基于比特币的私有信息检索支付协议

丁佳晨, 俞能海, 林宪正, 张卫明

中国科学技术大学, 中国科学院电磁空间信息重点实验室, 合肥 中国 230027

摘要 私有信息检索(PIR)是一种密码学工具, 使用户能够从远程数据库服务器中获取信息, 而不会让服务器知道用户获取了什么信息。PIR 方案基本上由两种类型组成, 即信息论私有信息检索(Information Theoretic PIR, IT-PIR)和计算性私有信息检索(Computational PIR, C-PIR)。IT-PIR 方案要求服务器之间不共谋。一旦服务器共谋, 就无法保证用户的隐私。共谋问题一直以来都没有一个比较好的解决办法。比特币和区块链的出现为解决公平和信任的问题提供了一种新方法。在本文中, 我们创新地使用区块链来处理 IT-PIR 中的共谋问题, 提出了一种基于比特币的 PIR 支付协议。在此支付协议中, 客户通过比特币交易支付服务费。我们通过比特币脚本控制交易兑现的条件, 使得如果服务方相互串通, 则使服务方受到利益损失。通过这种方式, 该支付协议可以在一定程度上降低共谋的可能性。

关键词 比特币; 区块链; 私有信息检索

中图法分类号 TP309.2 DOI 号 10.19363/J.cnki.cn10-1380/tn.2019.11.01

Bitcoin-based Payment Protocol for Private Information Retrieval

DING Jiachen, YU Nenghai, LIN Xianzheng, ZHANG Weiming

Chinese Academy of Science Key Laboratory of Electromagnetic Space Information, University of Science and Technology of China, Hefei 230027, China

Abstract Private information retrieval(PIR) is a cryptographic primitive that enables the client to get information from a remote database server without letting the server know what information the client fetched. PIR schemes basically consist of two types, information-theoretic PIR(IT-PIR) and computational PIR(C-PIR). IT-PIR schemes require that the servers are non-cooperating. Once the servers collude, the client's privacy cannot be guaranteed. The emergence of Bitcoin and blockchain provides a new way to solve the problem of trust. In this paper, we innovatively use blockchain to tackle the colluding problem of IT-PIR, propose a PIR payment scheme based on Bitcoin. In this payment scheme, the client pays the service fee by Bitcoin transactions. We control the conditions of the redemption of the transaction, makes the servers suffer a loss of profits if they collude with each other. In this way, this payment protocol can decrease the probability of colluding to some extent.

Key words Bitcoin; blockchain; private information retrieval

1 引言

信息检索(Information retrieval, IR)是用户从数据集合中检索所需信息的方法。用户根据自己的需求生成查询请求并将查询请求发送到存有数据库的服务器, 然后服务器将一个或多个相关结果返回给用户。随着互联网的普及, 在线的信息检索在人们的生活中变得非常普遍。同时, 用户的检索隐私也更难以得到保护。网络内容提供商会保存用户的检索记录, 通过这种方式完善用户肖像。这些网络服务提供商可以使用这些用户肖像来提供有针对性的广告或内容推荐来赚钱。在一些情况下, 用户希望他们的检索内容是私密的。比如, 投资者在查看股票市场数据库中某些股票当前的市场价值时可能不愿意透露他们对该股票感兴趣, 因为这可能会无意中影响该股票价格。当用户向服务器提交查询请求时, 他希望他的查询信息不会泄露。因此, 私有信息检索(PIR)的概念被提出, 它是一种密码学工具, 旨在保护用户在

通讯作者: 俞能海, 博士, 教授, Email: ynh@ustc.edu.cn。

本课题得到国家重点研发计划基金资助项目(No.2018YFB0804100)资助。

收稿日期: 2019-04-11; 修改日期: 2019-05-26; 定稿日期: 2019-11-04

检索信息时的隐私。一个简单的解决方案是, 当用户想要检索时, 他可以请求数据库的完整副本并自己在本地搜索他需要的信息。但是这种解决方案有着大量的通信开销, 在数据库很大时并不实用。PIR 的目的是在用户进行检索时保护查询的隐私的同时减少通信的开销。

PIR 协议主要分为两类。一类是信息论私有信息检索(IT-PIR), 是指可以提供信息论的安全性的 PIR 方案。1995 年, Chor 等人[1]提出了第一个 PIR 方案, 它能够隐秘地从复制的数据库中检索内容, 并且只需要少量的通信开销。他们的工作启发了该领域的一系列重要的后续研究[2-4]。Chor 等人[1]在中表明, 如果数据库存储在单个服务器中, 那么用户实现信息论 PIR 的唯一方法就是下载整个数据库, 这有着巨大的通信开销。一种降低通信的方法是在多个不共谋的服务器上存储复制数据库。另一类是计算性私有信息检索(C-PIR), 它提供较弱的安全性。只要服务器没有足够的计算能力来解决某个计算上的难解问题, 它就可以保证用户检索的隐私。1997 年, Kushilevitz 等人[5]在中提出了一种计算性 PIR 协议, 可以用单个服务器实现。基于二次剩余问题的困难性该方案可以被证明是安全的。通常, c-PIR 方案非常低效且不实用, 因为它使用一些诸如同态加密之类的密码学工具, 其计算速度很慢。

