密码学原语如何应用？解析密码学承诺的妙用

原创 廖飞强 微众银行区块链 5月20日

来自专辑

WeDPR隐私保护周三见

在不泄露明文的前提下，如何对隐私数据的内容进行承诺？密码学承诺的密文形式和普通的数据密文有何区别？隐私数据如何在密码学承诺的形式下依旧保持可用性？在量子计算的安全模型下，是否依旧可以构造安全可用的密码学承诺？

隐私保护方案设计中，除了保护隐私数据的机密性，确保密文形式隐私数据解读的唯一性也是重要的业务需求。业务流程中，很大程度会依赖隐私数据的具体数值，如果允许攻击者在自身利益驱动下，对处于密文形式的隐私数据进行任意解读，势必会对业务的整体公正性和有效性带来巨大影响。

以电子支付为例，一家银行为一位客户开具了一张面额1000元的电子支票，电子支票以密文形式交付给客户，流转过程中不会轻易泄露金额。然而，在使用时，银行也不希望客户能够

将这张电子支票解读成其他金额，如10000元。多兑现的9000元会造成银行的损失，银行甚至可能因此而停用整个密文电子支票业务。

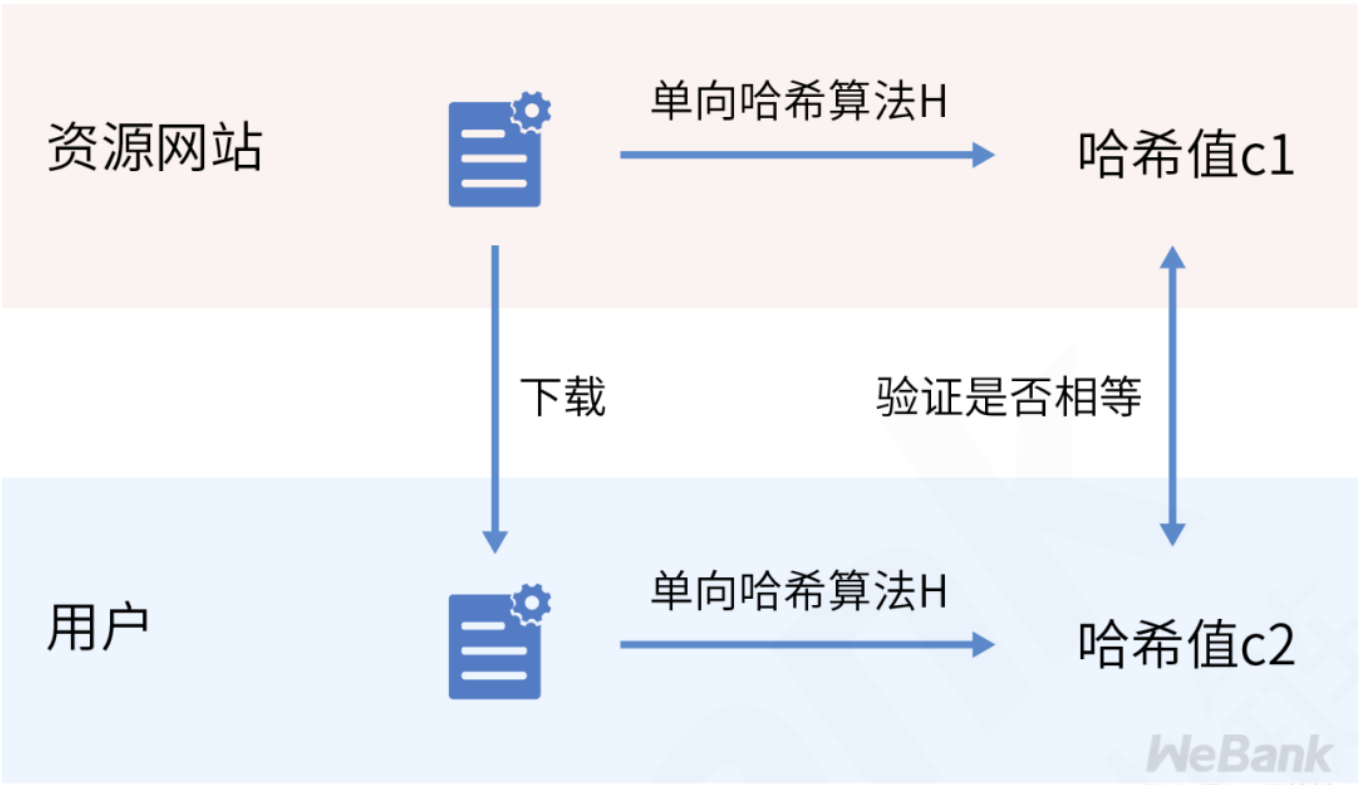
这里的**解读**与**解密**有一定区别，对密文数据解读不一定需要对密文数据进行解密。在上面示例中，当客户花费这张面额1000元的电子支票，解读时只需要证明电子支票的消费额小于未花费余额即可，而不需要解密未花费余额的具体数值。

