PAI 数据

Project PAI数据协议摘要

作者：
杜锦程 [1]
方丹博士 [2]
Mark Harvilla博士 [3]

摘要

Project PAI数据协议（“PAI数据”） 是对Project PAI区块链协议扩展的一个规范，其中包括一个对任意数据进行保护和提供访问的方法。在PAI币发展提案（PDP）2 [1]的背景下，本文定义了PAI数据支持的两种重要交易类型：存储交易（即便于数据存储和所有权证明的交易）和共享交易（即旨在实现授予和撤销指定接收人对数据访问权限的交易）。本文同时还提供了一个PAI数据协议与类似基于区块链的文件存储系统之间的比较分析。

[1] ObEN公司区块链研究员
[2] ObEN公司区块链研究员
[3] ObEN公司首席工程师

介绍

Project PAI数据协议（“PAI数据协议”） 是对扩展Project PAI区块链协议扩展的一个规范，其中包括一个保护和提供对任意数据进行保护和提供访问的方法。这是通过利用一个定制的OP_RETURN数据协议和椭圆曲线加密（elliptical encryption），并结合数据存储（例如，用于去中心化存储的由种子节点组成的分段式网络或用于中心化存储的专有数据存储）来实现的 [1]。

PAI数据背景

PAI数据协议是Project PAI为建立一个可持续的人文信息经济而努力的核心 [2]。有PAI数据，用户就能使用Project PAI协议来保护他们的数据并对此类数据进行授权。PAI Pass是一个目前正在开发的应用程序，它针对利用PAI数据的潜在用例提供了说明。PAI Pass是一个集成的单点登录服务和数据管理应用程序。PAI Pass平台让用户能使用一个独立于PAI Pass平台且用户认准的第三方数据存储协议（例如PAI数据）来输入和保护他们的数据（例如，姓名和年龄等）。

其他用例

虽然不能替代保护知识产权的传统方法（例如版权登记），但PAI数据协议对于希望在保留作者身份证据的同时将其内容商业化的知识产权开发者来说是极其有用的。如下所述，PAI数据提供了一种简单的方法来存储具有不可变时间戳的数据并能获得第三方的验证。因此，用PAI数据协议来保护专有数据（例如，商业机密和/或其他知识产权）以及此类数据的作者或所有者的身份，将产生强有力的证据，证明在时间戳中反映的时间或之前：（1）一件作品已经诞生；而且（2）确认了身份作者已被创建或至少知道此类数据。如果能通过一个认证系统（例如，PAI Pass）获得强化，则该功能将特别有用。

技术说明

PAI数据协议的主要构成有:

- Project PAI区块链
- OP_RETURN操作码
- 椭圆曲线加密(Elliptical Encryption)
- （数据）发送者 +（存储） 供应者 + （数据）接收者
- 数据存储（例如，由种子节点组成的分段式网络）

交易类型

存储交易

存储交易用于通过PAI区块链存储数据，并且如果所述交易包括关于声明所有权的一方的信息或者与诸如PAI Pass之类的平台一起使用的信息，则可以用于对存储的数据的所有权进行声明。存储交易是通过向自己广播一笔交易来创建的，其中OP_RETURN输出包含了一个数据的哈希标识符——可通过存储供应者进行检索。或者，例如，OP_RETURN可以包含数据本身的一个哈希或一些其他的标识符，从而将数据映射到专有数据存储中。

共享交易

共享交易用于通过PAI区块链在发送者和接收者之间共享数据。共享的交易是通过向预期接收者广播一笔交易来创建的，其中OP_RETURN输出包含了一个数据的哈希标识符——可通过存储供应者进行检索。

使用案例

本节概述了如何根据您的需要使用PAI币协议。

如果您需要...	***那么您应该使用...***
存储和/或声明数据的所有权（例如，数字资产、知识产权）	PAI存储交易
授予接收人数据权限	PAI共享交易（依据 PDP 2 [1]，带有一个“授予”操作ID （“grant” Operation ID））
撤消收件人使用数据的权限	PAI共享交易（依据 PDP 2 [1]，带有一个“撤销”操作ID （“revoke” Operation ID））