共谋是 IT-PIR 协议中的一个大问题。IT-PIR 协议要求不能全部的服务器都共谋。假设存在 N 个存有复制数据库的服务器, t-private IT-PIR 协议意味着它可以容忍 N 个服务器中最多 t 个服务器串通并且仍然能保护用户查询信息的隐私。通常我们有 $1 <= t <= N-1$。Goldberg 等人[6]在种提出了一种鲁棒的 PIR 协议, 它是 t-private 的, 并且当所有服务器共谋时它仍然是一个计算性的私有信息检索。在该方案中, 用户使用加法同态加密将检索请求加密发送到服务器。Devet 等人[7]也在中提出了一种混合的私有信息检索方案来处理共谋问题。结合 IT-PIR 和 c-PIR 是处理 IT-PIR 中的共谋问题的一种方法, 但如上所述, 由于使用了 c-PIR, 它是低效的并且依赖于计算上难解问题的安全性。

一些 IT-PIR 方案非常有效并且通信开销很低。如果我们能找到解决共谋问题的方法, 那么这些 IT-PIR 方案将更加实用。我们试图从经济利益的角度来抑制共谋。如果我们可以设计一个协议, 使得其中参与共谋的各方将遭受经济损失, 那么各方出于利益考虑就不会共谋。可以使用智能合约来设计协议来限制每个参与者的行为。其中智能合约是一组由计算机系统自动执行的数字定义的承诺。近些年出现的区块链系统具有很多优点, 如去中心化, 防篡改等, 使其成为实施智能合约的理想选择[8]。

1.1 本文工作

在本文中, 我们创新地将 PIR 和区块链结合起来, 提出了一种基于比特币的 PIR 支付协议。在此支付协议中, 用户通过比特币交易支付服务费用。我们通过控制交易兑现的条件, 使得如果服务方相互串通, 则服务方会遭受经济上的损失。通过这种方式, 该支付协议可以在一定程度上降低共谋的可能性。

1.2 本文框架

本文的其余部分安排如下。首先在第 2 节中介绍一些预备知识; 然后我们将在第 3 节提出基于比特币的 PIR 支付协议, 并在第 4 节中分析我们提议的支付协议的安全性和可行性; 在第 5 节中, 我们将模拟我们的协议; 最后, 在第 6 节中我们将简要总结并讨论我们未来的工作。

2 预备知识

在本节中介绍相关的预备知识。在 2.1 节中介绍比特币及比特币交易以及比特币脚本相关知识, 在 2.2 节介绍比特币限时承诺方案。

2.1 比特币及比特币交易

比特币是由中本聪在 2008 年提出的[9], 它被认为是第一个基于区块链的数字货币系统。在比特币系统中, 交易将比特币从一组地址转移到另一组地址。比特币网络中的节点广播它想要进行的一些交易。然后矿工将这些交易收集到块中, 检查其有效性, 并通过共识协议来将区块添加到区块链上。比特币系统中使用的共识协议是工作量证明(Proof of Work, PoW), 可以容易地将其理解为已经完成一定工作量的一个证明。每个人都能够将区块添加都区块链上, 一个人的工作量决定了他将区块附加到区块链上的可能性。由于工作量证明这种共识机制, 如果想篡改比特币的账本, 需要控制比特币网络中总计 51%的网络计算能力。因此, 比特币的两大优势是去中心化和不可篡改。除此之外, 比特币系统的脚本系统可以实现一些简单的智能合约。我们可以安全地在各种条件下实现比特币的转移。因此, 当多方想要进行金融交易时, 如果要求这笔交易需要在某些条件下才会触发, 那么使用比特币合约可以安全地实现这一目标。

比特币脚本可以用来设计公平性的协议。Andrychowicz 在[10]中使用比特币交易脚本来实现安全多方计算协议。他们使用比特币交易的时间锁来构建基于比特币的限时承诺方案, 该方案允许用户首先对秘密值做出承诺, 如果他在时间 t 之前没有公布他的秘密消息, 他将根据押金交易而受到相应的罚款。这种限时承诺方案是其安全多方协议的核心部分。Bentov 等人[11]在中解释了如何使用比特币网络来实现公平的多方协议。除此之外, 比特币脚本合约还可用于实现云计算中的公平支付。Huang 等人[12]在中为外包计算提出了一种公平的支付协议。Zhang 等人[13]在中介绍了BCpay, 它可以在没有任何第三方的情况下实现外包服务的公平支付。

比特币是基于交易而非基于账户的, 因此交易是比特币系统中最重要的部分。交易是一种数据结构, 它将比特币从未被使用的交易的输出(UTXO)发送到到一组地址。地址是公钥PK的哈希值(使用SHA-256 哈希函数), 用户拥有的地址以及对应的私钥 SK 。在比特币系统中, 公私钥对由椭圆曲线数字签名算法(ECDSA)生成。我们使用 PK_A 和 SK_A 表示用户 A 的一对公私钥, 而 $sig_A(m)$ 表示用户 A 使用私钥 SK_A 对消息 m 的签名。假设有另一个用户 B 拥有密钥对 (PK_B, SK_B), 并且 A 想要将价值 v 的比特币发送给 B, 这意味着将 v 从 PK_A 的地址转移到地址 PK_B。此交易 T_x 可表示为

$$T_x = (y, PK_B, v, sig_A(T_y, PK_B, v)),$$

其中, y 是输入交易 T_y 的索引, PK_A 必须是交易 T_y 的接收者。PK_B 是交易 T_x 的收款人, 即交易 T_x 可以兑换 SK_B。这是最基本的交易类型。