解决以上业务问题的关键，就在于密码学承诺的使用。密码学承诺有何神奇之处？且随本文一探究竟。

哈希承诺

在日常生活中，承诺无处不在。例如，预约打车成功后，司机和乘客之间就互相做了一个承诺。到了预约时间，乘客等车，司机接客，这就是在兑现承诺。

信息科学中也有类似的承诺技术存在。例如，某些网站在提供下载文件时，也会提供对应文件的单向哈希值（单向哈希算法相关内容可以参考第9论）。这里，单向哈希值便是一种对文件数据的承诺，以下称之为**哈希承诺**。基于下载的哈希承诺，用户可以对下载文件数据进行校验，检测接收到的文件数据是否有丢失或变化，如果校验通过，相当于网站兑现了关于文件数据完整性的承诺。



密码学承诺是一类重要的密码学原语，其中哈希承诺又是诸多技术中最简单的一种实现方式。

一般而言，密码学承诺的应用涉及**承诺方**、**验证方**两个参与方，以及以下两个使用阶段。

第一阶段为**承诺生成**（Commit）阶段，承诺方选择一个敏感数据v，计算出对应的承诺c，然后将承诺c发送给验证方。通过承诺c，验证方确定承诺方对于还未解密的敏感数据v只能有唯一的解读方式，无法违约。

第二阶段为**承诺披露**（Reveal）阶段，学术界通常也称之为**承诺打开-验证**（Open-Verify）阶段。承诺方公布敏感数据v的明文和其他的相关参数，验证方重复承诺生成的计算过程，比较新生成的承诺与之前接收到的承诺c是否一致，一致则表示验证成功，否则失败。

一个设计良好的密码学承诺具备如下特性：

- **隐匿性**：在打开关于v的承诺c之前，验证方不知道承诺方选择的敏感数据v。
- **绑定性**：在关于v的承诺c生成之后，承诺方难以将已承诺的敏感数据解释成另一个不同的数据v'。

所以，密码学承诺可以起到与日常生活中的承诺行为类似的效果，一旦做出承诺，就必须在披露阶段使用之前已经承诺的敏感数据。

对应地，在业务系统中，承诺生成阶段通常被用来生成密文形式的业务数据，而承诺披露阶段则多被用于在特定业务流程中进行数据校验。

除了直接公布敏感数据明文之外，承诺披露阶段所需的数据校验，也可以在不公布敏感数据明文的前提下，构造零知识证明来完成。相关内容将在后续零知识证明专题中展开。

具体回到哈希承诺，用户可以通过以下公式计算关于敏感数据v的承诺，其中H是一个密码学安全的单向哈希算法。

$$c = H(v)$$

基于单向哈希的单向性，难以通过哈希值H(v)反推出敏感数据v，以此提供了一定的隐匿性；基于单向哈希的抗碰撞性，难以找到不同的敏感数据v'产生相同的哈希值H(v)，以此提供了一定的绑定性。

