RC4 算法分析

于成俊

(南开大学网络空间安全学院 天津 300350)

(**专业：密码科学与技术 学号：2112066**)

摘 要 本文系统地介绍了 RC4 算法，首先描述了其基本流程，主要包括密钥调度算法和伪随机生成算法两个部分。然后进行了安全性分析，涵盖了前人对 RC4 伪随机性的研究，以及相关攻击的讨论，如弱密钥攻击、区分攻击和状态猜测攻击等。接着，探讨了 RC4 算法在实际应用中的场景和现状，介绍了它的辉煌与衰落，并详细讨论了 RC4 的相关改进，如利用椭圆曲线和 Serpent 函数的 S 盒进行的改进。通过对 RC4 算法的全面分析，本文旨在帮助读者更好地理解 RC4 算法的原理、应用以及当前面临的挑战，为密码学研究和网络安全领域的进一步探索提供参考。

关键词 RC4、伪随机性、PRGA、S 盒、密钥

Analysis of the RC4 Algorithm

Yu Cheng-jun

(College of Cybersecurity, Nankai University, Tianjin 300350, China)

Abstract This paper systematically introduces the RC4 algorithm. First, it describes its basic process, which mainly includes two parts: the key scheduling algorithm and the pseudo-random generation algorithm. Then, it conducts a security analysis, covering the previous research on the pseudo-randomness of RC4, as well as the discussion of related attacks, such as weak key attacks, differentiation attacks, and state guessing attacks. Then, it explores the scenarios and current status of the RC4 algorithm in practical applications, introduces its glory and decline, and discusses in detail the related improvements of RC4, such as the improvements made using the S-box of elliptic curves and Serpent functions. Through a comprehensive analysis of the RC4 algorithm, this paper aims to help readers better understand the principles, applications, and current challenges of the RC4 algorithm, and provide a reference for further exploration in the fields of cryptography research and network security.

Key words RC4, pseudorandomness, PRGA,S-box, key

1 引言

1.1 背景介绍

RC4（Rivest Cipher 4）是一种由罗纳德·里维斯特（Ronald Rivest）于 1987 年设计的流加密算法。凭借其实现简单、加密速度快等优点，RC4 在过去几十年中被广泛应用于多种加密场景，包括 SSL/TLS、WEP 和 Microsoft。然而，随着密码分析技术的不断发展，RC4 的诸多安全缺陷逐渐暴露，尤其是在高强度安全需求的场合中，其安全性受到严重质疑。

1.2 研究目的和意义

本文旨在全面探讨 RC4 加密算法，深入解析其工作原理，评估其优缺点，并总结前人在 RC4 攻击方面的详细分析，以揭示其安全性隐患。同时，本文将探讨如何改进 RC4 算法，分析其在现代加密技术中的地位，综述最新的研究进展和 RC4 算法的应用现状。通过这些分析和讨论，本文不仅为密码学

研究人员提供了 RC4 的理论基础，也为网络安全从业者在实际工作中更好地理解和应对 RC4 相关安全问题提供了重要参考。

1.3 论文结构

本文的结构安排如下：第一部分为引言；第二部分将简要介绍 RC4 算法的基本原理；第三部分将重点讨论 RC4 的各种攻击方法，深入剖析其原理和实际效果；第四部分将讨论 RC4 算法的应用场景，并分析其优缺点；第五部分将评估 RC4 在当前加密技术中的地位，并介绍最新的研究进展；最后，第六部分将总结全文的研究成果，并提出未来的研究方向和建议。

2 RC4 算法概述

RC4 算法的核心是通过伪随机数生成器（PRNG）生成密钥流，然后将密钥流与明文进行异或运算以生成密文。RC4 算法的具体步骤分为两个主要部分：密钥调度算法（Key Scheduling Algorithm，KSA）和伪随机生成算法（Pseudo-Random Generation Algorithm，PRGA）。

2.1 密钥调度算法（KSA）

KSA 用于初始化状态向量 S 并混合初始密钥，其具体步骤如下：

初始化：首先创建一个长度为 256 的状态向量 S，S 的每个元素 S[i]初始化为 i，即 S[0] = 0, S[1] = 1, ..., S[255] = 255。然后再创建一个长度与密钥相同的密钥向量 K。假设密钥长度为 L 字节，密钥为 K[0]到 K[L-1]。

置换：使用密钥 K 对状态向量 S 进行置换操作。置换过程如下：

过程 1. 置换

```
j = 0
FOR i FROM 0 to 255:
    j = (j + S[i] + K[i % L]) % 256
    SWAP S[i] AND S[j]
```

