arXiv:2305.04062v1 [cs.DC] 6 May 2023

一种基于区块链的平台，用于大规模模型的可靠推理与训练

Sanghyeon Park首尔国立大学首尔，大韩民国
lukepark@snu.ac.kr

Junmo Lee首尔国立大学首尔，大韩民国
junmo.lee@snu.ac.kr

Soo-Mook Moon首尔国立大学首尔，大韩民国 smoon@snu.ac.kr

摘要——随着人工智能（AI）持续渗透各个领域，人们日益关注人工智能驱动的推理和训练过程中的信任与透明度问题，尤其是潜在偏见及可追溯性方面的挑战。区块链等去中心化解决方案虽已被提出以应对这些问题，但在处理大规模模型时往往难以胜任，导致推理耗时且训练验证效率低下。

为克服这些局限，我们推出 BRAIN（基于区块链的可靠人工智能网络），这是一个专为确保大模型可靠推理与训练而设计的新型平台。BRAIN采用独特的两阶段交易机制，通过分离请求交易与响应交易，利用流水线处理实现实时处理。每个随机选出的推理委员会都会提交并揭示推理结果；智能合约达成共识后，便使用共识结果执行所请求的操作。此外，BRAIN通过随机选出的训练委员会开展训练。委员会成员提交提交交易和揭示交易，并附上各自的分数，从而依据分数的中位数进行本地模型聚合。

实验结果表明，BRAIN 能以合理的燃料费实现显著更高的推理吞吐量。具体而言，BRAIN 的每秒任务数性能是朴素单阶段实现的 454.4293 倍。

索引词—区块链、大规模模型、可验证随机函数、联邦学习

I. 引言

基于大规模深度人工神经网络的人工智能（AI）迅速发展，为各行各业带来了变革，并提供了众多强大的人工智能服务 [1], [7], [23], [39]。然而，这些网络通常依赖中心化服务器进行学习和推理，引发了有关信任的担忧，例如学习过程缺乏透明度，以及推理结果可能遭到操纵 [16], [26]。例如，人工智能算法中的偏见可能会对某些种族和族裔群体造成不成比例的影响，人脸识别系统 [8] 和刑事司法应用 [6] 中的问题便是明证。学习过程缺乏透明度，使用户难以评估这些服务的公平性及其伦理影响，进而削弱了人们对人工智能系统的信任。同样，中心化推理流程也面临挑战，因为推理结果可能遭到操纵或滥用，且人工智能生成的输出的真实性难以验证。例如，假新闻大量传播，以及人工智能生成的推理结果 [45], [53] 被不当利用。

表 I：BRAIN 与现有研究的比较

	计算引擎		信任机器		大型模型	No 硬分叉
	训练	推理	训练	推理		
DeepBrain [18]	●	●	○	○	●	×
Cortex [11]	○	●	○	●	○†	×
AI News [42]	×	×	○	○	●	●
Flchain [3]	●	○	●	○	●	×
Baffle [38]	●	○	○	○	◒‡	●
Fedcoin [27]	●	○	●	○	●	×
BFLC [25]	●	○	●	○	●	×
Blockflow [31]	●	○	●	○	●	●
DFL [44]	●	○	●	○	◒§	×
BlockDFL [37]	●	○	●	○	●	×
iDML [31]	●	○	●	○	●	●
ZEN [17]	×	×	○	●	○*	×
ZKP-FL [50]	●	○	●	○	○*	●
VOC [21]	●	○	●	○	○*	●
Y. Fan [15]	×	×	○	●	○*	×
BRAIN（我们提出的方案）	●	●	●	●	●	●

† 对交易吞吐量产生负面影响。此外，它无法支持推理时间超过区块间隔的模型。‡ 由于涉及巨额成本，在以太坊 [47] 上。§ 成本取决于模型大小。* 由于所需时间过长。

这种透明度不足会削弱用户对人工智能服务的信任，因为用户难以判断其真实性。这可能导致用户成本增加，因为他们可能需要另行进行搜索、审计或第三方评估，以确保服务可信。服务提供商也可能不得不投入更多营销工作，以说服用户相信其系统的可靠性。

A. 相关工作与挑战

人工智能训练和推理过程的中心化一直是一个重大问题，因此人们开始探索将去中心化技术作为解决方案。其中一种技术就是区块链。然而，事实证明，将区块链与人工智能相结合颇具挑战。尽管一些研究已朝着解决这一问题取得进展，但仅取得了部分成功。

[42] 提出了将区块链与人工智能结合使用的服务，但仅将区块链用作数据库和激励平台，并未提出直接提升人工智能训练和推理可信度的方法。 [38] 尝试通过区块链降低聚合器成本，但仍未能提供信任保障。此外，在区块链上存储模型权重的高昂成本，无法有效支持大型神经网络。与此同时，其他研究 [3], [27], [31] 通过

使用激励或罚没存款，引导训练朝正确方向进行，但未能为推理提供可信度。 [25] 提出了一种改进神经网络的方法，即评估并聚合委员会训练的神经网络，但同样未考虑推理。此外，在 [11], 中，尽管通过在区块链上执行推理，并允许所有参与节点验证推理结果，建立了信任，但由于节点必须等待且无法处理其他交易，因此无法支持推理耗时较长的大型神经网络。该方法也未涉及训练。借助 [18], ，参与节点的图形处理器可以出租并用于训练和推理，但无法保证结果的可信度。

零知识证明（零知识证明）可用于验证训练和推理，正如 [15], [17], [21],[50] 所示。它们不需要重复验证；只需提供证明即可验证操作是否正确执行。然而，由于证明生成时间极长，零知识证明仅适用于小型模型，例如 MobileNet [22]。

这些相关研究的总结见表 I。我们将支持程度标记为: 表示完全支持，表示部分支持，表示不支持， ×表示与相关功能无关。特别是，可以看出，区块链尚未被充分用作计算引擎或信任机器。即使提供了信任，也无法支持大模型，因为存在时间限制或成本过高的问题。

B. 贡献

我们提出的基于区块链的可靠人工智能网络（ BRAIN）有效解决了先前研究面临的挑战。据我们所知，BRAIN 是首个能够处理大型神经网络推理和训练，同时兼容现有区块链的区块链平台。

BRAIN 的主要贡献如下:

- 我们提出了 BRAIN，这是一种面向大规模神经网络的创新架构，旨在解决现有区块链中的延迟、冗余、验证和兼容性问题。
- BRAIN 的关键组件之一是无聚合器联邦学习。该方法利用智能合约实现异步模型更新，从而无需聚合器即可实现收敛。
- 仿真结果表明，BRAIN 在性能方面有所提升，且成本合理；我们还提供了优化超参数配置的指南。
- 我们强调区块链驱动的人工智能对于信任、透明度和可追溯性的重要性。此外，我们还重点介绍了 BRAIN 在各种应用中的潜在影响。

我们的实现代码和实验代码已公开发布，可供复现和进一步研究，详见 GitHub[1]。

II. 背景

本节介绍与 BRAIN协议 相关的重要背景知识。我们将简要讨论**可验证随机函数**，它支持密码学抽签；以及**联邦学习**，这是一种由多个参与者共同进行的训练方法。

A. 可验证随机函数

可验证随机函数（VRFs） [30] 是能够输出可验证伪随机值的函数。它们由三个函数组成：密钥生成、求值和验证。

- 密钥生成$(r) \rightarrow (pk, sk)$: 为给定的随机输入 r生成一对公钥 pk 和私钥 sk 。
- 求值$_{sk}(x) \rightarrow (y, \pi)$: 根据私钥 sk 和 x 作为输入，返回伪随机输出y 和证明 π 。
- 验证$^{pk}(x, y, \pi) \in$ {真，假}: 接受公钥 pk、伪随机输出 y、证明 π和消息 x 作为输入；如果 y 确实是由 sk 和 x 生成的输出，则返回真。否则，返回假。

密码学抽签是一种基于密码学计算而非预定义规则来选出成员的方法。参与者可以确认自己当选，但在可验证证明发布之前，无法确认其他人是否当选。在 BRAIN 中，密码学抽签通过可验证随机函数实现，与 Algorand [19] 类似。由于 sk 彼此之间不会共享，因此在结果提交之前，无法得知本轮哪些人当选委员会。这使得贿赂等恶意行为无法实施。

