密钥繁多难记难管理？认识高效密钥管理体系

原创 严强 微众银行区块链 4月22日

来自专辑

WeDPR隐私保护周三见

严强，SMU信息安全方向博士，信息安全顶级国际学术会议最佳论文奖获得者；曾作为Google隐私保护基础技术架构部门唯一来自中国的早期核心成员，领导研发的技术方案在Android和Google Play生态各大门户产品中全面集成投产。

密钥设置是否只要够安全就能够重复使用？定期修改密钥到底有没有必要？密钥不幸遗失

该如何恢复？素未谋面的双方，如何才能安全地进行密钥协商？

上一论我们了解到，基于密码学的隐私保护方案，其有效性很大程度上取决于能否有效管理好密钥。这里，我们将进一步分析密钥管理的具体范畴、每个操作环节的典型风险，以及应对手段。

密钥管理的对象是密钥本身或者用于生成密钥的密钥材料（常称之为**根密钥**），三类主要操作环节包括密钥的使用、保存和协商。

鉴于人类用户（以下简称用户）和计算机系统在自身能力上的差异，需要使用不同技术手段和治理手段来实现有效的密钥管理，以下将对三类操作环节一一展开分析。

密钥的使用

密钥的使用是指，用户基于根密钥，为不同业务操作生成实际使用的密钥的过程。这一过程不仅仅包括，直接使用预先设定好的密钥，如输入用户记忆中的用户口令，还包括使用经过一定变换后的密钥。

其主要风险是密钥泄露导致的非授权使用，可能会造成以下后果：

- 由该密钥加密的隐私数据泄露。特别是当所有历史隐私数据都只用一个密钥加密时，攻击者将有可能获得所有历史隐私数据明文。
- 使用该密钥通过身份验证入侵系统。不只是数据，攻击者可以获得用户所有的操作特权，例如，恶意修改系统访问控制参数、使用具有法律效应的数字证书对未授权的内容进行数字签名等。

针对这些在密钥使用环节的风险，核心的应对手段为**密钥轮转**，即每隔一定时间，生成一个新的密钥。

对于计算机系统，一般可以轻易生成新的随机密钥。新密钥与旧密钥可以没有任何关联，

其有效期限和密钥本身都将保存在高安全级别的存储介质中，以此最小化密钥暴露的风险。

然而，以上方案对于用户而言，可用性不高。

对用户来说，生成一个安全的新密钥，且能够记住和使用它，已经不是一件容易的事。如果再进一步要求用户记忆多个超长、随机、无关联的密钥，那用户很有可能被迫“变通”，使用不安全的手段，例如将密钥全部写在纸上、未受保护的手机APP里。

