下, 基于幅度比值密钥生成方法和基于相位差值密钥生成方法的信道信息泄露率, 并与传统 CSI 方法进行了对比, 结果如图 21、图 22 所示。

从图中可以看出, 本文方法的整体信道信息泄露率在 0.098 以下, 在动态场景中泄露率更低。而传统 CSI 密钥生成方法的信道信息泄露率波动较大, 在 0.068 到 0.348 之间。因为传统 CSI 方法利用单帧 CSI 量化生成密钥, 所以窃听者很容易捕获到单帧数据, 使得密钥泄露率升高。而本文提出的方法利用的是在每个子载波上的连续采样序列, 这对窃听者来说是一个挑战, 一旦没有捕获到连续的数据帧, 那么 Eve 获得的采样序列将与合法节点序列大不相同, 造成 Eve 伪造的密钥与合法通信双方密钥的不一致率飙升, 无法生成一致性密钥, 导致窃听失败。综上所述, 本文提出的密钥生成方法明显优于传统 CSI 方法。

3) 窃听方密钥错误率

根据第 6.2.5 所述公式, 计算两种方法在三种场景下的窃听方密钥错误率, 即 Eve 仿造的密钥与合



(a) 静态场景幅度比值
(a) Amplitude ratio in static scenario

(b) 静态场景相位差值
(b) Phase difference in static scenario

(c) 动态场景1幅度比值
(c) Amplitude ratio in dynamic scenario 1

(d) 动态场景1相位差值
(d) Phase difference in dynamic scenario 1

(e) 动态场景2幅度比值
(e) Amplitude ratio in dynamic scenario 2

(f) 动态场景2相位差值
(f) Phase difference in dynamic scenario 2

图 20 三种场景下窃听结果对比图

Figure 20 Comparison of the wiretap results in three scenarios

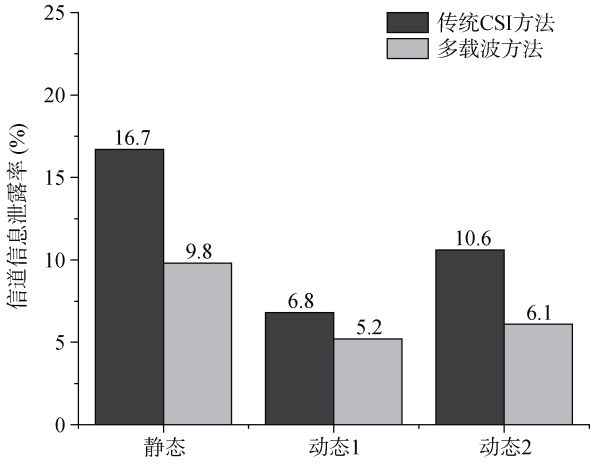


图 21 幅度比值信息泄露率对比图

Figure 21 Comparison chart of amplitude ratio information leakage rate

法双方提取密钥中不一致的比特数占密钥长度的比例, 实验结果如图 23 所示。

从图中可以看出, Eve 通过窃听合法节点的通信内容生成的密钥错误率均在 40%以上, 这说明在实际通信场景下, 即使窃听节点与合法节点十分接近, 也很难完全仿造出合法双方提取的密钥; 另一方面, 动态环境下窃听方密钥错误率高于静态环境, 这说明本文的方法在信道环境多变的场景下具有更高的安全性。

6.4 小结

结合本文提出的双载波商模型和子载波集合划分方法, 进行了大量的测试实验, 并对实验结果进行了全面分析。首先, 通过分析传统 CSI 方法和基于双载波商模型方法的互易性数据量和密钥生成速率, 验证了本文提出的可以显著提升 CSI 信息利用率和密钥生成速率; 其次, 通过分析密钥不一致率和随