B. 联邦学习

联邦学习是一种让客户端通过使用各自的数据训练本地模型，从而训练全局模型的技术。随后，聚合器收集本地模型的更新，并将其整合以创建下一个全局模型。一种广为人知的实现方法是联邦平均（FedAvg） [29],，该方法根据客户端持有的数据量，通过加权平均聚合客户端的更新。此外，FedAsync [49] 研究表明，即使在异步环境中，学习也能收敛。

BRAIN 旨在让参与者并行执行模型更新，其灵感源自 FedAsync 方法。不过，为了实现去中心化，BRAIN 使用智能合约而非聚合器。此外，BRAIN 采用与 BFLC [25] 类似的 K 折交叉验证技术来评估更新提案。该技术将每位参与者持有的本地数据集用作验证集，使系统能够抵御诸如投毒攻击 [5], [51]之类的攻击。

III. BRAIN：概述

BRAIN 是一个去中心化区块链 AI 平台，提供训练和推理服务。本节全面概述 BRAIN，重点介绍**参与者** 的角色，以及**超参数**在维持平台最佳运行状态方面的重要性。

A. 参与者类型

BRAIN 的生态系统由多种类型的参与者组成，每种参与者都可以承担多种角色:

用户。 用户是访问 BRAIN 提供的推理服务的个人。他们会花费 ETH 等资产，以

[1]https://github.com/BRAIN-chain

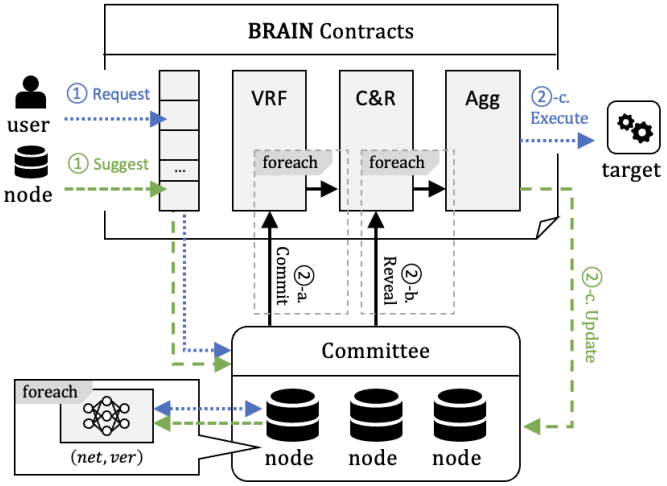


图1：该图展示了 BRAIN 平台中推理（蓝色虚线）和训练（绿色点划线）的统一流程。 ①参与者提交推理或模型更新评估请求，该请求随后被加入队列。 ②-a. 相应委员会（推理委员会或训练委员会）的每位成员提交一项承诺。 ②-b. 每位委员会成员揭示其提交的值（推理输出或模型评分）。 ②-c. 对于推理，任何人都可以使用推理输出执行所请求的操作。对于训练，BRAIN 节点会获得评估分数的中位数，并使用该值在本地计算更新后的模型。

通过交易提交推理请求，并指示如何处理生成的输出。在图1中， ①REQUEST 表示发起推理请求。

BRAIN 节点。节点是存入了足够资产的参与者，因此能够在 BRAIN 生态系统中为推理、训练或两者提供计算资源。存入资产使节点能够积极参与系统，从用户处赚取费用；与此同时，如果发生恶意行为，节点也可能面临智能合约罚没其存款的风险。本文假设所有节点存入的资产数量相同。此假设不失一般性，因为 BRAIN 不限制参与者数量，因此可以灵活地将存款分配到多个节点中。一个 BRAIN 节点可以属于推理委员会、训练委员会、两个委员会，或均不属于。

推理委员会。 推理委员会由通过可验证随机函数选出的 BRAIN 节点组成，负责处理推理请求。每个请求都会选出不同的节点配置。委员会成员因参与推理处理而获得激励。

训练委员会。训练委员会同样由通过可验证随机函数选出的节点组成，负责评估模型更新提案。它们会因参与评估模型更新而获得激励。

为防止不执行任务、只抄袭其他实体结果的搭便车者，BRAIN采用提交-揭示方案，在图1中标记为 ②-a 提交，随后是 ②-b 揭示。揭示后，达成一致的结果将用于人工智能服务的响应，标记为 ②-c 执行；或用于更新模型参数，标记为 ②-c 更新。

如果参与者从事损害协议的恶意或怠慢行为，他们将受到惩罚，以遏制此类负面行为。在 BRAIN 中，惩罚通过罚没已存入的资产来执行。非存款实体的用户也可能从事恶意行为，例如使用大量无意义的推理请求发起拒绝服务攻击[40], [41]。然而，可以通过要求攻击者消耗巨额资产来阻止这些尝试，其原理类似于区块链系统中交易费用的作用。此外，防止区块生产者实施恶意行为是底层区块链协议的责任，而非 BRAIN 的责任。

B. 超参数

在 BRAIN 中，称为超参数的预定义值 p在确保平台实现最佳运行方面发挥着至关重要的作用。

- E 和 d： E 表示可验证随机函数输入发生变化的纪元，而 d 用于调整委员会选举的概率。应适当设置 E，以防止可验证随机函数验证失败或法定人数形成缓慢。调整d 可以最大限度地减少选定节点的数量，从而降低成本，但会牺牲由冗余带来的安全性。推理和训练分别拥有各自的纪元（E_I 和 E_T）以及概率（d_I 和 d_T）。

- f：最终性因子确保可验证随机函数 种子不会因区块链分叉而改变。 f 的单位是轮次，每完成一次 ②-a 阶段，轮次就增加 1。

- Q_C 和 Q_R：它们分别决定了 ②-a 和 ②-b 中法定人数所需的节点数。根据 d调整它们，可以平衡活跃性与性能。

- T_C、 T_R、 T_E和 T_U：超时值控制着 BRAIN 的各个方面。应设置 ②-a 中的 T_C ，以取消不需要的请求，例如计算密集型请求。而 ②-c 中的 T_E 和 T_U 可以设为无穷大，因为每个请求都有自己的 $timeout$ 字段。 ②-b 中的 T_R 用于惩罚未能及时揭示的节点。

- R_C、 R_E、 R_U、 R_R和 R_S：这些符号分别表示成功完成提交、执行、更新、揭示和提议任务所获得的奖励。奖励激励节点执行分配给它们的任务。从本质上说，这些奖励的资金来自用户推理请求费用。建议将 R_C设为 0，以鼓励节点提交后续的揭示。

- P_C、 P_R和 P_S：具体而言， P_C 是对错误提交的惩罚， P_R 是对无效揭示或未揭示的惩罚，P_S 是对错误提议的惩罚。惩罚可防止攻击，并鼓励诚实推理和评估。应合理设置惩罚，以促进积极的资源使用。

本文通篇使用这些符号。特别是，在第七节中，我们将讨论并考察超参数如何影响 BRAIN 的性能及其服务的稳定性。

C. 设计目标与特性

如第 I-A 节所述，集成人工智能与区块链会带来各种挑战，尤其是在处理大模型时。为应对相应挑战，我们设计了若干项功能，简要介绍如下：

通过交易流水线降低延迟。 我们设计了一套能够最大限度降低人工智能相关延迟的系统，以避免对区块链整体性能造成负面影响。为此，我们实现了一个两阶段交易流程，如图1中的 ① 和 ② 所示，将请求交易与响应交易分离。此外， ② 阶段进一步划分为三个子阶段： ②-a、②-b 和 ②-c。这种方法支持流水线处理，从而实现高效的实时处理，减轻区块链的负担。

通过密码学抽签实现可扩展验证。我们实现了一种可扩展的验证机制，在信任需求与人工智能推理和训练的计算成本之间取得平衡。为此，BRAIN 利用基于可验证随机函数的密码学抽签，随机选出一部分节点组成委员会，负责执行推理或训练过程。这种方法通过消除冗余提高了验证效率，同时确保中心化实体无法操纵该过程。

采用无聚合器联邦学习的去中心化训练。 我们的目标是建立一个能够处理大规模模型并防止操纵行为的去中心化训练过程。为此，BRAIN 采用无聚合器联邦学习。由随机选出的节点组成的委员会，即训练委员会，负责对新提交的模型更新进行求值，并就其分数达成共识。这些达成一致的分数会保存于智能合约中，并作为权重用于计算集成模型。每个节点都在本地独立执行这一计算，无需依赖中心化聚合器。这种方法能够实现去中心化且高效的学习，即使存在恶意节点，也能保持系统的稳健性。