经过上述步骤，状态向量 S 被初始化并与密钥混合，准备进入生成密钥流的阶段。

2.2 伪随机生成算法（PRGA）

PRGA 用于生成伪随机密钥流，用以加密或解密数据。其具体步骤如下：

初始化：初始化两个索引变量 i 和 j，初始值均为 0。

生成密钥流：不断生成伪随机字节，直到产生足够的密钥流来加密或解密整个数据。生成过程如下：

过程 2. 生成密钥流

```
WHILE generating OUTPUT:
    i = (i + 1) % 256
    j = (j + S[i]) % 256
    SWAP S[i] AND S[j]
    K = S[(S[i] + S[j]) % 256]
    OUTPUT K
```

在上述过程中，输出的 K 值即为伪随机密钥流中的一个字节。

2.3 加密与解密

RC4 的加密和解密过程相同，均为将明文或密文与生成的密钥流进行异或运算。

如果要加密，则先生成与明文长度相同的密钥流，然后对明文的每个字节与密钥流对应字节进行异或运算，生成密文；如下：

过程 3. 加密

```
Ciphertext[i] = Plaintext[i] ^ KeyStream[i]
```

如果要解密，则先生成与密文长度相同的密钥流，然后对密文的每个字节与密钥流对应字节进行异或运算，恢复明文。如下：

过程 4. 解密

```
Plaintext[i] = Ciphertext[i] ^ KeyStream[i]
```

3 RC4 算法的安全性分析

RC4 算法是一种流密码系统，其生成的序列具有高度的随机性。尽管使用常规统计方法难以检测其输出，但完全随机化仍然无法实现，该算法的安全性依赖于生成伪随机密钥序列的密码生成器的特性[1]。

影响 RC4 算法的密钥流序列随机性的三个因素包括[2]：1)S 盒中的初始值分布均匀程度；2)索引指针 i，j 的分布均匀程度；3)根据指针输出的结果分布均匀程度。RC4 算法的内部状态由一个包含 256 Byte 的 S 盒和指针 i，j 组成，输出的密钥流序

列由 S 盒中的初始值和指针 i，j 共同确定，因此输出不重复的值至多Z_2^8个，范围较小，密钥流序列的随机性较差。

RC4 的伪随机性已经被广泛研究。Golic 曾研究 RC4 的线性统计弱点[3]，指出通过利用随机值可以将 RC4 的输出与随机序列区分开来。此结果后来被 Fluhrer 和 McGrew 进一步改进[4]，他们通过分析连续两个输出值和 i≡t(mod N)的已知值的分布，发现了微小的偏差，并利用信息理论证明这些偏差可以用来区分 RC4 和真正的随机序列。2013 年，AIFardan 等人在信息安全顶级会议 Usenix Security 发表了一篇论文，其中，他们利用统计的方法找到了 RC4 算法密钥流中存在的误差[5]，在此基础上，刘聪在 2016 年利用大数据平台[6]，进一步算出了 RC4 密钥流前 256 个每个位置上各字节出现的概率值，从而发现 RC4 算法的密钥流存在明显的单字节误差，例如，密钥流第二个字节输出为 0 的概率高达 1/128，远远高于其平均概率 1/256。因此 RC4 算法产生的密钥流并不是随机的，攻击者可能就会利用这个漏洞从而破解出明文。

由于 RC4 的伪随机性漏洞，RC4 易遭受多种攻击[7]，典型的有区分攻击、弱密钥攻击、错误引入攻击以及状态猜测攻击。

3.1 区分攻击

RC4 的区分攻击利用了 RC4 在生成密钥流时存在的非均匀性和偏差，来区分 RC4 生成的密钥流和真正的随机流，从而推测出一些密钥信息或部分明文。如果区分攻击的时间复杂度小于暴力破解的复杂度，就可以认定这个区分攻击是有效的[8]。例如，Mantin[8]等根据 RC4 输出的第一和第二个密钥子能被概率统计出来这一弱点，就设计了一个区分攻击方法，他们首先产生一组不同的输出密钥流，然后重点观察第一个和第二个密钥字，发现明文的第一个和第二个字分别以 P(Z_1=0)和 P(Z_2=0)的概率等于密钥流的第一个和第二个字，从而破解出明文。

3.2 弱密钥攻击

弱密钥攻击就是，找到 RC4 在特定密钥条件下产生的弱密钥，这类密钥会使 RC4 算法在初始化过程中产生显著的偏差，从而导致生成的密钥流具有非随机性，以此来进行攻击。FMS 攻击是最早对 RC4 提出的弱密钥攻击之一，它是由 Fluhrer、Mantin 和 Shamir[9]在 2001 年的研究中发现的，他们表明如果攻击者能够收集大量使用相同初始密钥（IV，Initialization Vector）和不同密钥的加密数据，就可以通过统计分析恢复密钥。除此之外，Klein[10]攻击等都属于弱密钥攻击。