在真正的比特币系统中, 交易可以有多个输入, 以防止单个输入没有足够的金额。交易中还可以有一个锁定时间 t, 使交易在时间 t 之后才生效。比特币脚本语言是基于堆栈的, 可以实现一些更复杂的交易。交易有一个输出脚本 π_x, 并且兑现此交易的方法是提供一个输入脚本以使得输出脚本的输出为 $true$。脚本可以是对交易正文的签名, σ。这个更复杂的交易可以表示为

$$T_x = (y_1, \cdots, y_n, \pi_x, v, t, \sigma_1, \cdots, \sigma_n).$$

此交易有多个输入交易 $(T_{y_1}, \cdots, T_{y_n})$,在时间 t 之前不能被兑现, $(\pi_1, \cdots, \pi_n)$分别是输出脚本。此交易只有在达到时间 t 之后才有效且 $(\sigma_1, \cdots, \sigma_n)$可以使 $(\pi_1, \cdots, \pi_n)$输出分别为 $true$。我们可以用以下形式表示此交易, 在本文的其余部分, 我们使用图 1 的形式来表示交易。

T_x (**in**: $T_{y_1}, \cdots, T_{y_n}$)
inscript: σ_1 … **inscript**: σ_n
outscript: $(body, arg)$: $\pi_x(body, arg)$
val: v B
timelock: t

图 1 比特币交易

Figure 1 Structure of Bitcoin transaction

2.2 比特币限时承诺方案

比特币限时承诺方案是 Andrychowicz 在[10]中提出的。我们的协议的思想与这个方案非常相似, 都是对于一笔交易有两种兑现方式, 一种是正常的兑现方式, 另一种是惩罚的兑现方式。在这个承诺方案中, 承诺方 C 必须对消息 m 做出承诺, 如果他在给定时间 t 内没有公布消息 m 的内容, 他将受到惩罚。承诺方创建一个价值为 v 的交易 $T_xCommit$, 如图 2 所示。

$T_xCommit$ (**in**: T_x)
inscript: $sig_C(T_xCommit)$
outscript: $(body, \sigma_1, \sigma_2, m)$: $\big(H(m) = h \wedge ver_C(body, \sigma_1)\big) \vee$ $(ver_C(body, \sigma_1) \wedge ver_P(body, \sigma_2))$
val: v B

图 2 承诺交易

Figure 2 Structure of *Commit*

此交易能被两种方式兑现。第一种方式是承诺方公布秘密消息 m, 满足其哈希值为 h, 这样承诺方可以拿回他的押金。 第二种方式是, 在时间 t 之后, 如果此交易尚未被兑现, 则接受承诺的一方 P 可以通过自己的私钥兑现这个交易。这样使得承诺方如果想拿回押金就需要在时间 t 之前公布消息。

3 基于比特币脚本的 PIR 支付协议

3.1 模型及定义

在 IT-PIR 协议中, 多个服务器存储相同的数据副本。当用户想要检索某些信息时, 相应的检索请求被发送到多个服务器, 并且在接收到所有请求的正确返回结果之后, 可以通过这些返回结果计算出他

想要检索的信息。对于服务器, 如果收到检索请求的所有服务器都串通并共享用户的检索请求, 则他们可以恢复用户的真实请求意图。因此, 共谋一直是 IT-PIR 协议中的一个大问题。

在支付协议中, 假设有多个服务方拥有自己的公共数据库。这些服务方形成了一个存储数据的联盟链。每一服务方都有一个节点, 它将存储整个链, 这意味着所有节点都将拥有链的数据的完整副本。因此, 这些节点可以用作 PIR 协议的服务器。我们假设这些节点是诚实但好奇的, 并且可能与其他节点串通以获取用户信息。除此之外, 假设用户不会与服务方串通。

在支付协议中, 有一个用户 U, 密钥 (PK_U, SK_U)。作为提供 PIR 服务的各方, 有 n 个节点代表对应方。这 n 个节点分别具有密钥对(PK_{S_i}, SK_{S_i})。除此之外, 这些方共享一对密钥 (PK_B, SK_B)。