与现有区块链的兼容性。 BRAIN 旨在兼容现有的合约感知型区块链，无需进行硬分叉，并能顺利集成到现有的高安全性网络中。

IV. BRAIN：推理

BRAIN 的推理组件旨在去中心化环境中提供信任和安全性。它具有多项功能，包括**两阶段交易**结构、**密码学抽签**以及**承诺-揭示**方案。

A. 两阶段交易

在区块链上实现推理的一种方式是将推理操作作为命令添加进去，如 [11] 所示。这样便可直接在区块链上调用推理，但由于复杂神经网络推理的处理速度比普通交易更慢，甚至慢于区块生成所需的时间，因此可能无法满足实时性要求。

BRAIN协议通过将请求分离，解决了区块链上人工智能任务处理延迟的问题。和响应拆分为两笔独立交易，解决了区块链上人工智能任务处理延迟的问题。第一阶段将推理请求推入队列，第二阶段则将其从队列中取出、执行推理，并通过单独的交易公开结果。这种两阶段配置使推理请求交易能够立即得到处理，让区块生产者无需等待响应即可继续处理区块中的下一笔交易。我们使用通过智能合约实现的优先队列来存放和处理请求交易。请求按优先级顺序加入队列并得到处理，用户可以通过支付更多资产来提高其请求的优先级。这类似于以太坊中的交易费用（燃料费）概念 [47]。

B. 密码学抽签

密码学抽签是 BRAIN协议 的一项关键功能，可确保推理委员会既有效又值得信任。它通过限制执行冗余操作的节点数量来降低成本并提升网络性能。攻击者若想伪造推理结果，就必须占据推理委员会的多数席位，并控制阶段 ② 的整个流程。然而，由于入选委员会成员的身份处于保密状态，这一目标难以实现。

由于 BRAIN 节点已经是区块链系统中的参与者，因此它们拥有自己的 pk 和 sk 密钥。因此，可以在可验证随机函数过程中使用这些密钥，而无需使用密钥生成函数生成新密钥。如图 2 msg所示，可验证随机函数中的验证和求值函数的输入由多个元素组成，用于确保推理结果的可靠性和稳定性。优先级最高的推理请求 q_1 包含在内，以表明它对应于正确的请求。存储在合约中的随机值 $seed_{r-f}$为消息引入随机性，以降低操纵该过程的可能性，例如 [19]。使用过去第 f个种子，而非最新的 $seed_r$，可最大限度地降低 msg 因分叉而发生变化的风险。 $\lfloor h/E_I \rfloor$ 根据预定义的超参数E_I提高 msg 的多样性，从而增加委员会成员，以确保推理过程的活性。

在合约上计算可验证随机函数有两种方式：1）通过合约实现完整的验证逻辑。但由于椭圆曲线运算复杂，这种方法的成本很高。2）使用 EVM 内置的预编译合约 ecrecover。与第一种方法相比，这种方式可显著降低成本。采用这种方法时，安全性会从 32 字节（哈希）降低到 20 字节（地址），但通常已足够。第七节 B 对这两种方法的相关成本进行了全面分析。

C. 承诺-揭示

② 阶段的提交-揭示方案旨在防止搭便车。如果没有这一方案，节点可能会查看其他节点交易中的推理结果，并提交数值相同的交易，而无需

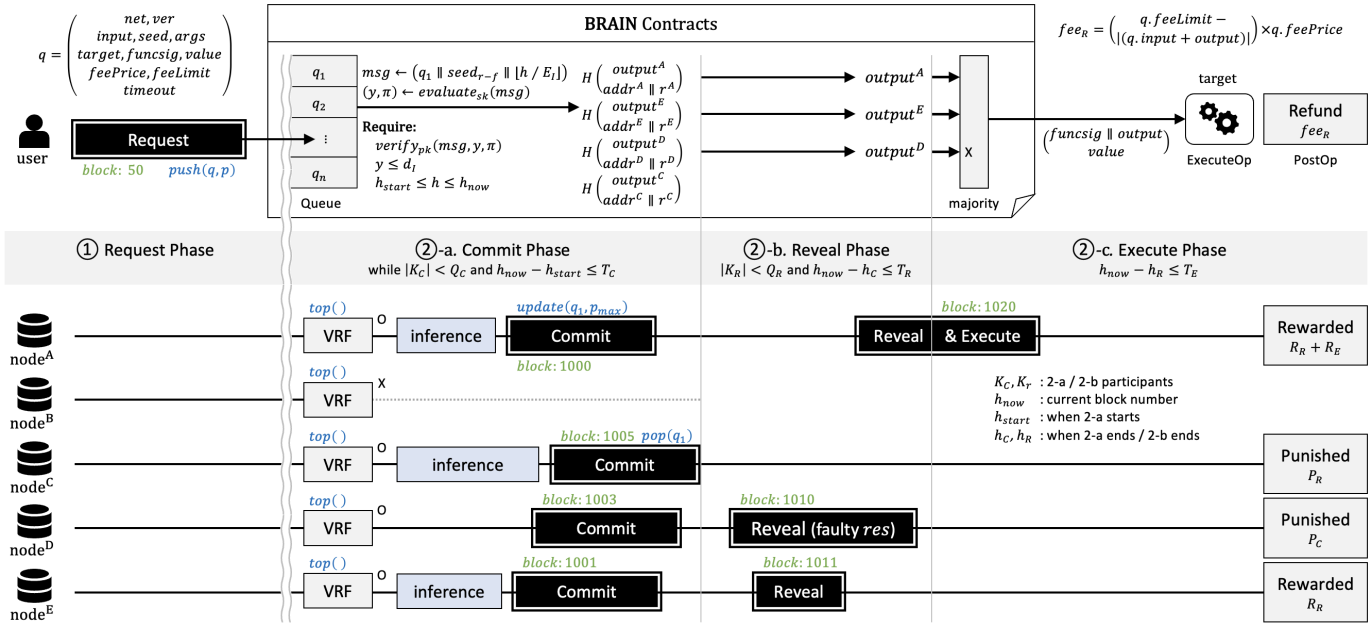


图 2：BRAIN 中推理验证流程概览。用户在阶段 ①发送推理请求。当请求从优先队列中弹出时，阶段 ②开始。随后，推理委员会在阶段 ②-a 提交结果。在阶段 ②-b，原始结果会被揭示。如果达到法定人数，则在阶段 ②-c 通过多数表决得出最终结果，并执行所请求的操作。奖励支付和惩罚在 POSTOP 步骤中进行。此外，合约会将剩余费用退还给用户。图中，双边框的黑色方块表示在区块链上公开记录的交易。绿色方块中的数字表示流程中事件的示例顺序。

实际执行所需的推理操作。BRAIN 中的提交-揭示方案确保带有随机值的结果保持隐藏，因此无法通过查看哈希值获取任何信息。此外，哈希输入包含一个地址，这实际上使其成为其他节点无法使用的承诺。因此，委员会中的节点必须执行所需操作，并在随后的揭示阶段证明自己已提交正确的结果，从而对其行为负责。

D. 推理过程

阶段 ①。如图 2 所示，两阶段交易的阶段 ① 会将推理请求 q 推入优先队列。请求中包含用于标识所请求神经网络模型的唯一标识符（net, ver）。除 $input$外，我们还使用随机 $seed$ 来处理模型的非确定性特征。请求中还会以序列化字节的形式包含其他人工智能相关参数，例如最大长度 和温度，并置于 $args$ 中以协调神经网络。此外，请求还指定了地址 $target$、函数 $funcsig$以及要与推理结果一同调用的原生资产金额 $value$。该请求会存储在队列中，供委员会稍后处理。指定 $timeout$ 是为了防止请求处理过晚。

请求的费用通常根据 $input$ 和 $output$的长度计算，类似于 OpenAI 的定价 [34]。但是，由于无法预先确定 $output$ 的长度，用户需要预先支付足够的费用。为此，系统使用 $feePrice$和 $feeLimit$作为参数采用费用定价机制，类似于 EIP-1559 [9]之前以太坊使用的燃料费系统。如果 $input$ 和 $output$ 的总长度超过指定的 $feeLimit$，请求将被取消。用户预先支付的总金额由 $feeLimit$乘以 $feePrice$确定，但实际支付金额计算为 $(input + output) \times feePrice$，剩余金额将在 ②-c阶段退款。请求的优先级由 $feePrice$确定。