如何让用户只记忆一个密钥，还能实现有效的密钥轮转，这里可以用到的关键技术是密钥派生函数（Key Derivation Function, KDF）。

KDF具有两个核心功能：

- 将一个短的用户口令延长到一个满足安全密钥长度的密钥。
- 由一个根密钥生成多个满足安全密钥长度的密钥。

典型的KDF，如IETF RFC 2898标准中的PBKDF2函数，可以表达成如下形式：

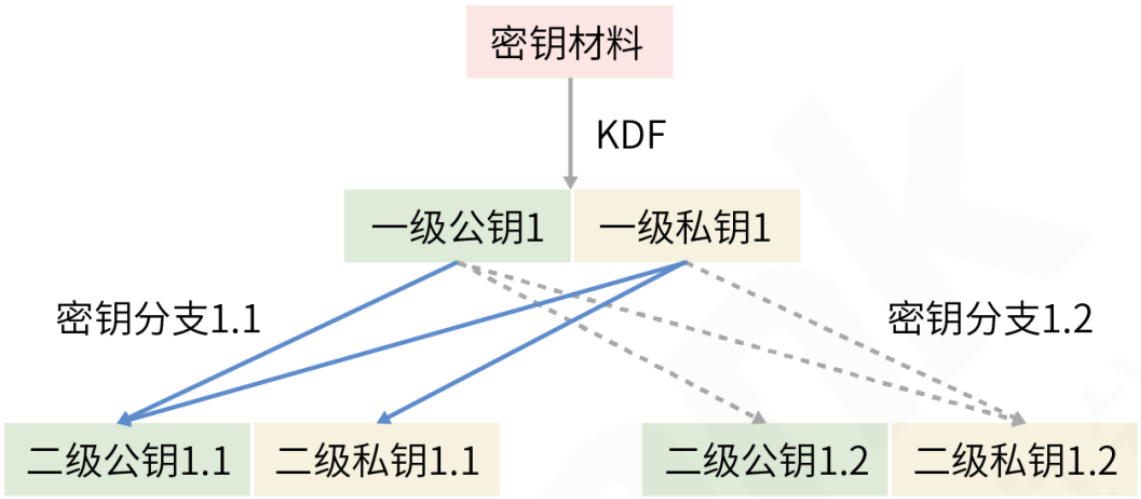
DerivedKey = PBKDF2(PRF, Password, Salt, IterationCount, DerivedKeyLength)

其中：

- PRF是一个随机数生成函数，负责在运算过程中生成一系列随机数。
- Password是用户口令。
- Salt是为了防止批量穷举攻击而设置的盐值，其作用等价于用户专属的随机种子。
- IterationCount是生成派生密钥时所需迭代计算的次数，可以通过刻意增加迭代计算的次数，加大攻击者穷举攻击的难度。
- DerivedKeyLength是派生密钥的长度，一般比用户口令长很多。

PBKDF2函数的五个输入中，用户只需要记忆用户口令Password，其他都可以由辅助的计算机系统来计算完成。用户口令可以根据用户的偏好设置，不影响用户体验。同时只要用户口令够长，安全性风险一般都比较可控。

除了PBKDF2之外，BIP-32标准中Hierarchical Deterministic 密钥钱包设计的核心也是KDF。区别在于BIP-32为椭圆曲线公钥密码学算法提供了特有的密钥派生规则，实现了子密钥树形扩展和中间节点公钥托管扩展，有兴趣的读者可以深入了解一下。



下一级公钥 = KDF1(上一级公钥) = KDF2(上一级私钥)

下一级私钥 = KDF3(上一级私钥)

> 注意：通过**上一级公钥**高效计算得出**下一级私钥**的KDF函数不应该存在。

KDF技术上实现了密钥轮转，在治理上也有必要采取一定的措施，进一步降低密钥使用时的风险。常见治理策略主要覆盖了两大方面的风险：

- 密钥的最短长度和最低复杂性：密钥长度不能太短，不能被常见的字典库轻易破解。
- 密钥的复用：建议定期更改密钥，且不能复用历史密钥。对于计算机系统，可以进一步要求为不同的系统用途设置不同的密钥。

关于第一条治理策略，业界一般都没有什么异议，但对于第二条中，用户需定期更改密钥的建议，近年来也有一些不同观点。

实践中往往发现，为了方便记忆，不少用户采用了不安全的变通方式，选用了有强关联的一组用户口令。例如，2019年使用的旧口令为“password2019”，2020年使用的新口令

为“password2020”，一旦旧口令泄露，也很容易推断出新口令。

反之，如果告知用户不需要定期更改口令，用户在心理上反而更有动力去设置一个更为复杂的口令。所以，实施定期更改用户口令的策略，不一定更安全。

除了以上治理策略，为了控制内部人员滥用密钥的风险，也很有必要将密钥的控制权分派到多个职能上相互约束的相关人员手中，只有当所有相关人员都同意使用时，才能正常使用，其背后的技术原理将在下一环节中提及。