哈希承诺的构造简单、使用方便，满足密码学承诺基本的特性，适用于对隐私数据机密性要求不高的应用场景。

对隐私数据机密性要求高的应用，需要注意哈希承诺提供的隐匿性比较有限，不具备随机性。对于同一个敏感数据v，H(v)值总是固定的，因此可以通过暴力穷举，列举所有可能的v值，来反推出H(v)中实际承诺的v。

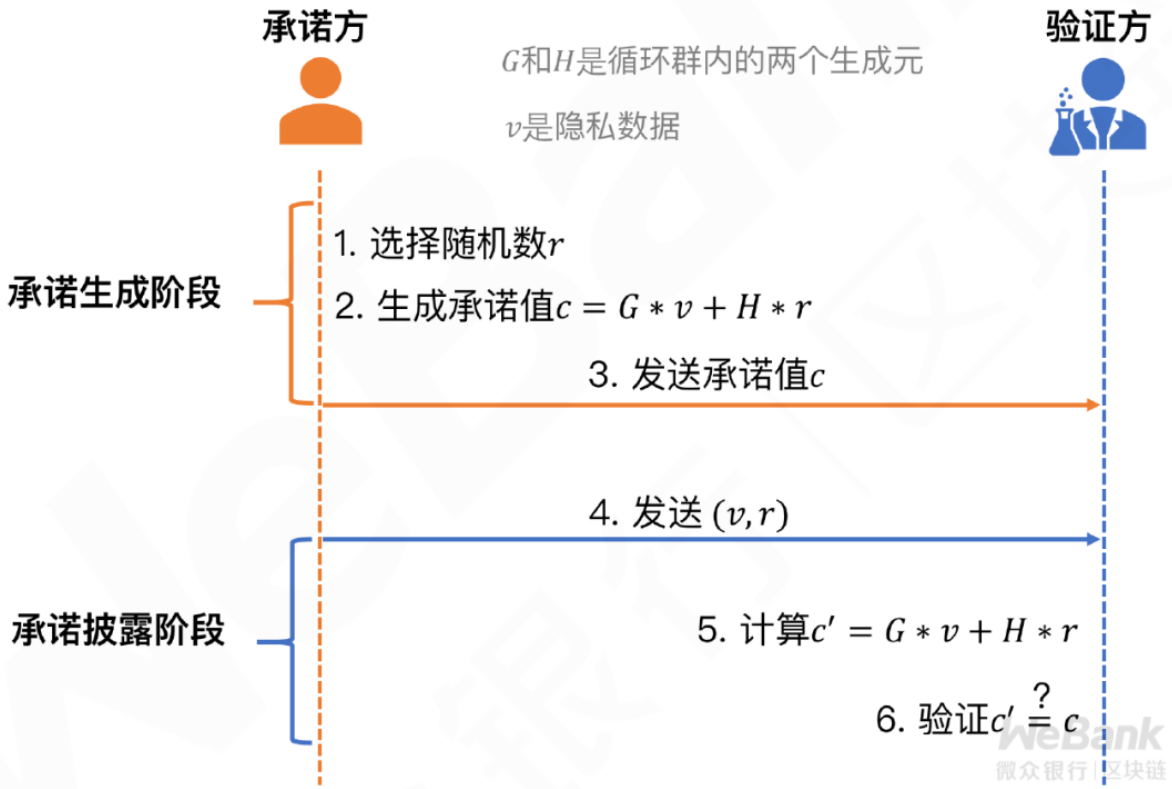
相比其他密码学承诺技术，哈希承诺不具有便于业务系统在密文形式对其处理的附加功能，例如，多个相关的承诺值之间密文运算和交叉验证，对于构造复杂密码学协议和安全多方计算方案的作用比较有限。

Pedersen承诺

Pedersen承诺是目前隐私保护方案中使用广泛的密码学承诺，相比哈希承诺，构造略微复杂，但提供了一系列优异的特性：

- 信息论安全（参见第4论）的理论最强隐匿性。
- 基于离散对数困难问题（参见第3论）的强绑定性。
- 具有同态加法特性的密文形式。

其具体构造如下：



有别于哈希承诺，对于同一个v会产生相同的承诺H(v)，Pedersen承诺通过引入随机致盲因子r，即便隐私数据v不变，最终的承诺c也会随着r的变化而变化，以此提供了信息论安全的隐匿性。