3.3 错误引入攻击

错误引入攻击（Fault Injection Attack）是一种主动攻击技术，攻击者通过故意引入硬件或软件故障来破坏加密系统的正常操作，以此获取敏感信息或加速破解密钥。该技术利用了加密算法在处理异常条件时可能暴露的内部状态或密钥信息，从而实施攻击。可以向 RC4 状态表的第 t 个位置引入错误[7]，然后判断错误引入后产生的错误密钥字的类型，就可探测出 PRGA 的初始状态S_0值的分布。

3.4 状态猜测攻击

状态猜测攻击（State Guessing Attack）主要针对流加密算法如 RC4。攻击者通过猜测并验证加密算法内部状态的部分信息，逐步推测出整个内部状态，从而恢复密钥或密钥流。如果攻击者能获得一定数量 PRGA 输出的密钥流序列，以此为基础推测出 PRGA 的初始状态值，则根据这个初始状态就可产生密钥流序列，RC4 就可被破解。

4 RC4 算法的应用场景及现状

由于 RC4 算法具有高效的加密速度和简单的实现，曾在网络通信的加密中被广泛应用。在 802.11 无线网络标准中，RC4 被用于 WEP（有线等效隐私）协议，以保护无线数据传输的安全。它还被用于一些 VPN 解决方案中，用于加密互联网上传输的数据。在早期的即时通讯软件中，RC4 用于加密通信内容，以确保用户消息的隐私。此外，RC4 还曾用于 SSL/TLS 协议中，为 Web 浏览器与服务器之间的通信提供加密保护。

但是由于 RC4 存在安全漏洞，它面临这越来越多诸如以上的攻击方法，这些攻击方法都威胁到了 RC4 的安全性。随着安全性要求的不断提高和密码学领域的发展，加上其安全性的缺陷使得它在现代加密中不再推荐使用。许多标准组织和协议已经弃用或建议弃用 RC4。例如，在 TLS 协议中，IETF（互联网工程任务组）于 2015 年发布了 RFC 7465，正式禁止在 TLS 中使用 RC4。在 WEP 协议中，WEP 已被更安全的 WPA（Wi-Fi 保护访问）和 WPA2 取代。

尽管 RC4 已逐渐被弃用，但关于 RC4 的研究

仍在继续，主要集中在以下几个方面。一方面，研究人员不断优化现有的攻击方法，以提高攻击效率和成功率。另一方面，对 RC4 内部工作机制和偏差问题进行更深入的理论分析，以更好地理解其安全性问题。此外，一些研究尝试对 RC4 进行改进，如 Spritz、RC4A、RC4+（RC4 的一个变种），试图解决 RC4 的一些安全问题，尽管这些改进在实际应用中的效果有限，但这些变种和改进对流加密算法的研究和发展提供了宝贵的经验和启示，同时也促进了密码学领域的不断进步。

5 RC4 算法改进

RC4 算法可以将具有较好非线性的 Serpent 函数的 S 盒引入 KSA 的形成中[6]，来加强随机性。具体实现方法是，前面实现过程都不变，只是在完成 swap(s[i],s[j])操作后，对此时 RC4 算法的 S 盒进行置换，让其每个字节的八位分成两组同时经过同一个 Serpent 置换 S 盒，八个置换 S 盒会被循环使用，改进后的 KSA 具有较强的随机性。

RC4 算法还可以引入椭圆曲线、MD5 哈希算法来进行改进[2]。具体方法是利用椭圆曲线生成随机大素数和随机整数，并通过 MD5 哈希算法处理生成 128 位消息摘要，再选取其中 64 位数据送入伪随机数生成器产生新的 64 位伪随机数作为密钥 Key，随后用该密钥初始化 S 盒，通过多次遍历产生 RC4 的初始状态，不断更新指针并进行非线性变换来输出密钥流序列。这些改进增强了 RC4 算法的随机性和安全性，能够有效抵抗区分攻击、“受戒礼”攻击和状态猜测攻击等。

针对错误引入攻击，可以在 RC4 运算过程中添加一个自我检查的步骤以防止错误引入攻击的实施[11]。主要的思想是通过比较加密算法在交换前后的状态变化来判断是否发生了错误引入攻击。如果自我检查步骤确认了错误引入攻击的存在，那么就立即终止算法的运行；如果一切正常，没有发现攻击迹象，那么就继续执行算法生成密钥序列。

除了以上方法，还有 Paul and Preneel[12]提出的 RC4 改进算法——RC4A，和一种基于“动态可变的置换组合”的改进算法——VMPC[13]。