阶段 ②. BRAIN 旨在安全地处理推理请求。它通过基于可验证随机函数（VRF）的加密抽选过程选出推理委员会，并采用提交-揭示方案先登记数值，之后再揭示这些数值。如图 2 所示，提交过程分为②-a 个阶段，而揭示过程分为 ②-b 个阶段。此外，揭示之后使用推理 $output$ 处理用户请求的步骤分为②-c 个执行阶段，其中合约将执行与最后一次揭示合并到单笔交易中。

阶段 ②-a. 节点使用可验证随机函数（VRF）自行验证其是否当选为特定推理请求的推理委员会成员。如果求值结果（256 位值 y）小于或等于难度条件超参数 d_I，该节点便有资格作为推理委员会成员参与。通过合约验证的节点有资格以推理委员会成员的身份注册推理结果。

哈希结果 $output$与地址 $addr$ 和随机值 r通过一笔COMMIT交易注册到合约中。要从 ②-a 进入 ②-b，必须发布超过Q_C 笔COMMIT交易。

阶段 ②-b. 执行揭示先前提交值的过程。为验证揭示的值是否与合约中先前登记的值相符，会将用于构造哈希值的相同组件—— $output$、 $addr$ 和 r ——传递给合约。合约使用这些值再次计算哈希；如果揭示的值与提交的值不同，则揭示交易将被拒绝。当至少有 Q_R 个节点（达到法定人数）揭示其提交的值后，流程即可进入阶段 ②-c。

阶段 ②-c. 由于 BRAIN 没有可信主体，因此最终结果是根据揭示值的多数结果得出的。随后，通过 EXECUTE 交易中的 EXECUTEOP 步骤，在目标地址 $target$ 上使用 $value$ ，执行用户请求的函数 $funcsig$ ，并采用推理 $output$ 达成一致的结果。通常，由于 $seed$ 固定，所有节点的结果都应一致。如果推理委员会中的多数成员是诚实的，就始终能够就真实答案达成一致；揭示违反这一点的结果的节点，则可被视为实施了恶意行为，并通过 POSTOP 步骤受到惩罚。此外，在同一步骤中，剩余费用将退还给用户。

由于阶段②中的子阶段**a**、**b**和**c** 彼此分离，推理处理可以采用流水线化方式进行。最高优先级推理请求 q_1 的阶段 ②-b和 ②- c可与下一个优先级请求 q_2的阶段 ②-a并行执行，从而实现低延迟服务。因此，推理交易的吞吐量仅受 Q_C、 d 和 T_C限制，它们决定了 ②-a阶段的处理速度。

V. BRAIN：训练

BRAIN采用联邦学习，由训练委员会使用**模型评分**对提出的模型进行求值，然后通过**无聚合器联邦学习**方法聚合更新提案，以得出全局模型。此方法允许任何人在任何时候提出更新；由于这些更新会经过验证并随后聚合，模型因此能够持续改进。

A. 模型评分

一种直接的训练验证方法是让当选的委员会节点使用固定种子和相同的数据执行相同的学习过程。虽然这种方法能够确保模型更新的可靠性，但由于训练成本高昂，许多节点会执行重复操作，因此效率低下。

为了解决这些问题，BRAIN 采用了以往研究中使用的技术 [25], [31]。每个节点都会在自己的数据集上评估模型后报告分数，这与 K 折交叉验证类似 [25]。为确保评分公平，提出更新的节点不得参与委员会，即使该节点已通过可验证随机函数获得资格也不例外。

分数通常基于准确率。中位数用于就全局评分达成共识，因为除非多数节点具有恶意，否则这种方法能够提供可靠的结果。BRAIN 的模型评分方法简化了训练委员会的职责，使其侧重于评估模型的性能，而不是验证训练过程。

B. 无聚合器联邦学习

在传统联邦学习中，中心聚合器会计算并共享一个全局模型，该模型是各项更新的加权平均。然而，BRAIN 没有中心聚合器，也没有服务器。相反，每个节点都可以根据合约中存储的分数，使用加权移动平均（WMA）在本地聚合模型。尽管 [38] 使用合约存储权重，但正如第 I-A 节所述，这种方法受高额交易费用限制，且无法支持大模型。无论模型大小如何，BRAIN 都能通过仅存储模型的分数而不存储模型本身，大幅降低上传成本。这种方法与 [44],采用的声誉概念以及 [37], [52]使用的投票机制类似。

算法1 无聚合器联邦学习

要求： 窗口大小 $n \geq 2$，阈值 $score_{th} \leq a_r$
确保： 本地计算的全局模型 M

过程AAFLWMA(n)
  初始化 $r \leftarrow 0, \overline{M_0} \leftarrow M_0, a_0 = 1$
  对于所有新的 $(M_r,\ a_r)$ 以及 $r \geq 1$ **do**
    $\overline{M_r} \leftarrow (1-\alpha)\overline{M_{r-1}} + \alpha M_r$,
    其中

$$\alpha = \begin{cases} \frac{a_r}{a_{r-n+1}+\ldots+a_r} & \text{如果 } r \geq n-1 \\ \frac{a_r}{a_0+\ldots+a_r} & \text{否则} \end{cases}$$

如算法1所述，初始全局模型 M_0被设为 $\overline{M_0}$。对于每个第 r 个全局模型 M_r，节点会计算前 n 个模型（包括当前更新建议 M_r）的加权移动平均，并使用相应的分数 a_{r-n+1}、……、a_r作为权重。由于所有节点都从相同的模型开始，并使用相同的分数整合更新，因此本地得出的全局模型彼此一致，除非某个节点故意遗漏更新或错误使用权重。

将加权移动平均方程代入 FedAsync 的形式 [49],，即可证明该技术的收敛性；FedAsync 已通过数学证明能够收敛。

证明。 由于 $\overline{M_{r-1}} = \frac{a_{r-n+1}M_{r-n+1}+\ldots+a_{r-1}M_{r-1}}{a_{r-n+1}+\ldots+a_{r-1}}$，

$$\begin{aligned}\overline{M_r} &= \frac{(a_{r-n+1}M_{r-n+1} + \ldots + a_{r-1}M_{r-1} + a_rM_r)}{(a_{r-n+1} + \ldots a_r)} \\ &= (1 - \frac{a_r}{a_{r-n+1} + \ldots a_r})\overline{M_{r-1}} + (\frac{a_r}{a_{r-n+1} + \ldots a_r})M_r\end{aligned}$$

取 $\alpha = \frac{a_r}{a_{r-n+1}+\ldots+a_r}$，可得

$$\overline{M_r} = (1-\alpha)\overline{M_{r-1}} + \alpha M_r$$

其收敛性已由 [49] 证明。与$\alpha = \frac{a_r}{a_0+\ldots+a_r}$ 类似，其中 $r < n-1$, 不失一般性。 □

C. 训练过程

阶段 ①. 如图1所示，在训练过程的第一阶段，模型更新 M 通过 SUGGEST 交易发布到区块链，并推送到先进先出队列进行处理。我们使用循环队列来提高存储成本效率，详情见第七节-B。所提议的交易包含一个标识符（net, ver），用于指定带有版本号的目标更新神经网络，以及一个 $timeout$ 字段，用于指示该更新是否过时。 ver 值计算为神经网络权重的哈希。模型更新通过 IPFS [4]等任意网络共享，而非区块链，使每个节点都能通过获取权重的哈希并将其与合约中存储的 ver 进行比较，来验证共享模型。

阶段 ②。 在 BRAIN 训练过程的第二阶段，训练委员会会处理存储在队列中的模型更新提案。此阶段与推理过程类似，但有几个关键区别。首先，在阶段 ②-a 和②-b 中隐藏和揭示的值是 $score$s，而不是推理 $output$s。其次，在阶段 ②- c 中使用中位数，而不是多数。此外，训练过程中不需要 EXECUTEOP 步骤。在 POSTOP 步骤中，执行更新会获得奖励 R_U。如果达成一致的 $score$超过某个阈值，提出模型更新的节点将获得奖励 R_S。另一方面，如果未达到该阈值，该节点将受到惩罚 P_S。