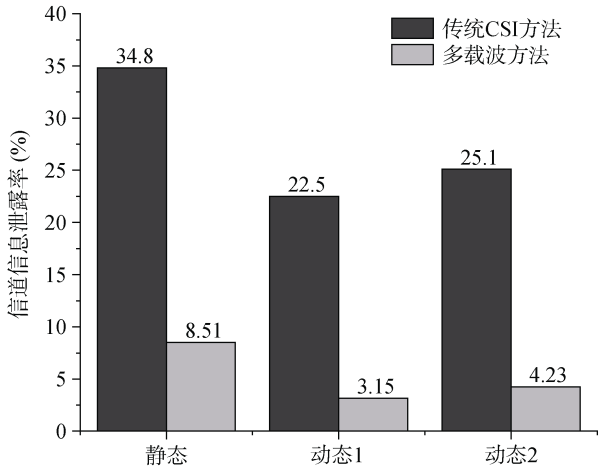


图 22 相位差值信息泄露率对比图

Figure 22 Comparison chart of phase difference information leakage rate

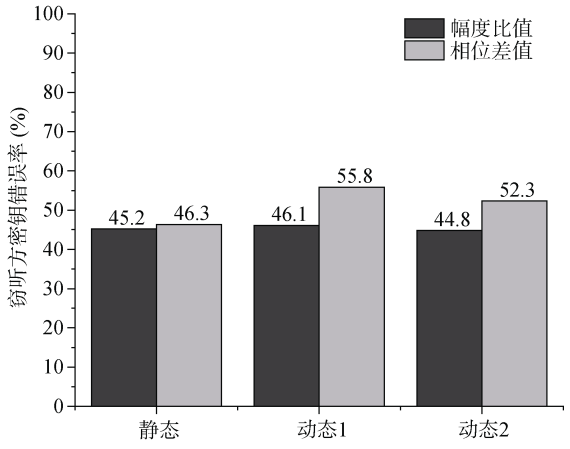


图 23 窃听方密钥错误率对比图

Figure 23 Comparison chart of eavesdropper's key error rate

机性, 验证本文提出的方法在三种场景下的密钥不一致率均低于20%, 在可纠错的范围内, 并且最终生成的密钥通过了随机性测试; 然后, 通过分析传统 CSI 方法、基于双载波商模型方法以及基于双载波商模型和集合划分方法的密钥更新率, 说明本文提出的集合划分方法可以有效加快密钥更新速率; 最后, 对比了传统 CSI 方法和本文提出的方法在不同场景下的密钥安全性, 实验结果说明本文的方法具有更低的信道信息泄露率和更高的窃听密钥错误率, 验证了基于双载波商模型和集合划分的密钥生成方案在实际通信场景中应用的可行性。

7 总结

在本文中, 首次提出了双载波商数模型, 开创性地使用 CSI 商数作为生成密钥的随机源, 并且在保证双方密钥一致率的情况下, 显著提升了 CSI 信息利用率和密钥生成速率。随后, 针对基于双载波商模型密钥生成方法中涌现出的大量密钥重复段问题, 本文依据相关系数, 提出了一种子载波划分集合方法, 并提取内具有代表性的序列, 可以在一定程度上消除密钥重复段, 加快密钥更新率, 增强密钥安全性。最后, 本文搭建了一个验证性实验系统来检验我们所提方法的可行性与效果。

实验证明, 利用双载波商模型可以将提升密钥生成速率提升 3～9 倍, 子载波集合划分方法可以降低相近轮次生成密钥的相关性, 将密钥更新率加快 2～3 倍, 提升合法通信双方会话的效率和安全性。

本文聚焦密钥生成技术中的速率提升研究, 力求尽可能高效的利用捕获到的 CSI 信息, 以助力于提升通信安全和效率的研究, 希望本文的工作可以为无线信道密钥生成技术领域添砖加瓦。

参考文献

[1] Mathur S, Trappe W, Mandayam N, et al. Radio-telepathy: extracting a secret key from an unauthenticated wireless channel[C]. *Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008:128-139.

[2] Ren K, Su H, Wang Q. Secret Key Generation Exploiting Channel Characteristics in Wireless Communications[J]. *IEEE Wireless Communications*, 2011, 18(4): 6-12.

[3] Li G Y, Yu J B, Hu A Q. Research on Physical-Layer Security Based on Device and Channel Characteristics[J]. *Journal of Cryptologic Research*, 2020, 7(2): 224-248.
(李古月, 俞佳宝, 胡爱群. 基于设备与信道特征的物理层安全方法[J]. *密码学报*, 2020, 7(2): 224-248.)

[4] Li G Y, Hu L, Staat P, et al. Reconfigurable Intelligent Surface for Physical Layer Key Generation: Constructive or Destructive?[J]. *IEEE Wireless Communications*, 2022, 29(4): 146-153.

[5] Chen C, Jensen M A. Secret Key Establishment Using Temporally and Spatially Correlated Wireless Channel Coefficients[J]. *IEEE Transactions on Mobile Computing*, 2011, 10(2): 205-215.

[6] Xiao S F, Guo Y F, Huang K Z, et al. Cooperative Group Secret Key Generation Based on Secure Network Coding[J]. *IEEE Communications Letters*, 2018, 22(7): 1466-1469.

[7] Qin D R, Ding Z. Exploiting Multi-Antenna Non-Reciprocal Channels for Shared Secret Key Generation[J]. *IEEE Transactions on Information Forensics and Security*, 2016, 11(12): 2693-2705.

[8] Zhang S J, Jin L, Lou Y M, et al. Secret Key Generation Based on Two-Way Randomness for TDD-SISO System[J]. *China Communications*, 2018, 15(7): 202-216.

[9] Aldaghri N, Mahdavifar H. Physical Layer Secret Key Generation in Static Environments[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 2692-2705.

[10] Jin L, Zhang S J, Lou Y M, et al. Secret Key Generation with Cross Multiplication of Two-Way Random Signals[J]. *IEEE Access*, 2019, 7: 113065-113080.

[11] Yao J L, Li C, Ren H P, et al. Chaos-Based Wireless Communication Resisting Multipath Effects[J]. *Physical Review E*, 2017, 96(): 032226.