密钥的保存

密钥的保存是指，用户将密钥保存在存储介质中，并在特定的情况下，从存储介质恢复出之前保存的密钥。

其主要的风险是因保存不当导致密钥泄露或遗失，除了上一环节中提到的后果之外，可能会额外造成以下后果：

- 由该密钥加密的隐私数据无法被解密。
- 由该密钥保护的权益无法被兑现。

针对这些在密钥保存环节的风险，核心的应对手段为**物理隔离**和**密钥分片**。

前者指的是，密钥保存的环境应该是一个与恶意环境隔离的安全环境。后者指的是，密钥保存时不应该整存整取，而是进行分片，由多个信任方分别保存，必要时还需要实现多地容灾恢复。

对于计算机系统，安全硬件模块和高物理安全的服务器房间是实现物理隔离常见的手段，必要时，保存密钥的设备可以一直保持离线状态，杜绝意料之外的非授权访问。

对于密钥分片，可以使用密码学秘密分享算法来实现。最常用的密码学秘密分享算法是

Shamir秘密分享算法，由以色列密码学家Adi Shamir在其1979年的论文『How to share a secret』中提出。

Shamir秘密分享算法的核心思想是，将密钥的值设为一个N阶随机多项式中的常量参数，然后在该随机多项式上随机选M个点的坐标，这些坐标就是关于密钥的分片。

这些分片具有以下特性：

- 如果攻击者获得分片总数小于N，攻击者无法获得任意关于密钥的信息。
- 如果所有可能的分片总数M大于N，通过其中任意N个分片，使用拉格朗日多项式插值算法恢复出随机多项式之后，便可有效地恢复出密钥。

$$y = a_n * x^n + a_{n-1} * x^{n-1} + \ldots + a_1 * x^1 + a_0$$

密钥分片过程：

选定一个随机多项式 y，设置a_0= 需要进行分片保护的密钥

选取多项式曲线上N个线性无关的点作为密钥分片(x_1, y_1), (x_2, y_2), …, (x_n, y_n)

密钥恢复过程：

根据多项式性质，获得以上N个点一定能解出所有的多项式参数a_n, a_{n-1}, …, a_0

其中，a_0就是恢复出的密钥

相比计算机系统，用户对于前一项物理隔离的要求可能更容易实现。相比读取存储介质中的数据，在未进行威逼利诱的前提下，用技术手段直接提取用户记忆中的口令目前要困难得多。

但对于第二项密钥分片的要求，则需要配合各类托管技术，使用计算机辅助手段生成和保存高安全性的密钥分片。具体的技术分类和比较，可以参考上一论密钥托管相关内容。

无论采用哪一种技术，一般情况下，用户最少需要记忆一个用户口令。但如同房门钥匙一样，遗忘用户口令并不罕见，尤其是在账户数目和相关口令总数繁多的情况下。

如果服务提供商提供有效的重设服务，相比将用户口令全部写在纸上或手机APP里，通过该服务重设用户口令，泄露风险可能更低。

密钥的协商

密钥的协商是指，多个用户或系统远程协商即将在交互过程中所使用的密钥。

作为社会性生物，和陌生人交换信息，是人类生活中必不可少的组成部分。在当前数据驱动的时代，计算机系统通过跨域交换信息实现更大的价值发掘，是现代信息化商业核心业务模式之一。

在这些信息交换过程中，最典型的应用之一是隐私数据的端到端加密传输，为此需要生成**一次性密钥**对信息进行加密保护。在缺乏可信信道的前提下，能否与交互方安全地完成密钥协商，对于隐私数据的保护尤为关键。