为了方便对协议的理解, 先在表 1 中介绍相关记号。

表 1 协议相关记号

Table 1 Relative Notations

记号	含义
T_x (**in**: $T_{y_1}, \cdots, T_{y_n}$)	交易 T_x 的输入交易为 $(T_{y_1}, \cdots, T_{y_n})$
$sig_C(T_x Commit)$	C使用私钥SK_C对交易$Commit$进行签名
inscript: $\emptyset, sig_U(T_x Charge_i)$	输入脚本的参数, $\emptyset$表示第一个参数为空, 第二个参数为 U 对交易$Charge_i$的签名
$ver_{S_i}(body, \sigma_1)$	使用 S_i 的公钥 SK_{S_i} 验证签名 σ_1 是否正确
$ver_{S_i}(body, \sigma_{S_i}) \wedge ver_U(body, \sigma_U)$	符号 $\wedge$ 表示与, 在输出脚本中, 表明兑现交易的条件是两个签名都验证通过
$ver_{S_i}(body, \sigma_{S_i}) \vee ver_U(body, \sigma_U)$	符号 $\vee$ 表示或, 满足一个条件就可以
$checkmultisig_2[PK_1, PK_2, PK_3]$ $(\sigma_1, \sigma_2, \sigma_3)$	验证多重签名, 在三个签名中只要有两个就验证通过
val: v B	交易的金额为价值为 v 的比特币

3.2 PIR 支付协议

协议的主要步骤如图 3 所示。

Setup: 用户 U 根据他想要检索的内容和相应的 PIR 协议构造自己的查询请求。查询请求表示为 $(q_1, \cdots, q_n)$, 相应的比特串为$(m_1, \cdots, m_n)$。用户还将生成一个随机数 R, 并将 R 发送到各个服务器。这些比特串与随机数 R 组合来生成一组密钥对 (PK_{m_i}, SK_{m_i})。对组合比特串进行两次 $SHA256$ 计算得到私钥:

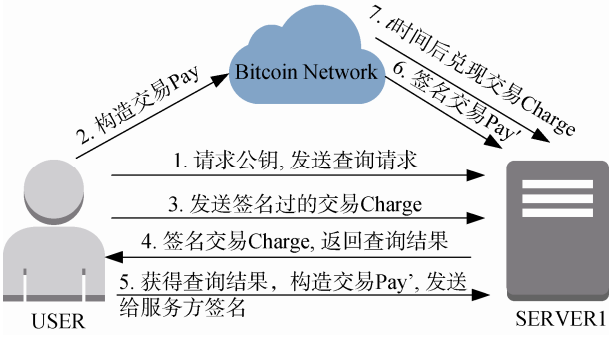

图 3 协议步骤

Figure 3 Procedure of the protocol

$$SK_{m_i} = SHA256\,(SHA256\,(m_i \,||\, R)),$$

然后根据私钥生成相应的公钥。

用户向服务器请求服务器的公钥 $(PK_B, PK_{S_1}, \cdots, PK_{S_n})$。然后, 用户将查询请求的比特串和随机数 R 发送到相应的服务器。

Payment: 用户 U 创建比特币交易来为 PIR 服务支付相应的费用。对于每个服务器, U 将创建一个价值为 v 的交易 Pay_i 并发布在比特币网络中。交易结构如图 4 所示。

$T_x Pay_i$ (**in**: T_x)
inscript: $sig_U(T_x Pay_i)$
outscript $(body, \sigma_1, \cdots, \sigma_n, \sigma_U, \sigma_B, \sigma_{S_i})$: $(ver_{m_1}(body, \sigma_1) \wedge \cdots \wedge ver_{m_n}(body, \sigma_n) \wedge ver_B(body, \sigma_B)) \vee (ver_{S_i}(body, \sigma_{S_i}) \wedge ver_U(body, \sigma_U))$
val: v B

图 4 交易Pay_i结构

Figure 4 Structure of Pay_i

交易 Pay_i 有两种兑现方式。在正常情况下, 在时间 t 之后, 此交易将由交易 $Charge_i$ 兑现, 对应着输出脚本中符号 “∨”后面的兑现方式。交易结构如图 5 所示。

交易的主体 $Charge_i$ 由用户 U 创建, 然后用户对交易进行签名并将其发送到对应的服务器 S_i。

Computation and Redeem: 服务方 S_i 检查用户发送的交易 $Charge_i$ 后对交易进行签名。交易 $Charge_i$ 只能通过服务器 S_i 的私钥的签名兑现。服务方 S_i 根据查询请求 m_i 计算出结果并返回给用户。然后 S_i 可以在时间 t 之后通过发布并兑现交易 $Charge_i$来获得服务费用。交易 Pay_i 也可以通过另

一个交易 $Collude$ 兑现(交易结构如图 6 所示), 这种方式需要服务方之间共谋, 因为需要一组私钥 $(SK_{m_1},\cdots,SK_{m_n},SK_B)$ 的签名。

$T_xCharge_i$ (**in**: T_xPay_i)
inscript: $\emptyset,\cdots,\emptyset,sig_U(T_xCharge_i)$, $\emptyset,sig_{S_i}(T_xCharge_i)$
outscript $(body,\sigma_1)$: $ver_{S_i}(body,\sigma_1)$
val: v B
timelock: t

图 5　交易$Charge_i$结构

Figure 5　Structure of $Charge_i$

$T_xCollude$ (**in**: T_xPay_i)
inscript: $sig_{m_1}(T_xCollude),\cdots,$ $sig_{m_n}(T_xCollude),\emptyset,sig_B(T_xCollude),\emptyset$
outscript $(body,\sigma_1)$: $ver_{S_i}(body,\sigma_1)$
val: v B