例子: 存储交易

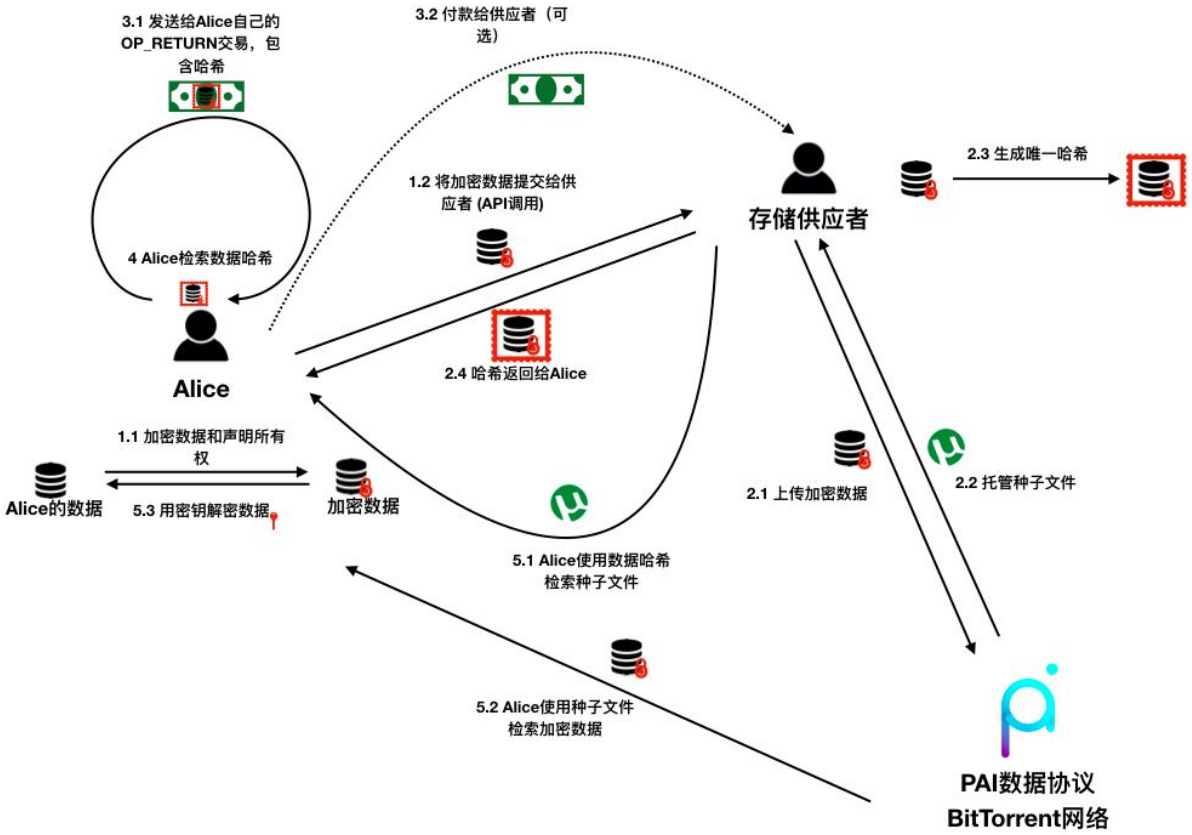

图1. 存储交易流程

如图1所示，一方（例如，Alice）可以通过PAI数据交易存储数据并声明对所存储数据的所有权。方法如下：

1. Alice用她的公钥加密数据，包括对所有权或作者身份的声明。此加密数据只能由拥有相应私钥的Alice进行解密。
2. Alice选择存储供应者，并且通过协议指定的一系列标准API调用（即“供应者API”），将加密数据提交给供应者。存储供应者：
    a. 将数据上传到PAI数据BitTorrent网络。
    b. 托管相应的种子文件。
    c. 生成与数据关联的唯一哈希。
    d. 哈希返回给Alice。
3. Alice创建一个PAI数据协议交易，有两个输出：
    a. 发送给Alice自己的OP_RETURN交易：此输出包含OP_RETURN操作码字段中加密数据的哈希。
    b. 付款给供应者（可选）：此输出根据各方之间的离线协议补偿存储供应者托管种子文件的服务。
4. Alice接收OP_RETURN交易并检索数据哈希