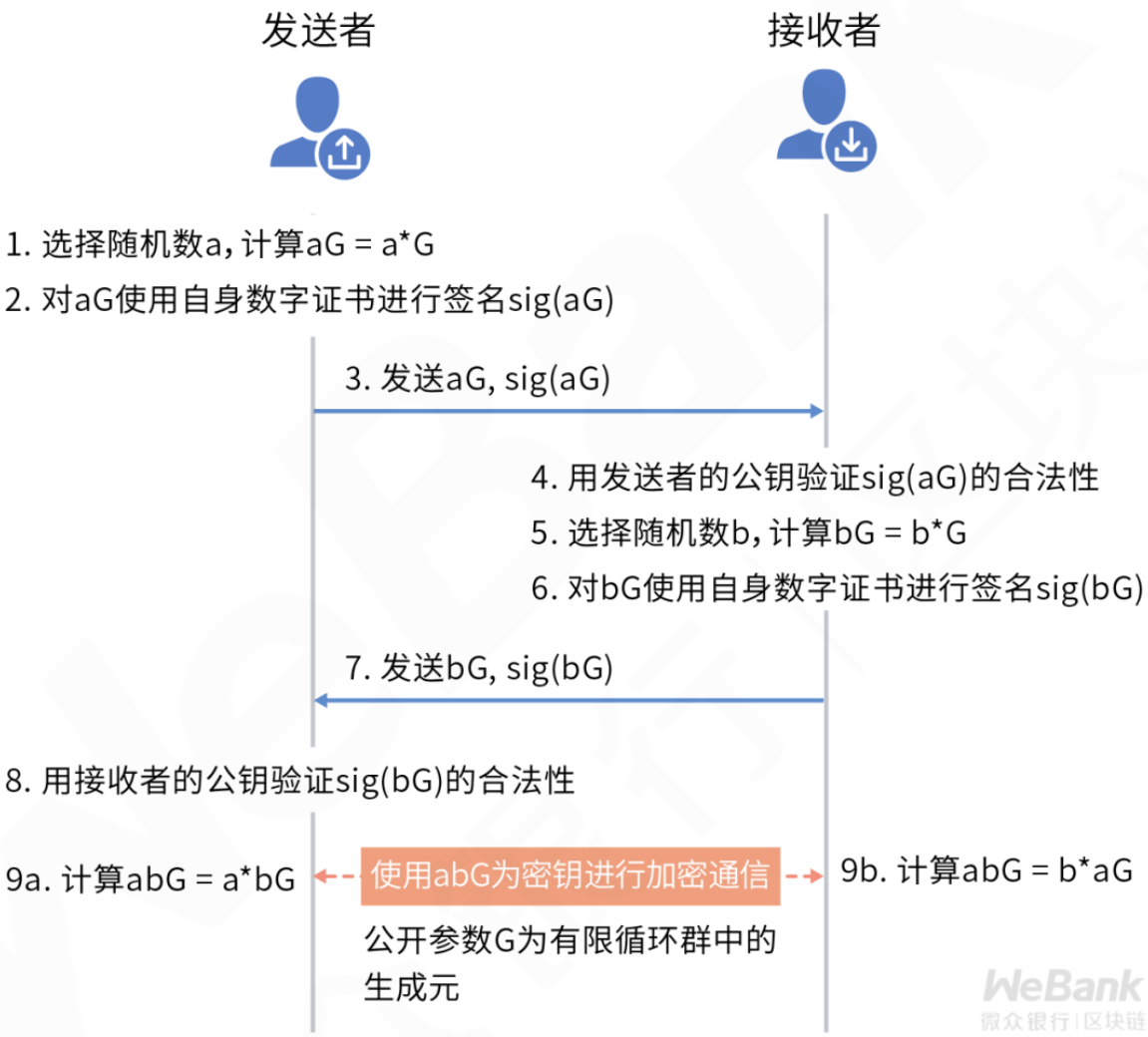
其主要的风险是恶意通信环境中的密钥截取或篡改，除了密钥使用环节中提到的后果之外，可能会额外造成以下后果：