Pedersen承诺在构造中采用了离散对数运算，因此也赋予其加法同态性。可以通过两个分别关于v1和v2的Pedersen承诺c1和c2“相加”，得到的新承诺便是关于v1+v2的Pedersen承诺。

- Pedersen承诺的算法为C，r_1 和 r_2 为独立的随机数，计算两个承诺值：

  关于$v1$的Pedersen承诺 $c_1 = C(v_1, r_1) = G * v_1 + H * r_1$

  关于$v2$的Pedersen承诺 $c_2 = C(v_2, r_2) = G * v_2 + H * r_2$

- 根据加法同态性，关于$v_1 + v_2$的Pedersen承诺可以通过将关于 v_1 和 v_2 的Pedersen承诺相加获得：

$$
\begin{aligned}
c_3 &= c_1 + c_2 \\
&= (G * v_1 + H * r_1) + (G * v_2 + H * r_2) \\
&= G * (v_1 + v_2) + H * (r_1 + r_2) \\
&= C(v_1 + v_2, r_1 + r_2)
\end{aligned}
$$

WeBank 微众银行|区块链

除了能够构造关于v1+v2的Pedersen承诺之外，Pedersen承诺还可以用来构造v1*v2、v1 || v2等更复杂的Pedersen承诺，通过基于离散对数的通用零知识证明系统，来证明新产生的承诺满足与原始承诺c1和c2之间存在指定的约束关系。

在实际业务中，Pedersen承诺自带的加法同态性，配合零知识证明获得约束关系证明功能，在区块链中可以有广泛的应用，目前主要以隐匿账本的形式，提供灵活的隐私数据的密文上链存证和交易密文数值关联性的第三方验证。

具体方案设计中，相关业务方在链下完成业务交互之后，将对应的数值变化表达成Pedersen承诺，再将对应的承诺数据上链，这个过程中无需披露任何隐私数据明文。

上链之后，非相关的第三方虽然难以通过Pedersen承诺的密文形式反推出隐私数据明文，但可以验证承诺之间的约束关系，核实业务交互的合法性，例如，验证隐匿转账发生之后，依旧满足会计平衡、外汇交易中使用了正确汇率进行跨行对账等。

值得注意的是，Pedersen承诺产生的密文形式，与通过普通加解密算法生成的数据密文有一定相似性，在计算过程中都使用敏感数据v，致盲因子r的作用和密钥的作用也有一些相似，均用以混淆最后的密文输出。