图 6　交易$Collude$结构

Figure 6　Structure of $Collude$

由于密钥 SK_{m_1} 是从消息 m_i 生成的, 所以知道这些消息的人也可以通过计算得到相应的私钥。此外, 所有服务方都知道 SK_B 。如果服务方想要知道用户想要检索哪些信息, 则他们必须彼此串通并且彼此共享他们获得的消息 m_i 。因此, 如果一个服务方知道所有查询请求比特串, 以及私钥 SK_B 他就可以兑现交易 Pay_i 。

最后, 用户再根据服务方返回的结果还原出最终结果。

这个版本的支付协议似乎在一定程度上抑制了共谋, 但仍然存在一个问题。共谋的服务方可能不通过共享他们获得的查询字符串, 而是共享他们的计算的结果, 以此来避免利益的损失。如果掌握了全部的计算结果, 他们还是可以得到用户想要查询的结果。因此, 必须修改支付协议来解决这个问题。假设服务器计算的返回结果是 $(r_1,\cdots,r_n)$。 根据我们之前生成 SK_{m_i} 的方法, U 使用 r_i 和 R 生成一组私钥 $(SK_{r_1},\cdots,SK_{r_n})$。如果将之前交易 Pay_i 中的兑现条件的私钥由请求比特串生成的私钥换成由计算结果比特串生成的私钥, 那么就能防止服务方分享计算结果。因此我们需要将交易 Pay_i 转化成一个新的交易 Pay' 。要实现这个转换, 用户需要生成新的交易 Pay_i' 来兑现之前的交易 Pay_i , Pay_i' 的输出脚本与 Pay_i 的输出脚本类似, 交易结构如图 7 所示。

T_xPay_i' (**in**: T_x)
inscript: $sig_{m_1}(T_xPay_i'),\cdots,sig_{m_n}(T_xPay_i'),$ $sig_B(T_xPay_i')$
outscript $(body,\sigma_1,\cdots,\sigma_n,\sigma_U,\sigma_B,\sigma_{S_i})$: $(ver_{r_1}(body,\sigma_1)\wedge\cdots\wedge ver_{r_n}(body,\sigma_n)\wedge ver_B(body,\sigma_B))\vee$ $(ver_{S_i}(body,\sigma_{S_i})\wedge ver_U(body,\sigma_U))$
val: v B

图 7　交易Pay_i'结构

Figure 7　Structure of Pay_i'

要兑现 Pay_i , U 需要私钥 SK_B 的签名, 所以 U 将此交易发送给任意一个服务器以获得签名。对于服务器不对此交易进行签名的情况, 我们将在后面的安全性分析中进行讨论。

兑现交易 Pay' 有两种方式, 第一种方式与 $Charge$ 相同。 第二个类似于 $Collude$, 用私钥 SK_{r_i} 替换私钥 SK_{m_i} , 交易结构如图 8 所示。这种情况下, 如果一个服务方掌握足够的计算结果 r_i , 则他可以兑现交易 Pay'。

$T_xCollude'$ (**in**: T_xPay_i')
inscript: $sig_{r_1}(T_xCollude'),\cdots,$ $sig_{r_n}(T_xCollude'),\emptyset,sig_B(T_xCollude'),\emptyset$
outscript $(body,\sigma_1)$: $ver_{S_i}(body,\sigma_1)$
val: v B

图 8　交易$Collude'$结构

Figure 8　Structure of $Collude'$

上述协议适用于可以容忍 $n-1$ 个服务方共谋的 PIR 协议, 这意味着如果不是所有服务方都串通, 那么它们无法恢复用户的查询请求。还有一些 PIR 协议只能容忍 $l-1$ 个服务方串通。所以我们必须在之前的协议中做一些改变。我们将在 Pay_i 和 $Collude$ 中使用多重签名方案, 如图 9 所示。

如果想通过交易 $Collude$ 兑现 Pay_i , 需要有私钥 SK_{m_i} 的签名以及从一组 n 个私钥 $(SK_{m_1},\cdots,SK_{m_n})$中任意 l 个私钥的签名。通过这种

方式, 只有参与共谋的服务方的服务费用会被交易 $Collude$ 兑现, 诚实的服务方仍然可以容果正常方式获得应得的费用。交易结构如图 10 所示。

$T_x Pay_i$ (**in**: T_x)
inscript: $sig_U(T_x Pay_i)$
outscript $(body, \sigma_1, \cdots, \sigma_l, \sigma_U, \sigma_B, \sigma_{S_i})$: $(checkmultisig_l[PK_{m_1}, \cdots, PK_{m_n}]\ (\sigma_1, \cdots, \sigma_l) \wedge ver_B(body, \sigma_B)) \vee (ver_{S_i}(body, \sigma_{S_i}) \wedge ver_U(body, \sigma_U))$
val: v B

图 9　多重签名方案的交易Pay_i结构

Figure 9　Structure of Pay_i with Multisignature

$T_x Collude$ (**in**: $T_x Pay_i$)
inscript: *Any l signature out of* $\left(sig_{m_1}(T_x Collude), \cdots, sig_{m_n}(T_x Collude)\right)$, $\emptyset, sig_B(T_x Collude), \emptyset$
outscript $(body, \sigma_1)$: $ver_{S_i}(body, \sigma_1)$
val: v B