与推理过程类似，训练过程中的评估也可以流水线化。这样可以高效地进行更新，因为第一条更新建议 M_1 的子阶段 ②-b 和第二条建议 M_2 的阶段 ②- a 可以同时启动。

VI. BRAIN：实现细节

本节概述 BRAIN 平台的实现细节，重点介绍推理过程中采用的回退机制。该平台旨在处理各种情形，例如未达到法定人数或发生轮次超时，从而提供一个稳健且实时的去中心化人工智能平台。

在阶段 ②-a 中，如果超过超时时间 T_C 的区块通过后仍未达到法定人数 Q_C ，流程将进入算法2的 **a-回退** ，并取消推理请求 q，将所有资产返还给用户。最后一笔满足法定人数及必要要求的 COMMIT 交易会对队列执行 pop 操作，结束阶段 ②-a 并开始阶段 ②-b。

如果在超时 T_R之前，来自节点的响应数量始终未达到法定人数 Q_R，则执行 **b回退** 。如算法3所示，根据服务预期提供的安全性与活性之间的权衡，预先定义以下两种方法之一：

1) b-I-回退。尽管未达到法定人数，只有已揭示的值进入阶段 ②-c。安全性可能降低，但服务的活性会提高。

2) b-II-回退。流程到此结束。推理请求 q 被设为次高优先级$p_{max}-1$，并重新放入优先队列，以便另一组推理委员会立即处理该请求。在这种情况下，可以提供较高水平的安全性，但活性会有所降低。

算法 2 ②-a. 提交阶段伪代码

要求： $top()$ ≠空值
过程 RESULTCOMMIT(Q_C, T_C, d_I)
  $q_1 \leftarrow top()$
  对于 区块高度 $h \leftarrow h_{start}$, ... **执行**
    对于 $k \in \{$节点$\} - K_C$ 并行 **执行**
      $msg \leftarrow (q_1 \parallel seed_{r-f} \parallel \lfloor h/E_I \rfloor)$
      $(y, \pi) \leftarrow$ 求值$_{sk}(msg)$
      如果 $y \le d_I$ **那么**
        $output^k \leftarrow$ 推理(q_1)
        $H \leftarrow H(output^k \parallel addr^k \parallel r^k)$
        调用 提交$^k_{q_1}(msg, y, \pi, H)$
    如果 $|K_C| \ge Q_C$ **则中断** ▷②-a 的结尾
    否则，如果 $h - h_{start} > T_C$ **则中断** ▷ a-回退
function推理(q)
  return $q.net_{q.ver}(q.input, q.seed, q.args)$
要求： 验证 $_p{}^k(msg, y, \pi)$, $y \le d_I$, $h_{start} \le h \le h_{now}$
交易 提交$^k_q(y, \pi, h, H)$ 将 H^k_q存储至 BRAIN 合约; $K_C \leftarrow K_C \cup \{k\}$
  如果 $|K_C| == 1$ **那么** $update(q, p_{max})$
  如果 $|K_C| \ge Q_C$ **那么** ▷② 的结尾-a
    $pop(q)$; $h_C \leftarrow h_{now}$
    $(seed_r, \pi) \leftarrow$ 求值$_{sk}(seed_{r-1} \parallel r)$
  否则如果 $h_{now} - h_{start} > T_C$ **那么** ▷ a-回退
    $pop(q)$; $K_C \leftarrow \varnothing$
    $seed_r \leftarrow H(seed_{r-1} \parallel r)$
    退款 $(q.feeLimit - |q.input|) \times q.feePrice$
    退款 $q.value$ ETH

算法 3 ②-b. 揭示阶段伪代码

需要： $K_C \neq \varnothing$，阶段 ②-b 回退类型 $\tau \in \{\tau_I, \tau_{II}\}$
过程 RESULTREVEAL$_q$(Q_R, T_R, τ)
对于区块高度 $h \leftarrow h_C$, ...**执行**
  对于 $k \in K_C - K_R$ 并行执行 **执行**
    调用 揭示^k $_q(output^k, addr^k, r^k)$
  如果 $|K_R| \ge Q_R$ **则中断** ▷②-b 的末尾
  否则如果 $h - h_C > T_R$ **则中断** ▷ b回退
要求： H $(output \parallel addr \parallel r)$ $== H^k_q$, $addr == addr^k$
交易 揭示$^k_q(output, addr, r)$
  将 $output$ 存储至 BRAIN 合约; $K_R \leftarrow K_R \cup \{k\}$
  如果 $|K_R| \ge Q_R$ **则调用** 执行 ▷②-b 结束
  else if $h_{now} - h_C > T_R$ **then**
    如果 $\tau == \tau_I$ **则** $h_R \leftarrow h_{now}$ ▷ b-I-fallback
    否则 $push(q, p_{max} - 1)$; $K_C \leftarrow \varnothing$; $K_R \leftarrow \varnothing$
    ▷ b-II-回退

VII. 实验

为了评估 BRAIN 的性能和有效性，我们进行了一系列实验：**推理交易吞吐量**和**合约开销分析**。这些实验的结果揭示了 BRAIN 的性能和可扩展性，并展示了其处理大型神经网络的能力。

A. 推理交易吞吐量

本节将评估 BRAIN 平台提供链上推理的性能，重点关注每秒任务数和超时次数等指标。

本实验考察的关键变量包括：请求交易的频率 $freq$，默认值为 $0.0577 =$（$5752729/99638953$），该值根据 2022 年第二季度以太坊上交易最频繁的服务 OpenSea [35] 与交易总量的比率得出。我们还考虑了推理请求的超时值 $q.timeout$，默认值为 20 个区块，在以太坊上约为 4 分钟。另一个值得关注的变量是当选推理委员会成员的难度级别 d_I。降低 d_I 会降低个人当选的概率。我们使用固定的基础难度 2^{255}，进行实验，这意味着每位成员有 50%（$= 2^{255}/2^{256}$）的当选概率。作为比较，实验中使用的 2^{253} 当选概率为 12.5%（$= 2^{253}/2^{256}$）。我们将 BRAIN 节点总数设为 21，这一设定参考了 EOS 区块链的 BFT-DPoS 共识机制 [14]。这一选择兼顾了去中心化、治理和可扩展性，同时确保网络安全性 [33]。因此，通过阶段 ②-a、Q_C 所需的最少节点数上限为 21。Q_C 的默认值设为 11，约为 21 的一半，且超过 21 的 50%。更高的法定人数值 15 与 EOS 类似，详见第七节-A3。

为确保其他设置不会影响实验性能，我们对其进行了控制。BRAIN平台使用的区块链规范基于以太坊；截至2023年1月13日，区块间隔为12.06秒，每个区块平均包含155笔交易。为了衡量平台性能，我们根据 [24]中报告的EVM交易性能时间，将非推理交易的执行时间量化为1毫秒。用于推选推理委员会的输入消息发生变化的纪元 E_I固定为8个区块。每个阶段 ② 周期的时长均视为无限。在区块链系统中，普遍接受这样一种假设：超过半数的参与节点是诚实的 [32], 。因此，BRAIN的推理和学习系统利用多数值和中位数值，始终能够正确运行。然而，尽管结果的完整性不会受到影响，故意扣留已提交输出的恶意攻击仍可能影响系统带宽。攻击者数量越多，法定人数比率相应提高，这可用于衡量此类影响，并会导致带宽降低、超时增加。相关指标将在第七节 A.3 节中讨论。

在本实验中，我们认为 Q_R 始终等于 Q_C。这是因为我们可以通过提高 Q_C来考虑恶意节点的存在；而本实验中不存在恶意节点，因此假设 $Q_R = Q_C$代表最坏情况，即最低吞吐量。这一假设能够稳健地解释我们的结果，并为系统性能提供保守估计。

在我们的实验中，我们使用了 GPT-J-6B 的 float16 版本 [46]，，这是一个拥有 60 亿参数的 GPT-3 [7] 开源版本。推理过程中未应用任何优化技术。所有用于推理的测试数据集均取自 SAMSum [20]。评估共使用了 819 个数据集。为确定每个请求进入优先队列时的优先级，我们从帕累托分布 [2], [36] 中抽样取值，以模拟真实世界中的 $feePrice$分配。优先级的离散范围为 0 到 1000，数值越大表示优先级越高。