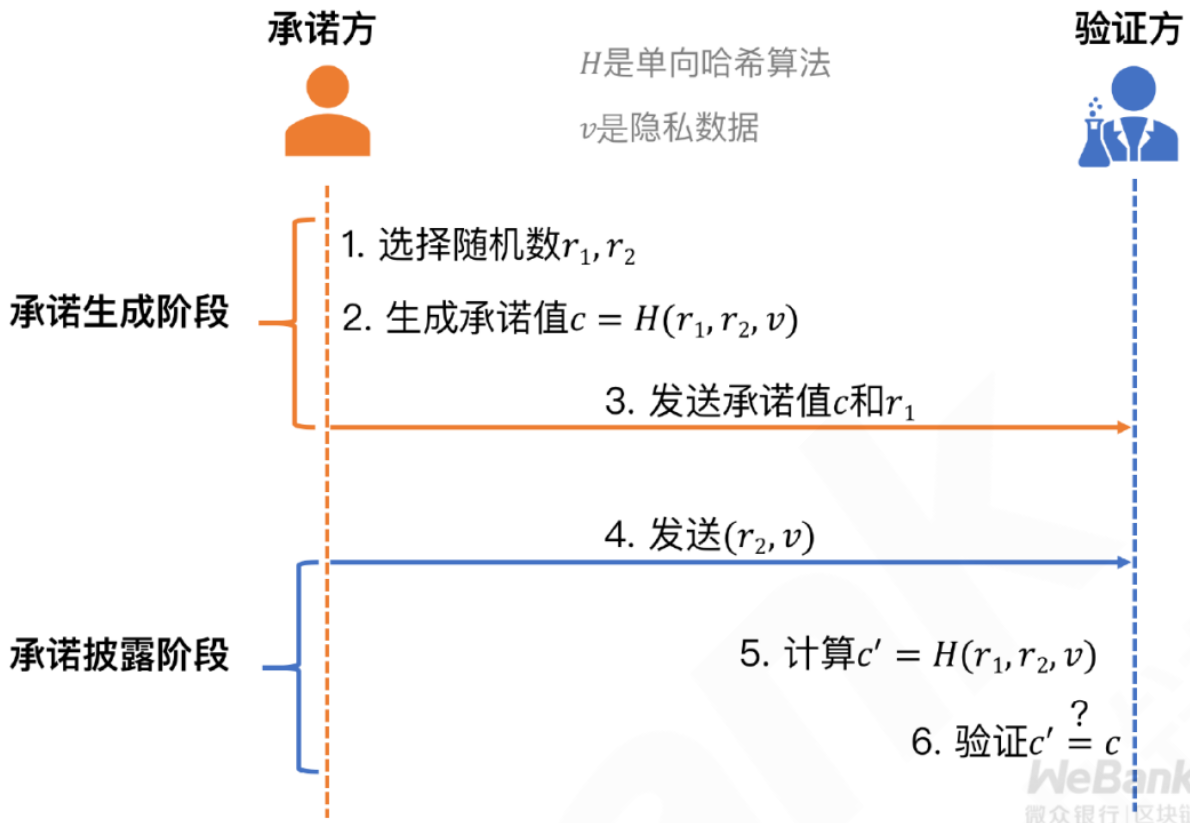
但不同的是，密码学承诺不提供解密算法，如果只有r，无法有效地提取出敏感数据v的明文，只能通过暴力穷举所有可能的v值的方法逐一验证，试图通过匹配的承诺值来破解v的明文。

所以，Pedersen承诺重在“承诺”，适用于数据属主向第三方证明承诺中的敏感数据满足一定的约束关系，由于不直接提供解密功能，不能直接支持需要互不透露敏感数据明文的多方协同计算，这一点与密码学领域的同态加解密算法有很大区别，切勿混淆概念。