- 由该密钥保护的隐私数据被篡改。例如，金融交易中的收款人、付款金额。

- 由该密钥保护的其他密钥泄露。例如，通过加密信道传输的后续协议中所约定使用的密钥。

针对这些在密钥协商环节的风险，核心的应对手段为**认证交换**和**认证分发**。

对于计算机系统，根据业务场景和部署环境安全假设的不同，认证密钥交换的协议有很多不同的类型，最常用的是基于公钥证书体系和Diffie-Hellman密钥交换协议。

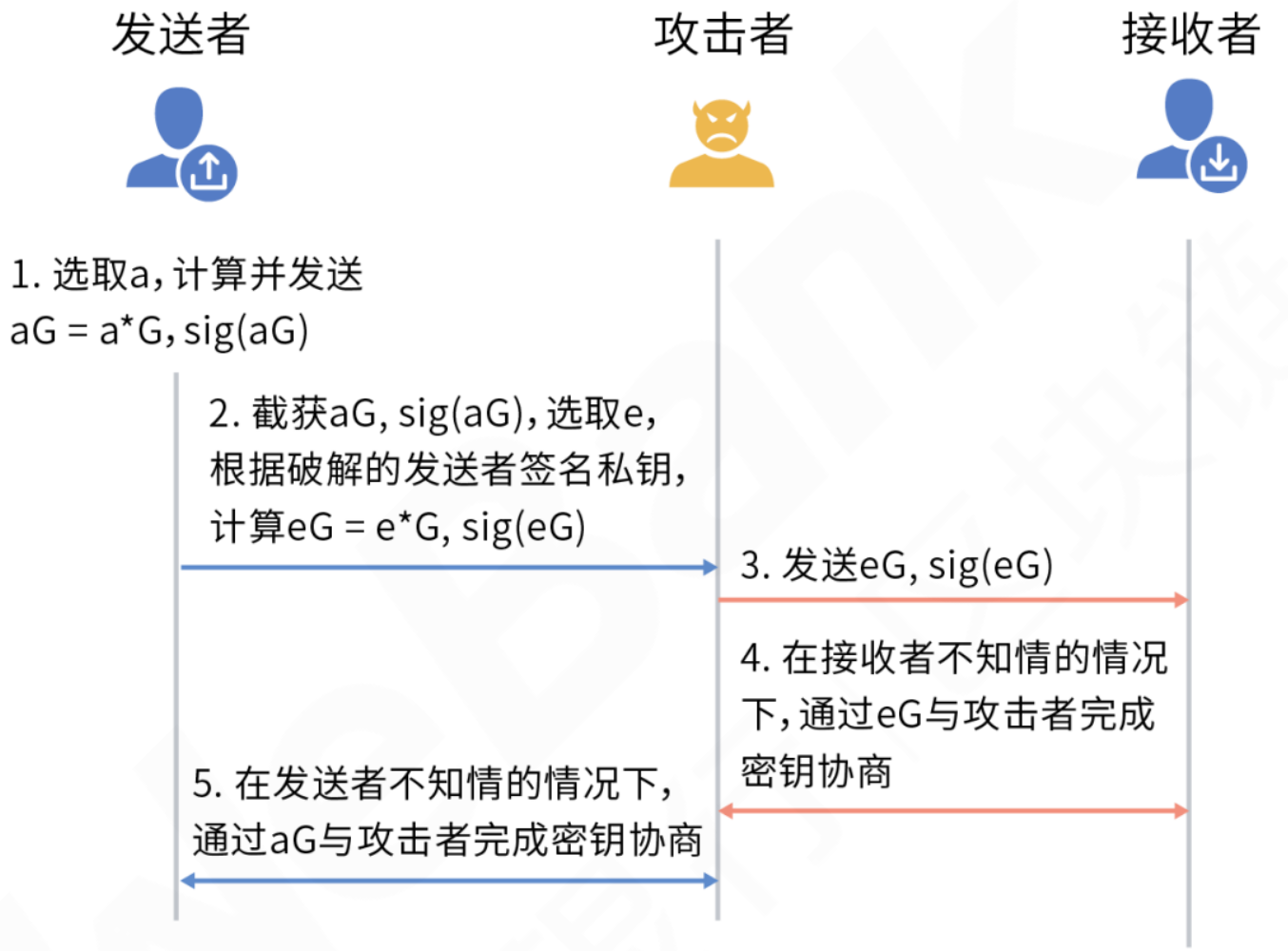


关于认证密钥分发，经典密码学相关方案有基于公钥证书体系的密钥封装（Key Encapsulation Mechanism, KEM）、基于可信中间代理和代理重加密（Proxy Re-Encryption）的密钥密文转换等。比较创新的技术方案包括在第6论中提到的基于量子纠缠理论的量子密钥分发技术，我国的墨子号量子科学实验卫星已经实现了千公里级星地双向量子纠缠分发原型实验。