[12] da Cruz P I, Suyama R, Loiola M B. Increasing Key Randomness in Physical Layer Key Generation Based on RSSI in LoRaWAN Devices[J]. *Physical Communication*, 2021, 49: 101480.

[13] Hashem S, Jasim A. A proposed modification on RC4 algorithm by increasing its randomness[J]. *Journal of Al-Rafidain University College for Sciences*, 2017 (1): 349-372.

[14] Soni A, Upadhyay R, Kumar A. Wireless Physical Layer Key Generation with Improved Bit Disagreement for the Internet of Things Using Moving Window Averaging[J]. *Physical Communication*, 2019, 33: 249-258.

[15] Li G Y, Hu A Q, Zhang J Q, et al. Security Analysis of a Novel Artificial Randomness Approach for Fast Key Generation[C]. *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017: 1-6.

[16] Zeng K, Wu D, Chan A, et al. Exploiting Multiple-Antenna Diversity for Shared Secret Key Generation in Wireless Networks[C].

2010 Proceedings IEEE INFOCOM, 2010: 1-9.

[17] Xiao L, Hong S Y, Xu S Y, et al. IRS-Aided Energy-Efficient Secure WBAN Transmission Based on Deep Reinforcement Learning[J]. *IEEE Transactions on Communications*, 2022, 70(6): 4162-4174.

[18] Zhang J Q, Marshall A, Woods R, et al. Efficient Key Generation by Exploiting Randomness from Channel Responses of Individual OFDM Subcarriers[C]. *IEEE Transactions on Communications*, 2016: 2578-2588.

[19] Pincus S. Approximate Entropy (ApEn) as a Complexity Measure[J]. *Chaos*: *An Interdisciplinary Journal of Nonlinear Science*, 1995, 5(1): 110-117.

[20] Ponnaluri S, Azimi-Sadjadi B, Hue Y K, et al. A Practical Wireless Reciprocity-Aware Key Establishment Protocol[C]. *MILCOM 2016 - 2016 IEEE Military Communications Conference*, 2016: 1107-1113.

[21] Deek L, Garcia-Villegas E, Belding E, et al. Intelligent Channel Bonding in 802.11n WLANs[J]. *IEEE Transactions on Mobile Computing*, 2014, 13(6): 1242-1255.

[22] Wei Y C, Zeng K, Mohapatra P. Adaptive Wireless Channel Probing for Shared Key Generation[C]. *2011 Proceedings IEEE INFOCOM*, 2011: 2165-2173.

[23] Soni A, Upadhyay R, Kumar A. Low Complexity Preprocessing Approach for Wireless Physical Layer Secret Key Extraction Based on PCA[J]. *Wireless Personal Communications*, 2022, 125(3): 2865-2888.

[24] Tan Q Y, Huang S L, Liu S J. Retracted a Method for Detecting Amplitude-Phase Joint Characteristic Parameters of Wireless Channel for Generating Key Parameters[J]. *Complexity*, 2021, 2021(1): 9951742.

[25] Tse D, Viswanath P. Fundamentals of Wireless Communication[M]. Cambridge, UK: Cambridge University Press, 2005.

[26] Qian K, Wu C S, Yang Z, et al. PADS: Passive Detection of Moving Targets with Dynamic Speed Using PHY Layer Information[C]. *2014 20th IEEE International Conference on Parallel and Distributed Systems*, 2014: 1-8.

[27] Ye C X, Mathur S, Reznik A, et al. Information-Theoretically Secret Key Generation for Fading Wireless Channels[J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(2): 240-254.

[28] Wallace J. Secure Physical Layer Key Generation Schemes: Performance and Information Theoretic Limits[C]. *2009 IEEE International Conference on Communications*, 2009: 1-5.

[29] Guo D K, Xiong J, Gao Y W, et al. Design and Implementation of Key Extraction Scheme Based on Wireless Channel State Information[J]. *Journal of Signal Processing*, 2021, 37(3): 336-348.
(郭登科, 熊俊, 高玉威, 等. 基于无线信道状态信息的密钥提取方案设计与实现[J]. *信号处理*, 2021, 37(3): 336-348.)

姜禹 于 2009 年在东南大学信号与信息处理专业获得博士学位。现任东南大学网络空间安全学院副教授。研究领域为无线感知、物理层安全。研究兴趣包括: 物理层安全、无线网络安全、RFID 技术。Email: jiangyu@seu.edu.cn

胡爱群 于 1993 年在东南大学信号与信息处理专业获得博士学位。现任东南大学网络空间安全学院教授。研究领域为无线物理层安全、移动通信内生安全理论与技术。研究兴趣包括: 无线网络安全、物理层安全技术。Email: aqhu@seu.edu.cn

王禹淳 于 2020 年在吉林大学计算机科学与技术专业获得理学学士学位。现在东南大学网络与信息安全专业攻读硕士学位。研究领域为物理层安全、密钥分发。研究兴趣包括: 物理层安全、无线信道密钥。Email: wangyuchun@seu.edu.cn