图 10　多重签名方案的交易$Collude$结构

Figure 10　Structure of $Collude$ with Multisignature

以上是基于比特币的完整的 PIR 支付协议。核心思想是, 如果服务方共谋并分享他们获得的信息以便在时间 t 之前得知用户的真实查询请求意图, 他们都可以通过交易 $Collude$ 来获得本应该支付给其他服务方的费用。由于比特币的匿名性, 他们无法区分是哪个服务方兑现了这些交易拿走了这些费用。因此这些参与共谋的服务方出于利益的考虑会兑现这些交易。一旦服务方共谋, 将有一方兑现这些交易拿走费用, 其他服务方将会失去相应的服务费用。通过使用我们的支付协议, 这些服务提供方出于利益的考虑不会进行共谋行为。

4　安全性分析

4.1　协议安全性分析

在本节对基于比特币的 PIR 支付协议的安全性进行分析。根据文献[12, 14]中的安全性分析, 分析本文中提出的支付协议是否满足下列的安全性质。

(1) 完整性: 完整性指协议的结构完整, 能完成协议想要达成的目的, 也就是用户获得查询结果并能保障自己的查询隐私; 服务方可以得到相应的服务费用。

(2) 公平性: 利用比特币交易脚本实现公平性。无论是用户还是服务方, 如果不按照协议的正常流程执行, 会受到利益上的损失或者终止协议来保护其他参与方的利益。以此来保障参与方按照协议的正常流程执行。

(3) 合理性: 协议的假设以及协议整体流程需要在现实中是合理可行的。

定理 1　基于比特币的PIR支付协议满足完整性。

证明. 协议结构要完整。在正常情况下, 用户和所有服务方将按照协议的流程执行操作。用户发送查询请求和相关参数给服务方, 并将交易发布在比特币网络。服务方根据查询请求计算查询结果并返回给用户。在时间 t 之后, 服务方将通过交易 $Charge$ 兑换交易 Pay'。所有服务方都将获得应得的费用。用户会获得正确的返回结果, 还原出自己想要检索的结果。

定理 2　基于比特币的PIR支付协议满足公平性。

证明. 在假设中, 用户不会与服务方串通, 服务方将诚实地返回正确的计算结果。但他们是好奇的, 可能会相互串通以获取用户的私人信息。

考虑公平性, 协议参与方都有可能不按照协议正常流程执行。

(1) 用户不按照协议正常流程执行:

用户不构建支付交易 Pay 进行支付的话, 服务方不会提供私有信息检索服务, 协议终止。

用户不将交易 Pay 兑现, 转化为交易 Pay'。这一步的目的是为了防止服务方共享计算结果, 保护用户的隐私。所以用户不会选择不将交易Pay转化为交易 Pay'。如果用户不转化交易的话, 协议终止。服务方依然可以得到费用, 用户的隐私则不再能得到保障。

(2) 服务方不按照协议正常流程执行:

在协议的第二步中, 用户将 $Charge_i$ 签名后发送给服务方 S_i, S_i 需要对交易进行签名然后发布。如果 S_i 不对交易签名, 则协议将中止。服务方无法获得服务费用。

如果服务方选择共谋, 共享他们获得的检索请求字符串。其中一个服务方可以通过交易 $Collude$ 来兑现所有支付给每个服务方的交易 Pay, 这样其他服务方将失去其应得服务费用。出于利益的考虑, 服务方不会选择共谋。因此, 如果用户的查询消息的价值没有服务方的服务费用价值高的话, 服务方将不会选择共谋。这里需要定义用户查询消息的价值, 服务方通过用户的检索记录构建用户肖像来投放定向

广告, 通过这种方式获利。据此可以计算得到平均每个用户查询信息的价值。假设知道用户的秘密查询消息的平均收益是 V_s, 我们可以将 PIR 服务的查询费用 v 设置为大于 V_s。

如果服务方选择共谋, 共享其计算结果。与共享检索请求时类似, 一个服务方可以通过交易 $Collude'$ 来兑现所有支付给每个服务方的交易 Pay', 这样其他服务方将失去其应得服务费用。在用户将交易 Pay 兑现并转化为交易 Pay' 时, 需要私钥 SK_C 的签名。如果在此步骤时没有任何一个服务方按照协议要求对此交易进行签名, 这样他们彼此分享他们的计算结果, 他们似乎仍可以知道用户 U 的查询信息并同时获得用户支付的费用。但实际上, 知道所有返回结果的服务器仍然可以兑现应支付给其他服务方费用的交易。因为服务方都知道私钥 SK_C, 所以如果他们分享他们的计算结果, 一定有一个服务方会签名交易 Pay' 并兑现交易以获得更多的钱。

如果服务方提前协商, 以避免某一服务方支取其他服务方应得的收入。由于比特币的匿名性, 他们无法区分所有共谋方中的哪一方兑现了交易。所以很难用押金或其他方法设计一个方案来阻止其中一个服务方拿走这些费用。因此, 如果用户和各服务方都按照支付协议的流程执行, 则用户可以在不泄露个人隐私的情况下获得 PIR 的服务, 并且服务方也可以获得其相应的服务费。