在技术手段的协助下，用户往往对于密钥协商是无感的。例如，我们在使用浏览器时，通

常不会意识到，对不同的网站建立安全连接时，会根据不同网站的不同数字证书，进行密钥的认证交换，并最终使用不同的一次性密钥与不同的网站通信。

但用户也需要警惕，核实认证的有效性。一旦数字证书失窃或者过期，攻击者就有机会假扮服务提供方，截获用户的隐私数据，篡改用户的操作请求，实施经典的中间人攻击。



一旦任意一方的数字证书失窃，签名私钥被泄露，攻击者就有可乘之机。

密钥管理的三大环节中，存在着大量的风险点。由于密钥是密码学中的最高机密，窃取密钥往往是攻击者的首要目标。任何一个环节出现问题，对应隐私保护方案的整体有效性，都会受到严重影响。

本论的分享主要关注技术方案和治理策略层面（参见以下总结表），在实际方案部署时，

工程实现和其他相关层面的保护也很关键，会需要更完备的组合策略，从多个维度上提供层次化的保障。

	特有风险举例	技术方案	治理策略
密钥使用	• 隐私数据泄露 • 权益滥用	• 密钥轮转	• 密钥复杂性最低要求 • 密钥复用限制 • 操作人员制约机制
密钥保存	• 隐私数据无法解密 • 权益无法兑现	• 物理隔离 • 密钥分片	• 操作人员物理访问限制 • 多地容灾恢复
密钥协商	• 隐私数据篡改 • 密钥泄露	• 认证交换 • 认证分发	• 证书有效性核实

正是：密钥管理谨小慎破解，技术治理合璧显神功！

有效的密钥管理是使用基于密码学的隐私保护方案的重要前提。无论隐私保护方案内部设计多么精妙，任何在密钥使用、保存、协商环节中出现的疏漏，都会使之功亏一篑。

除了传统的安全性分析，针对用户的可用性分析也至关重要。在实际隐私保护应用中，高于常规人类认知记忆能力的要求，都会促使用户使用不安全的变通手段，导致最终效果大打折扣。

有效的密钥管理需要在多个维度上融合技术方案和治理策略，同时实现安全性和可用性之间的平衡和优化。

了解完密钥相关的重要原则和管理技术，自下一论开始，我们将分享在实际的密码学算法中如何使用这些密钥，深入解析隐私保护相关的密码学原语，欲知详情，敬请关注下文分解。

---END---

《隐私保护周三见》

“科技聚焦人性，隐私回归属主”，这是微众银行区块链团队推出《隐私保护周三见》深度栏目的愿景与初衷。每周三晚8点，专家团队将透过栏目和各位一起探寻隐私保护的发展之道。

栏目内容含括以下五大模块：关键概念、法律法规、理论基础、技术剖析和案例分享，如您有好的建议或者想学习的内容，欢迎随时提出。

栏目支持单位：零壹财经、陀螺财经、巴比特、火讯财经、火星财经、价值在线、链客社区

往期集锦

第1论 | 隐私和效用不可兼得？隐私保护开辟商业新境地

第2论 | 隐私合规风险知几何？数据合规商用需过九重关

第3论 | 密码学技术何以为信？深究背后的计算困难性理论

第4论 | 密码学技术如何选型？初探理论能力边界的安全模型

第5论 | 密码学技术如何选型？再探工程能力边界的安全模型

第6论 | 密码学技术如何选型？终探量子计算通信的安全模型

第7论 | 密码密钥傻傻分不清？认识密码学中的最高机密

上下滑动查看更多

长按二维码关注

微众银行区块链

白皮书下载 | 订阅干货 | 进群交流 | 合作联系

文章已于2020-04-22修改