1) 推理响应延迟：为了测量大型神经网络脱离 BRAIN 后的性能，我们在一台配备单张 RTX 3090 显卡、运行 Ubuntu 22.04.1 的系统上进行了实验，测量推理所需的时间。我们使用固定种子进行了 10 次实验，以计算平均耗时。记录到的最短推理时间为 0.0975 秒，最长时间为 50.6394 秒，平均耗时为 18.5421 秒。

根据每个请求的推理时间，我们随后实现了一个模拟器，用于测量 BRAIN 在区块链上的延迟。我们将推理请求相较于一般交易的 $freq$ 设为默认值 0.0577，其他超参数也采用默认值。实验模拟了 100 次并取平均值。结果表明，BRAIN 在收到请求后，至少需要 2 个区块，最多需要 23 个区块，平均需要 7.1653 个区块（$\sigma = 3.5919$）才能收到响应。显然，至少需要 2 个区块，因为提交-揭示方案要求至少包含两笔交易：提交和揭示。它们必须位于不同的区块中，因为每个 BRAIN 节点只有在识别到阶段 ②-a 结束后，才能发送揭示交易。

由于以太坊的区块间隔为 12.06 秒，这对应的最短时间为 24.12 秒，最长时间为 277.38 秒，平均时间为 86.4135 秒。将其与单台计算机上的推理耗时进行比较后，我们发现，使用 BRAIN 可靠地执行人工智能推理所需的时间约为 67.8714 $=$（$86.4135 - 18.5421$）秒。该值受到法定人数 Q_C 的显著影响，不同 Q_C 值下的推理性能可见后续第七节-A3。

2) 每秒任务数：由于 BRAIN 中的推理操作在链下执行，因此推理交易不会产生链上计算延迟。因此，假设交易执行的情况与通常相同，则每秒交易数（TPS）的测量值也相同（$1/0.001 = 1000$）。

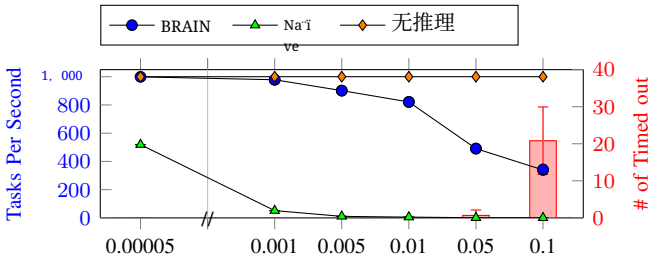


图 3：展示推理请求频率与每秒任务数（折线）以及超时请求数（柱状图）之间关系的图表。

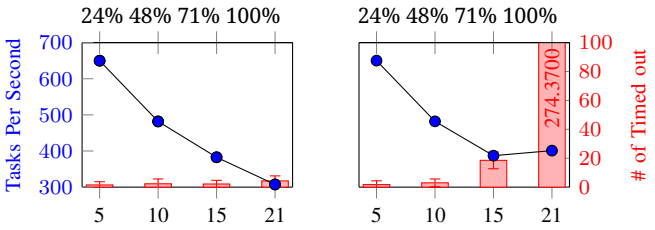
时间需达到 1 毫秒。为了准确求值 BRAIN 的去中心化推理系统的性能，我们定义了每秒任务数 指标。该指标统计每秒完成的任务数量，并将与单个请求相关的所有交易归为一个任务。每个任务可能由多个交易组成，其完成时间取这些交易的总执行时间。与推理无关的交易则分别视为独立任务。

图 3 展示了不同推理请求 $freq$ 下 100 次模拟的平均值和标准差。BRAIN 的每秒任务数性能以蓝色圆圈标记的折线图表示。随着 $freq$ 增加，每秒任务数会减少，因为处理一个推理请求需要将多个交易（包括 REQUEST 以及所有提交和揭示交易）聚合为单个任务。然而，得益于 BRAIN 的流水线处理能力，这种下降并非线性下降，而是有所缓解。

作为对比，我们纳入了朴素实现的结果。该实现只是在链上等待大型神经网络的响应，而不使用两阶段交易结构。如图 3 中绿色三角形标记的折线图所示，朴素实现的每秒任务数明显更低，凸显了 BRAIN 的有效性。当活动级别设为 $freq = 0.0577$ 时，BRAIN 平均每秒可处理 458.0193 ($\sigma = 0.8641$) 个任务。相比之下，朴素实现的平均性能仅为 1.0079。这表明，BRAIN 处理任务的速度是朴素方法的 454.4293 倍。

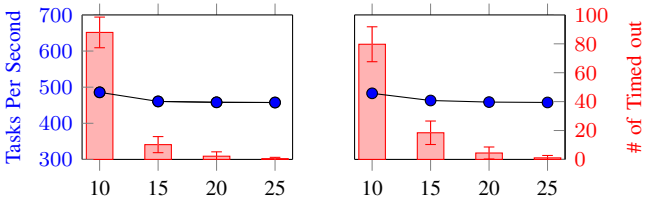
此外，随着推理请求频率的增加，超时请求的数量也相应增加，尤其是在超过阈值$freq = 0.1$之后。然而，通过调整超参数可以缓解这一问题，例如，如图4所示，增大 d_I，或如图5所示，增加每个请求的 $q.timeout$。

3) 不同 Q_C上的每秒任务数： 图 4 展示了每秒任务数和超时次数随法定人数 Q_C（即阶段 ②-a 中的法定人数数量）变化的实验结果。随着单次推理所需节点数量增加，推理委员会产生的总体交易需求也随之增加，从而导致每秒任务数下降。相反，较小的法定人数会带来更高的性能。



(a) Q_C ($d_I = 2^{255}$)　(b) Q_C ($d_I = 2^{253}$)

图4：不同 Q_C 下每秒任务数和超时次数。上方 x 轴表示 Q_C 占节点总数的百分比。(a) 展示可验证随机函数选举概率为 50% 时的结果，(b) 展示概率为 12.5% 时的结果，具体取决于 d_I。



(a) $q.timeout$ ($d_I = 2^{255}$)　(b) $q.timeout$ ($d_I = 2^{253}$)

图5：不同推理请求 $q.timeout$ 下的每秒任务数和超时次数。(a) 显示选举概率为 50% 时的结果，(b) 显示选举概率为 12.5% 时的结果，基于 d_I。

然而，需要注意的是，设置法定人数大小时，平衡安全性与性能之间的权衡至关重要，因为降低法定人数大小虽能提升性能，但也可能威胁推理结果的可靠性。

如图4a所示，增加法定人数会略微增加超时次数，但在 $E_I = 8$ 和 $q.timeout = 20$ 取默认值的环境中，BRAIN 仍可处理大量法定人数。然而，如图4b所示，在难度较高的 2^{253}，中，当法定人数达到 15 时，超时次数会显著增加；达到 21 时，平均超时次数为 274.37，因为在此超参数设置下平台会崩溃。调整难度 d_I 而不是 Q_C 可以减少超时次数，但这会降低计算资源效率并增加单次推理成本。因此，在设置 T_C、E_I、d_I、$q.timeout$ 和 Q_C 等超参数时，必须根据预期的服务激活程度 $freq$，结合所需的性能、安全性和超时容忍度，确定 BRAIN 能够处理的适当范围。

4) 不同 $q.timeout$的超时次数：请求中包含的超时值对应于以区块为单位测量的有效期。图5 展示了缩短 $q.timeout$ 会增加更多请求超过其到期时间的可能性，尤其是在该值低于某个阈值时。在默认值环境中，当 $q.timeout \leq 10$ 时，超时情况会变得严重。

与此同时，图5a和图5b分别比较了难度级别为 2^{255} 和 2^{253}，时的实验结果。尽管——

表 II：以太坊和 Polygon 上的 Gas 消耗

方法	min	max	avg	美元（平均值）	
				以太坊	Polygon
可验证随机函数					
验证	1543493	1862450	1643712	$ 32.504	$ 0.077
验证$_{fast}$	106360	352838	150715	$ 2.980	$ 0.007
承诺-揭示：无哈希函数和有哈希函数 H					
提交	44825	62072	45732	$ 0.904	$ 0.002
提交$_H$	44861	44897	44895	$ 0.888	$ 0.002
揭示	27831	796620	87124	$ 1.723	$ 0.004
揭示$_H$	47355	47391	47389	$ 0.937	$ 0.002
（循环）队列和优先队列					
push	51324	68424	51345	$ 1.015	$ 0.002
pop	29013	46113	29034	$ 0.574	$ 0.001
入栈$_{prior}$	84353	137955	91699	$ 1.813	$ 0.004
出栈$_{prior}$	34909	116942	100861	$ 1.995	$ 0.005
①推送$_{prior}$ =				$ 1.813	$ 0.004
②-a. 验证$_{fast}$ + 提交$_H$ +pop$_{prior}$ =				$ 5.863	$ 0.014
②-b. 揭示$_H$ =				$ 0.937	$ 0.002