定理3 基于比特币的PIR支付协议满足合理性。

证明. 在定理 2 中我们表明了共谋的服务方会失去其应得的服务费用, 以此来说明服务方出于利益考虑不会共谋。得出这个结论是基于一个假设——协议的参与方都会优先考虑自己的利益, 对于用户来讲就是保护自己的检索隐私; 对于服务方来讲就是自己获得的收入, 并且参与方都不会做令自己利益受损的行为。这个假设在现实中是合理可行的, 所以协议中通过利益来抑制共谋满足合理性。

4.2 不足

协议通过服务方不会做损害利益的行为这个假设来处理共谋问题, 这个假设在现实是有合理性的, 然而并不是绝对的。因此协议不是绝对安全的协议, 不能保证服务方绝对不会共谋, 而是提供一种从经济利益的角度防止共谋的方法。

协议只能在时间 t 内防止共谋行为, 时间 t 之后服务方共谋不会受到损失。对于只需要一段时间内保护查询隐私的用户, 适合于此协议。对于需要一直保护查询隐私的用户, 则不适合此协议。用户很多需要保密的检索内容都是具有一定时效性的, 比如用户查询某个科研想法是否有人实现, 在用户自己完成这项科研工作后, 这个查询便不需要继续保密。因此在一段时间内防止服务方共谋来提供保密的私有信息检索是具有实用性的。

5 协议实施以及评估

我们已经实施并测试了我们的支付协议。模拟了三个服务方和一个用户的情况。我们在支付协议中构建了交易并将其发布在比特币测试网中。这些交易是使用 cryptos[15]创建的, 这是一个用于数字货币的构建签名和交易的 Python 库。为简单起见, 我们使用 P2SH(Pay to Script Hash)交易。例如, 在交易 Pay_i 中, 输出脚本内容如图 11 所示。

4 <Pubkey m1> <Pubkey m2> <Pubkey m3> <Pubkey B> 4 OP_CHECKMULTISIG	第一部分
OP_ROT OP_ROT	第二部分
2 <Pubkey S1> <Pubkey U> 2 OP_CHECKMULTISIG	第三部分
OP_BOOLER OP_VERIFY	第四部分

图 11 交易Pay_i输出脚本
Figure 11 Out-script of Pay_i

脚本的第一部分检查堆栈顶部的 4 个签名是否合法, 这部分对应于兑现此交易的第一种方法。脚本第二部分将前两个参数移动到堆栈顶部。脚本的第三部分对应于兑换此交易的第二种方法。最后的部分 <OP_BOOLOR OP_VERIFY> 检查两种方法中的一种是否验证正确。如果我们使用 P2SH 交易, 则脚本如图 12 所示。

OP_HASH160 [20 bytes hash value of the script] OP_EQUAL

图 12 P2SH 交易输出脚本
Figure 12 Out-script of P2SH transaction

输出脚本将替换为兑换脚本的哈希, 并且当接收者使用此交易输出时, 将在输入脚本中提供完整的输出脚本。因此交易发送者不需要在交易中提供复杂的输出脚本。通过使用 P2SH 交易, 我们可以直接使用 P2SH 比特币地址进行支付。在我们的模拟中, 交易都可以按协议设想的方式被兑现。

在比特币测试网中, 交易很快就会得到确认, 但在真正的比特币网络中, 交易将在一段时间后得到确认。一些交易甚至需要几天都得不到确认。通

常, 支付的交易费用越高, 交易确认的速度就越快。因此, 如果我们想减少整个协议的执行时间, 我们需要支付更多的交易费用。但即使支付跟多的交易费用, 比特币网络的交易吞吐量还是太低, 因此更可靠的方式是使用另一个具有更高吞吐量的区块链平台。本文中采用比特币主要是因为比特币是目前最广为人知的加密数字货币平台, 方便理解协议的构造。

在支付协议中, 需要在比特币网络上发布大量交易。因此, 它将增加用户和服务提供商的通信开销。我们需要保持 PIR 方案的通信开销低于下载整个数据库的开销。除了请求消息和返回的结果, 附加通信包括随机数 R 和 $4\times N$个交易, 其中 N 是服务器的数量。对于固定数量的服务器 , 额外的通信可以看作是一个常数 C , 因此我们可以找到一个适当的的 PIR 方案, 以确保总通信开销是低于下载整个数据库的。

6 结论

在本文中, 提出了一种基于比特币的 PIR 支付协议。我们首次提出利用比特币和区块链的优良特性来解决 IT-PIR 中的共谋问题。在我们的协议中, 用户通过比特币交易支付服务费, 并且该交易可以在不同条件下兑换。我们设计将用户的搜索信息构建为私钥, 并将知道用户的搜索信息与撤销应支付给服务提供商的费用相关联。为了获得相应的费用, 计算方不愿意与其他方共享用户的搜索信息。从而在一定程度上减少了共谋的发生。

在未来的工作中, 我们将继续关注区块链在 PIR 中的应用。除了比特币, 以太坊[16]是另一种流行的加密货币, 它是最早提供图灵完全智能合约的区块链之一。智能合约在以太坊虚拟机(EVM)中以字节码执行, 它可以遵循更复杂的计算和事务逻辑。如今一些区块链系统可以实现更好的匿名性, 例如 Zcash 和 Monero。他们分别使用零知识证明和环签名来保护交易中用户的隐私。Ishai 等人[17]在中, 提出了一种基于匿名网络的 PIR 方案, 可以尝试使用智能合约以及区块链的匿名性来构建可行且有效的 PIR 方案。

参考文献

[1] Chor B, Goldreich O and Kushilevitz E, et al. “Private information retrieval,” *Proceedings* of *IEEE 36th Annual Foundations of Computer Science*(FOCS’95), pp. 41-50,1995.