量子承诺

为了应对量子计算可能带来的风险，寻求经典密码学承诺技术的替代品，后量子密码学承诺也是重要的研究方向之一。比较典型的方案有量子比特承诺。

量子比特承诺(Quantum Bit Commitment)是基于量子力学原理构造的比特承诺方案，具体实现可以抽象为一个带随机输入的单向哈希算法。

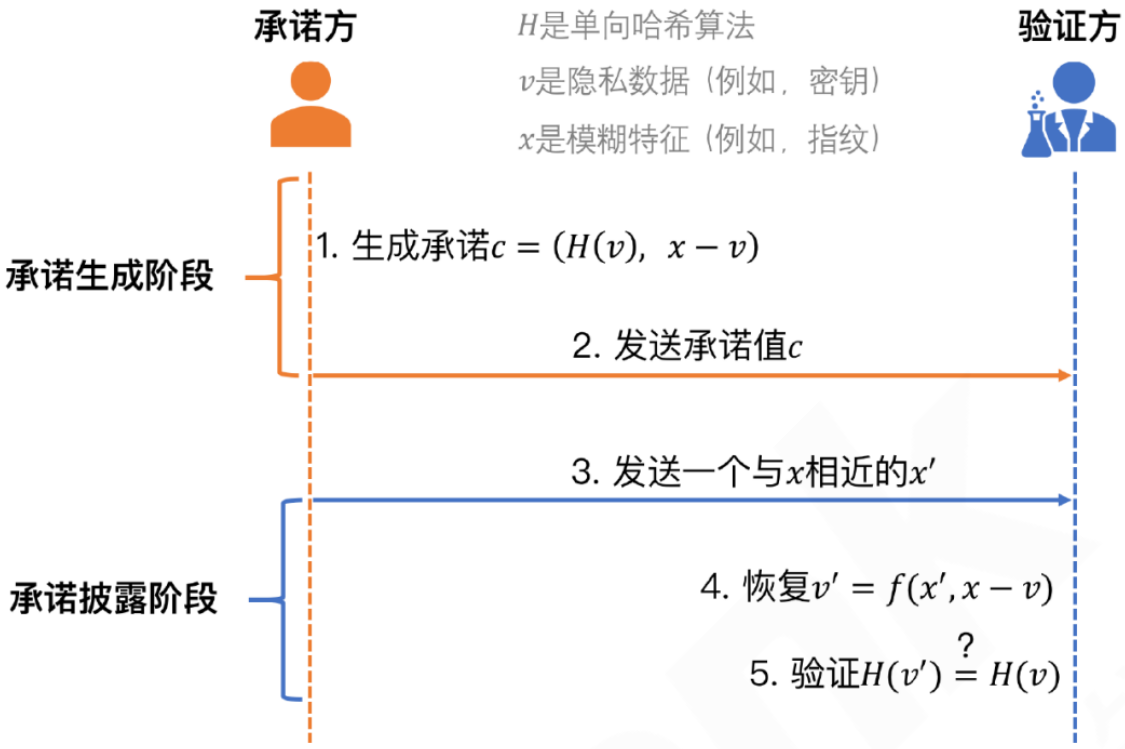


根据单向函数的单向性，承诺方向验证方发送r1和c后，验证方不知道v，满足对v的隐匿性。另外，由单向哈希的抗碰撞性可知承诺方难以找到r2′和v′，使H(r1，r2，v) = H (r1，r2′，v′)，因此承诺方难以违约，满足对v的绑定性。

量子比特承诺的构造看似简单，但实际实现需要借助量子协议完成计算，同时也有一定的理论局限性。

早在1996年，Hoi-Kwong Lo和Hoi Fung Chau团队、Dominic Mayers团队分别独立地证明了**不存在满足信息论安全的理论最强绑定性的量子比特承诺方案**。这个不存在性被称为**MLC no-go定理**。其主要原因是，如果验证方完全没有任何承诺的信息，那么承诺方可以通过量子纠缠随意地改变承诺内容，而验证方既不能阻止也不能发现承诺方的违约行为。

总体而言，后量子密码学承诺的研究尚处于早期阶段，充满了各类挑战，目前难以直接应用到实际业务系统中。除了量子比特承诺之外，基于模糊算法的量子模糊承诺也是一类热门研究方向，目标应用领域为生物特征识别相关的隐私安全系统。将来不排除有更实用的方案面世，以此消解量子计算可能带来的冲击，我们将持续保持关注。



f是一个模糊数据恢复算法，当输入x'值接近x值时，其输出$v' = v$

正是：业务数据精确至毫厘，密码承诺隐匿遁无形！

密码学承诺的隐匿性和绑定性是隐私保护方案设计中常用的关键特性，在保障隐私数据机密性的同时，也保证了密文形式隐私数据解读的唯一性。对于业务系统设计而言，密码学承诺为隐私数据提供了另一种高效的密文表达方式。

本论中，我们重点介绍了哈希承诺和Pedersen承诺，在往后的文章中，我们还会进一步介绍其他重要的密码学承诺，例如zk-SNARKs零知识证明系统中使用的多项式承诺、向量承诺等。

对于需要在数据的密文形式上直接进行运算和交叉验证的业务，只要不涉及互不透露数据明文的多方协同计算，相比现有同态加密算法，以Pedersen承诺为代表的密码学承诺往往可以提供更好的性能。这一优势与密码学承诺的同态性密不可分，如何构造和应用同态性，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

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