尽管在法定人数固定的情况下，降低被选为可验证随机函数的概率可能会因满足法定人数所需时间变长而增加超时的可能性，但实验结果显示差异很小。这表明，在为 E、Q_C和 $freq$设置适当参数时，BRAIN能够在可接受的范围内应对d_I 的变化。相反，如图4b所示，当 Q_C 和d_I 的设置不当时，超时次数会大幅增加。

B. 合约开销分析

我们分析了 BRAIN 中为确保真实性、安全性和低延迟而采用的各种技术所产生的开销。具体而言，我们考察了**可验证随机函数**、**承诺-揭示**方案以及**队列**。我们在表 II 中评估了这些技术在以太坊 [47] 和 Polygon [43] 上的相关燃料费消耗及相应成本。这为这些技术所产生的开销提供了定量分析。以 2023年1月13日为准，我们采用以太坊和 Polygon 上分别为 14 gwei（$14 * 10^{-9}$ ETH）和 51.6 gwei（$51.6 * 10^{-9}$ MATIC）的燃料费价格。

关于可验证随机函数（VRF），我们观察到，使用基于 ecrecover 的验证$_{fast}$ 函数可显著降低成本，在以太坊上每次验证的平均成本为 2.98 美元，而完整验证的成本为 32.5 美元。

关于承诺-揭示方案，我们在 SAMSum [20]测试数据集上测量了 819 个实际推理结果的提交和揭示交易成本。我们很容易观察到，提交过程中是否对推理 $output$ 使用哈希函数，燃料费的差异都微不足道。然而，揭示过程中是否使用哈希函数，燃料费则存在明显差异，因为在交易中使用哈希值可以使燃料费趋于均匀。以太坊上与承诺-揭示方案相关的额外开销被认为是切实可行的，约为 1 至 2 美元。

我们测量了 $push$ 和 pop 操作的 Gas 成本，以评估循环队列和优先队列的效率。结果表明，优先队列的燃料费略高于普通队列。不过，总体而言，以太坊上的开销较低，范围为 $0.5 至 $2.0。

VIII. 应用

BRAIN 通过提供安全性、可追溯性和去中心化，为各行各业带来了潜在的应用。

无偏见的生成式人工智能。 生成式人工智能模型因其能够创建逼真的高质量内容而受到广泛关注 [1], [7], [23], [39]。尽管它们具有创造性表达的潜力，但也带来了训练数据中潜在偏见相关的挑战。为解决这一问题，BRAIN 提供了一种去中心化训练方法，使无聚合器联邦学习者能够加入训练，并根据其被认为最值得信任的私有数据集对模型进行求值。

推荐系统。 人工智能驱动的内容推荐被广泛应用，包括相关电影、视频和音乐推荐 [12], [48]。然而，用户往往无法判断这些推荐究竟来自真实的人工智能推理，还是被操纵的广告。借助 BRAIN 上利用区块链可追溯性的推荐系统，用户可以验证所推荐的内容是否源自实际推理。

智能 NFT。 非同质化代币（NFT） [13] 已成为在区块链上表示独特数字资产的流行方式，常用于头像和个人资料图片 [10], [28]。通过集成 BRAIN，NFT 创作者可以将人工智能模型嵌入其数字资产中，从而实现动态动画和聊天等智能功能。BRAIN 的去中心化特性确保这些人工智能驱动的 NFT 不受中心化控制。

IX. 结论

BRAIN 是一个由合约驱动的去中心化平台，旨在支持大规模神经网络的训练与推理，为各种基于人工智能的去中心化服务提供支持。为确保可靠性和安全性，BRAIN 采用了两阶段交易结构和基于可验证随机函数的委员会抽选机制。此外，我们提出了一种无聚合器联邦学习机制，无需依赖中心化服务器，从而提高成本效益。

我们的实验展示了这些特性在每秒任务数性能方面的有效性，与朴素实现相比，提升了 454.4293 倍。通过微调超参数，我们成功地在性能、安全性和超时容忍度之间取得平衡，同时有效控制超时次数和每秒任务数。此外，我们还证明了 BRAIN 带来的额外 Gas 成本开销很低。

请注意，目前 BRAIN 不包含隐私保护特性，而这也不是本文的研究重点。在未来的研究中，我们计划探索增强 BRAIN平台隐私性的方法。

致谢

本研究部分得到了 Kakao Brain 公司的资助。

参考文献

[1] Andrea Agostinelli、Timo I Denk、Zal´an Borsos、Jesse Engel、Mauro Verzetti、Antoine Caillon、Qingqing Huang、Aren Jansen、Adam Roberts、Marco Tagliasacchi 等人。MusicLM：从文本生成音乐。arXiv 预印本 arXiv:2301.11325，2023年。 [2] Barry C Arnold。帕累托分布。Wiley StatsRef：在线统计参考资料，第 1–10 页，2014年。 [3] Xianglin Bao、Cheng Su、Yan Xiong、Wenchao Huang 和 Yifei Hu。FLChain：一种支持可审计联邦学习并兼顾信任与激励的区块链。发表于 2019年第五届大数据计算与通信国际会议（BIGCOM），第 151–159 页，2019年。 [4] Juan Benet。IPFS——一种内容寻址、版本化的点对点文件系统。arXiv 预印本 arXiv:1407.3561，2014年。 [5] Battista Biggio、Blaine Nelson 和 Pavel Laskov。针对支持向量机的投毒攻击。arXiv 预印本 arXiv:1206.6389，2012年。 [6] Adrienne Brackey。《Northpointe 的 COMPAS 算法中的种族偏见分析》。博士论文，杜兰大学科学与工程学院，2019年。[7] Tom Brown、Benjamin Mann、Nick Ryder、Melanie Subbiah、Jared D Kaplan、Prafulla Dhariwal、Arvind Neelakantan、Pranav Shyam、Girish Sastry、Amanda Askell 等人。语言模型是少样本学习者。《神经信息处理系统进展》，33 1901，2020年 1877：–。[8] Joy Buolamwini 和 Timnit Gebru。Gender Shades：商业性别分类中的交叉性准确率差异。载于 Sorelle A. Friedler 和 Christo Wilson 编，《第一届公平性、问责性与透明度会议论文集》，机器学习研究论文集第 81 卷，第 77–91 页。PMLR，2018年2月23–24日。[9] Vitalik Buterin、Eric Conner、Rick Dudley、Matthew Slipper、Ian Norden 和 Abdelhamid Bakhta。EIP-1559：ETH 1.0 链的费用市场变更。
https://eips.ethereum.org/EIPS/eip-1559，2019年。 [10] Simone Casale-Brunet、Mirko Zichichi、Lee Hutchinson、Marco Mattavelli 和 Stefano Ferretti。NFT 头像对社交网络社区的影响。发表于 2022年 ACM 社会公益信息技术会议论文集，第 283–291 页，2022年。 [11] Ziqi Chen、Weiyang Wang、Xiao Yan 和 Jia Tian。区块链上的 Cortex-AI。Cortex Labs Pte. Ltd.，新加坡，技术报告 C，201803307:2018，2018年。 [12] Heng-Tze Cheng、Levent Koc、Jeremiah Harmsen、Tal Shaked、Tushar Chandra、Hrishi Aradhye、Glen Anderson、Greg Corrado、Wei Chai、Mustafa Ispir 等人。用于推荐系统的宽度与深度学习。发表于第一届推荐系统深度学习研讨会论文集，第 7–10 页，2016年。 [13] William Entriken、Dieter Shirley、Jacob Evans 和 Nastassia Sachs。ERC-721：非同质化代币标准。https://eips.ethereum.org/EIPS/eip-721，2018年。 [14] EOSIO。《EOS.IO 技术白皮书 v2》。
h
t
t
p
s://
github.co
m/EOSIO/Documentation/blob/master/TechnicalWhitePaper.md。[15] Yongkai Fan、Binyuan Xu、Linlin Zhang、Jinbao Song、Albert Zomaya 和 Kuan-Ching Li。基于零知识证明验证卷积神经网络预测的完整性。《信息科学》，2023年。[16] Yongkai Fan、Guanqun Zhao、Xia Lei、Wei Liang、Kuan-Ching Li、Kim-Kwang Raymond Choo 和 Chunsheng Zhu。SBBS：一种用于雾环境中物联网数据可信度的安全区块链方案。《IEEE 物联网期刊》，8(11):9268–9277，2021年。[17] Boyuan Feng、Lianke Qin、Zhenfei Zhang、Yufei Ding 和 Shumo Chu。ZEN：用于可验证零知识神经网络推理的优化编译器。密码学电子预印本档案，2021年。 [18] DeepBrain Chain Foundation。《DeepBrain Chain 白皮书》。
https://www.deepbrainchain.org/assets/pdf/DeepBrainChainWhite_paper en.pdf，2017年。[19] Yossi Gilad 等人。Algorand：为加密货币扩展拜占庭协议。发表于第二十六届操作系统原理研讨会论文集，第 51–68 页，2017年。[20] Bogdan Gliwa、Iwona Mochol、Maciej Biesek 和 Aleksander Wawer。SAMSum 语料库：用于抽象式摘要的人工标注对话数据集。arXiv 预印本 arXiv:1911.12237，2019年。[21] Jonathan Heiss、Elias Gr¨unewald、Stefan Tai、Nikolas Haimerl 和 Stefan Schulte。通过可验证链下计算推进基于区块链的联邦学习。发表于 2022年 IEEE 国际区块链会议（Blockchain），第 194–201 页。IEEE，2022 年。