5. 如果Alice希望提出所有权声明或以其他方式检索数据，（1）使用数据哈希从存储供应者处检索种子文件；（2）从PAI数据BitTorrent网络处下载数据；（3）使用她的私钥解密数据。

例子: 共享交易

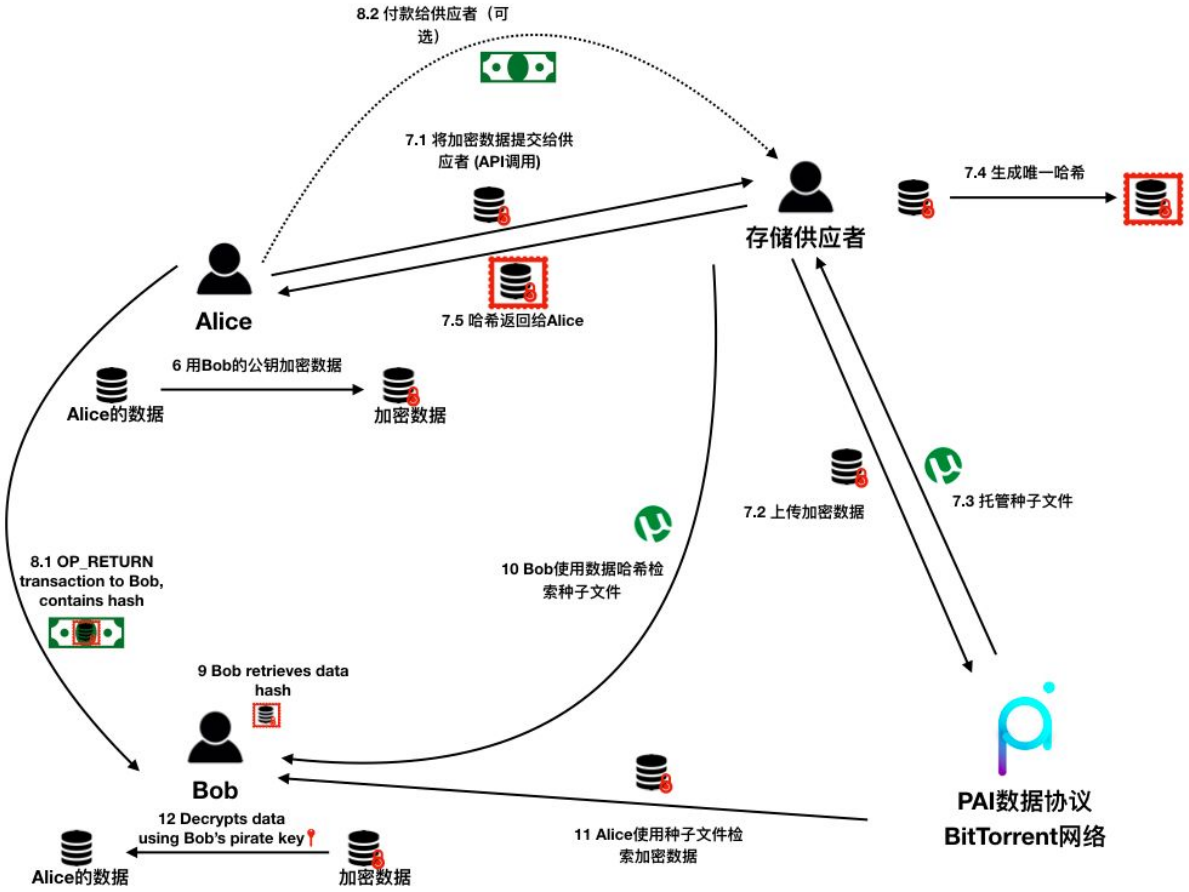

图2. 共享交易流程

如图2所示，双方（例如数据发送者Alice和数据接收者Bob）可以以下文的方式通过PAI数据协议交易来互相分享数据。

6. Alice用Bob的公钥加密数据。此加密数据只能由拥有相应私钥的Bob进行解密。
7. Alice选择存储供应者，并且通过协议指定的一系列标准API调用（即“供应者API”），将加密数据提交给供应者。存储供应者：
    a. 将数据上传到PAI数据BitTorrent网络。
    b. 托管相应的种子文件。
    c. 生成与数据关联的唯一哈希。
    d. 哈希返回给Alice。
8. Alice创建一个PAI数据协议交易，有两个输出：
    a. 发送给Bob的OP_RETURN交易：此输出包含OP_RETURN操作码字段中加密数据的哈希。