[2] Beimel A, Ishai Y, and Kushilevitz E. “General constructions for information-theoretic private information retrieval,” *Journal of Computer and System Sciences*, 71(2): pp. 213-247, 2005

[3] Beimel A, Ishai Y and Kushilevitz E, “Breaking the O (n/sup 1/(2k-1)/) barrier for information-theoretic Private Information Retrieval,” *The 43rd Annual IEEE Symposium on Foundations of Computer Science*(FOCS’02), 2002. Proceedings, pp. 261-270, 2002.

[4] Yekhanin S. “Private information retrieval[M]//Locally Decodable Codes and Private Information Retrieval Schemes,” Springer, Berlin, Heidelberg, pp. 61-74,2010.

[5] Kushilevitz E and Ostrovsky R, “Replication is not needed: Single database, computationally-private information retrieval,” *Proceedings 38th Annual Symposium on Foundations of Computer Science*(FOCS’97), pp. 364-373,1997.

[6] Goldberg I, “Improving the robustness of private information retrieval,” *2007 IEEE Symposium on Security and Privacy* (SP’07), pp. 131-148, 2007.

[7] Devet C and Goldberg I, “The best of both worlds: Combining information-theoretic and computational PIR for communication efficiency,” *International Symposium on Privacy Enhancing Technologies Symposium. Springer*, Cham, pp. 63-82, 2014.

[8] C.G.Ma, J.An and W.Bi, “Smart Contract in Blockchain”, *Netinfo Security*, 2018, 18(11): pp. 8-17.
(马春光, 安婧, 毕伟, “区块链中的智能合约”, *信息网络安全*, 2018, 18(11):8-17.)

[9] Nakamoto S, “Bitcoin: A peer-to-peer electronic cash system,” http://bitcoin.org/bitcoin.pdf, 2008.

[10] Andrychowicz M, Dziembowski S and Malinowski D, “Secure multiparty computations on bitcoin,” *2014 IEEE Symposium on Security and Privacy*(SP’14), pp. 443-458, 2014.

[11] Bentov I and Kumaresan R, “How to use bitcoin to design fair protocols,” *International Cryptology Conference. Springer*, Berlin, Heidelberg, pp. 421-439, 2014.

[12] Huang H, Chen X and Wu Q, “Bitcoin-based fair payments for outsourcing computations of fog devices,” *Future Generation Computer Systems*(FGCS’18), 78: pp. 850-858, 2018.

[13] Zhang Y, Deng R and Liu X, “Outsourcing service fair payment based on blockchain and its applications in cloud computing,” *IEEE Transactions on Services Computing*, 2018.

[14] J.X.Wu, Y.Gao and Z.Y.Zhang, “A Multi-Party Privacy Preserving Fair Contract Signing Protocol based on Blockchains”, *Journal of Cyber Security*, 2018,3(3): pp. 43-50, 2018.
(吴进喜, 高莹, 张宗洋, “基于区块链的多方隐私保护公平合同签署协议”, *信息安全学报*, 2018, 3(3) : pp. 43-50, 2018.)

[15] “Pycryptotools, Python library for Crypto coins signatures and transactions”, https://pypi.org/project/cryptos/1.36/

[16] Wood G, “Ethereum: A secure decentralised generalised transaction ledger,” https://ethereum.github.io/yellowpaper/paper.pdf,

2014.

[17] Ishai Y, Kushilevitz E and Ostrovsky R, “Cryptography from anonymity,” *IEEE Symposium on Foundations of Computer Science* (FOCS'06), pp. 239-248, 2006.

丁佳晨 于 2016 年在中国科学技术大学信息安全专业获得学士学位。现在中国科学技术大学学校信息安全专业攻读硕士学位。研究领域为区块链相关。研究兴趣包括: 区块链、密文计算。Email: djccmf@mail.ustc.edu. cn

林宪正 于 2010 年在台湾交通大学获得博士学位。现任中国科学技术大学网络空间安全学院教授。研究领域为纠错码技术、资料储存码, 区块链。研究兴趣包括: 信道编码算法、区块链。Email: sjlin@ustc.edu.cn

俞能海 于 2004 年在中国科学技术大学信息与通信工程专业获得博士学位。现在中国科学技术大学信息处理中心主任, 多媒体与通信研究实验室主任。研究领域为图像处理与视频分析、媒体内容安全、网络通信与安全等。Email: ynh@ustc.edu.cn

张卫明 于 2005 年在中国人民解放军信息工程大学获得博士学位。现任中国科学技术大学网络空间安全学院教授。研究领域为多媒体安全、信息隐藏、人工智能安全等。研究兴趣包括: 信息安全。Email: zhangwm@ ustc.edu.cn