[22] Andrew G Howard、Menglong Zhu、Bo Chen、Dmitry Kalenichenko、Weijun Wang、Tobias Weyand、Marco Andreetto 和 Hartwig Adam。MobileNets：面向移动视觉应用的高效卷积神经网络。arXiv 预印本 arXiv:1704.04861，2017年。[23] Midjourney Inc. Midjourney。https://www.midjourney.com，2022年。[24] Jae-Yun Kim、Junmo Lee、Yeonjae Koo、Sanghyeon Park 和 Soo-Mook Moon。Ethanos：基于账户的区块链全节点的高效自举。载于第十六届欧洲计算机系统会议论文集，第99–113页，2021年。[25] Yuzheng Li 等。具有委员会共识的基于区块链的去中心化联邦学习框架。IEEE Network，35 241，2021 1 234( ): – 。[26] Wei Liang、Yongkai Fan、Kuan-Ching Li、Dafang Zhang 和 Jean-Luc Gaudiot。工业区块链网络环境中的安全数据存储与恢复。IEEE Transactions on Industrial Informatics，16(10):6543–6552，2020年。[27] Yuan Liu、Zhengpeng Ai、Shuai Sun、Shuangfeng Zhang、Zelei Liu 和 Han Yu。Fedcoin：用于联邦学习的点对点支付系统。载于《联邦学习》，第125–138页。Springer，2020年。[28] Yuga Labs LLC。无聊猿游艇俱乐部。
https://boredapeyachtclub.com/ #/.[29] Brendan McMahan、Eider Moore、Daniel Ramage、Seth Hampson 和 Blaise Aguera y Arcas。基于去中心化数据进行深度网络通信高效学习。载于《人工智能与统计学》，第1273–1282页。PMLR，2017年。[30] Silvio Micali、Michael Rabin 和 Salil Vadhan。可验证随机函数。载于第40届计算机科学基础年会（第99CB37039号），第120–130页。IEEE，1999年。[31] Vaikkunth Mugunthan、Ravi Rahman 和 Lalana Kagal。Blockflow：去中心化、保护隐私且可问责的联邦机器学习。载于 Javier Prieto、Alberto Partida、Paulo Leit˜ao 和 Ant´onio Pinto（编），《区块链与应用》，第233–242 页，Cham，2022年。Springer International Publishing。[32] 中本聪。比特币：一种点对点电子现金系统。Decentralized business review，第21260页，2008年。[33] Ivan on Tech。访谈 Dan Larimer——EOS、Cardano、以太坊、委托权益证明、Steemit、加密货币泡沫、加密货币的未来。https://youtu.be/ o7HQlcl-LlQ?t=33m17s。[34] OpenAI。定价。
https://openai.com/pricing。[35] Inc Ozone Networks。OpenSea。
https://opensea.io/。[36] 维尔弗雷多·帕累托。《政治经济学教程》，第2版，Rouge，洛桑，2:305，1897年。[37] Zhen Qin、Xueqiang Yan、Mengchu Zhou、Peng Zhao 和 Shuiguang Deng。Blockdfl：基于区块链的完全去中心化联邦学习框架，2023年。[38] Paritosh Ramanan 和 Kiyoshi Nakayama。Baffle：基于区块链的无聚合器联邦学习。载于2020年IEEE国际区块链会议（Blockchain），第72–81页，2020年。[39] Robin Rombach、Andreas Blattmann、Dominik Lorenz、Patrick Esser 和 Bj¨orn Ommer。使用潜在扩散模型进行高分辨率图像合成，2022年。[40] Muhammad Saad、Laurent Njilla、Charles Kamhoua、Joongheon Kim、DaeHun Nyang 和 Aziz Mohaisen。通过内存池优化防御基于工作量证明的区块链系统中的分布式拒绝服务攻击。载于2019年IEEE国际区块链与加密货币会议（ICBC），第285–292页。IEEE，2019年。[41] Muhammad Saad、My T Thai 和 Aziz Mohaisen。海报：通过内存池优化遏制针对基于区块链的加密货币的分布式拒绝服务攻击。载于2018年亚洲计算机与通信安全会议论文集，第809–811页，2018年。[42] Zonyin Shae 和 Jeffrey Tsai。用于可信新闻的人工智能区块链平台。载于2019年IEEE第39届国际分布式计算系统会议（ICDCS），第1610–1619页，2019年。[43] Polygon technology。Polygon。https://polygon.technology/。[44] Yongding Tian、Zhuoran Guo、Jiaxuan Zhang 和 Zaid Al-Ars。DFL：高性能区块链联邦学习。arXiv 预印本 arXiv:2110.15457，2021年。[45] Ruben Tolosana、Ruben Vera-Rodriguez、Julian Fierrez、Aythami Morales 和 Javier Ortega-Garcia。深度伪造及其他：人脸篡改与伪造检测综述。Information Fusion，64:131–148，2020年。[46] Ben Wang 和 Aran Komatsuzaki。GPT-J-6B：一个拥有60亿参数的自回归语言模型。https://github.com/kingoflolz/ mesh-transformer-jax，2021年5月。

[47] Gavin Wood 等。以太坊：一种安全的去中心化通用交易账本。以太坊项目黄皮书，2014年。[48] Shiwen Wu、Fei Sun、Wentao Zhang、Xu Xie 和 Bin Cui。推荐系统中的图神经网络：综述。*ACM Computing Surveys*，55(5):1–37，2022年。[49] Cong Xie、Sanmi Koyejo 和 Indranil Gupta。异步联邦优化。*arXiv* 预印本 *arXiv:1903.03934*，2019年。[50] Zhibo Xing、Zijian Zhang、Meng Li、Jiamou Liu、Liehuang Zhu、Giovanni Russello 和 Muhammad Rizwan Asghar。基于零知识证明的区块链实用联邦学习。*arXiv* 预印本 *arXiv:2304.05590*，2023年。[51] Chaofei Yang、Qing Wu、Hai Li 和 Yiran Chen。针对神经网络的生成式投毒攻击方法。*arXiv* 预印本 *arXiv:1703.01340*，2017年。[52] Haoxiang Yu、Hsiao-Yuan Chen、Sangsu Lee、Sriram Vishwanath、Xi Zheng 和 Christine Julien。iDML：激励式去中心化机器学习。*arXiv* 预印本 *arXiv:2304.05354*，2023年。[53] Xichen Zhang 和 Ali A Ghorbani。在线虚假新闻概述：特征、检测与讨论。*Information Processing & Management*，57(2):102025，2020年。