    b. 付款给供应者（可选）：此输出根据各方之间的离线协议补偿存储供应者托管种子文件的服务。
9. Bob接收OP_RETURN交易并检索数据哈希。
10. Bob利用数据哈希从存储供应者处取得种子文件。
11. Bob从PAI数据协议的BitTorrent网络处下载数据。
12. Bob用自己的私钥解密数据。

比较 — 其他基于区块链的文件存储系统

基于区块链的文件存储系统主要有两类 [3]：

1. 存储证明 / 容量证明（如 Permacoin, Spacemint）：单个节点没有足够空间存储完整数据集时，通过证明部分子集的持有真实性与完整性来生成新的区块
2. 维持一条区块链用以充当存储提供者与存储购买者之间的中间代理人 （例如 Storj, Filecoin （IPFS））

这些基于区块链的文件存储解决方案普遍修改了协议的共识机制；而PAI币并没有。例如，它们用到了以下共识机制：

- 存储证明（如 Permacoin, Spacemint, Burstcoin, Chia）：证明验证者保留了一定量的存储空间。
- 复制证明（如 Filecoin/IPFS [4]）：证明验证者在唯一的物理存储设备上复制了部分数据。
- 时空证明（Filecoin/IPFS）：证明部分数据被存储了一段时间。

PAI数据协议的优势

存储供应者进入门槛较低

在大多数其他基于区块链的数据存储解决方案中，一个节点供应存储的动机是区块奖励。节点可以供应的存储或带宽越多，它挖到下一个区块并获得区块奖励的概率就越高。这会导致节点之间的存储/带宽竞争，就像工作量证明（Proof of Work）中的哈希算力竞争一样，并且可能逐渐淘汰只具有小存储容量的节点。

相比之下，在PAI数据协议中，存储供应者通过数据提交者的交易被支付（可选）。如果与数据提交者达成了脱链协议，那么没有大量可用磁盘空间的单个存储供应者仍可以直接通过中继数据受益。这有助于降低存储供应者的进入门槛，从而鼓励更多的网络参与。无论数据本身是否继续由供应者提供，作为数据保管链的不可变记录，授予和撤销操作是永久嵌入区块链中的。

存储支付更灵活

对于存储供应者的可选支付为分发费用提供了更大的灵活性。例如，供应者可以根据其可访问的空间，带宽，冗余级别和保留策略向存储供应者免费支付不同费用。或者，供应者可以基于中继的数据量向存储供应者付费。如果数据传输速度很慢，从接收者的角度来看，提交者可能会提高他们愿意支付的费用，从而鼓励更多的供应者参与并提高速度。

数据传输效率更高

一旦数据接收者取得了数据，基于数据发送者与数据接收者之间的脱链协议，数据发送者可以通过撤销交易要求存储提供者（可能有多个）删除他们的本地备份。这可以极大地提高系统效率，因为存储供应者们可以复用他们有限的存储空间以支持不同的数据传输任务。请注意，在任何情况下数据都是用接收者的公钥加密的，因此对存储供应者来说是不可读的。

参考文献

[1] Alex Waters, Mark Harvilla, Patrick Gerzanics, PDP 2, “PAI Data Storage and Sharing,” 2018, https://github.com/projectpai/pdps/blob/master/pdp-0002.mediawiki
[2] Project PAI, “Project PAI Technical Whitepaper,” https://projectpai.com/assets/files/whitepaper/whitepaper-tech.pdf, 2018
[3] Paulus Nicolas Meessen, “Long term data storage using peer-to-peer technology,” Radboud University, 2017
[4] Protocol Labs, “Filecoin: A Decentralized Storage Network,” https://filecoin.io/filecoin.pdf, 2017