6 总结与展望

这篇文章详细介绍了 RC4 加密算法的流程以及相关的安全性分析，探讨了它容易受到的攻击以及相关的改进。曾经，RC4 因其实现简单和加密速度快而被广泛应用于各种加密场景，但由于存在弱密钥问题和可预测的密钥流问题，逐渐走向了淘汰。希望将 RC4 作为经典案例纳入密码学教材，通过研究其成功与失败的经验教训，帮助新一代密码学家更好地理解流加密算法的设计和分析。基于对 RC4 的经验教训，期望未来的密码算法设计者能够开发出更安全和高效的新型流加密算法，以满足现代信息安全的需求。

致 谢 首先，我想表达对贾老师这一学期的辛勤教学表示诚挚的感谢。贾老师的教学方式不仅严谨细致，而且充满了热情和耐心，让我对网络安全方面的知识产生了浓厚的兴趣。在贾老师的课堂上，我不仅学到了丰富的网络安全理论知识，还深入了解了各种安全技术和实践。整个学期的学习过程中，贾老师不仅传授了知识，更是激发了我们对网络安全的求知欲和探索精神。本学期的课程不仅让我在学业上有所收获，更让我在人生道路上有了更广阔的视野和更坚实的基础。

同时，我要由衷感谢我的家人，因为他们始终是我最坚实的后盾。在我忙碌于学习的时候，是他们默默地支持和关心着我，让我能够安心地专注于大学学习。他们的无私关爱和期待，成为我不断前进的动力源泉。家人的支持和理解让我有信心面对挑战，勇敢追求梦想。我深知家人的付出和支持是我成长道路上最宝贵的财富，我会倍加珍惜，并以优异的成绩回报他们的期待和关爱。

此外，我还要感谢我的同学们和朋友们。在论文撰写过程中，我们互相鼓励、互相帮助，共同度过了许多难忘的时光。他们的陪伴和支持，让我在学术研究中感受到了温暖和力量。

我还要感谢学校的图书馆，为我提供了丰富的学术资源和良好的研究环境。让我能够查阅到大量的文献资料，使我的研究工作得以顺利进行。

最后，我要再次向所有给予我帮助和支持的人们表达我最深的感激之情。在大学学习的过程中，每一步都离不开你们的鼓励、指导和建议。你们的无私付出和热情帮助，使我在面对困难时不再感到孤单和无助，而是充满了信心和勇气。

参 考 文 献

[1] 黄少青. RC4 算法的安全性分析[硕士学位论文]. 北京邮电大学,北京, 2009 年

[2] 陈虹.刘雨朦.肖成龙.郭鹏飞.肖振久.基于椭圆曲线的改进RC4 算法. 辽宁 葫芦岛: 辽宁工程技术大学 软件学院, 计算机应用: 1001-9081(2019) 08-2339-07, 2019 年

[3] Golic. Linear statistical weakness of alleged RC4 keystream generator. In EUROCRYPT:Advances in Cryptology:Proceedings of EUROPCRYPT 1997

[4] Fluhrer, McGrew. Statistical analysis of alleged RC4 keystream generator. In FSE:Fast Software Encryption, 2000

[5] AIfardan N J, Bernstein D J, Paterson K G, et al. On the security of RC4 in TLS. In:Proceedings of the 22nd USENIX conference on Security. 2013,305-320

[6] 刘聪. 基于密钥流的RC4 算法安全性分析与改进[硕士学位论文]. 湖南大学,湖南, 2016 年

[7] 孟毛广. RC4 流算法的研究与改进[硕士学位论文]. 合肥工业大学, 安徽合肥, 2014 年

[8] Mantin I. Predicting and distinguishing attacks on RC4 keystream generator[M]//Advances in Cryptology-Eurocrypt 2005. Springer Berlin Heidelberg, 2005:491-506.

[9] Fluhrer, S. Mantin, I. &Shamir, A.(2001). Weaknesses in the Key Scheduling Algorithm of RC4. Selected Areas in Cryptography.

[10] Klein, A. (2005). Attacks on the RC4 stream cipher. Designs, Codes and Cryptography, 48(3), 269-286.

[11] 胡亮.迟令.袁巍.李宏图.初剑锋.RC4 算法的密码分析与改进. 吉林长春: 吉林大学 计算机科学与技术学院, 吉林大学学报（理学版）: 1671-5489(2012) 03-0511-06

[12] Paul S. Preneel N. A new weakness in the RC4 key-stream generation: an approach to improve the security of the cipher. In: Proceedings of Fast Software Encryption (FSE), LNCS. Heidelberg: Springer. 2008

[13] Maitra S, Paul G. Analysis of RC4 and proposal of additional layers for better security margin[M]//Progress in Cryptology-INDOCRYPT 2008. Springer Berlin Heidelberg, 2008:27-39