Network3 Economic Model Whitepaper

Network3

August 22, 2024

Abstract

In the data-driven digital era, the importance of data has become increasingly prominent, and the rapid development in the field of Artificial Intelligence (AI) in particular marks the dawn of a new era. Since the introduction of generative AI models such as GPT-3, they have led the explosive growth of AI with amazing capabilities and wide application prospects. However, this growth has also brought with it significant challenges, especially in terms of arithmetic requirements and costs. The Network3 project was born out of this backdrop, aiming to solve the arithmetic challenges in the AI industry and the wider field through a decentralized approach. This whitepaper details Network3's token economy model, which aims to build a decentralized, efficient, and sustainable distributed machine learning ecosystem.

1 Introduction

1.1 Summary

Network3 is a Depin platform focused on privacy security and efficient distributed computing, providing data transfer, arithmetic sharing, and security assurance services for distributed AI. Network3 integrates an efficient anonymous certificateless signcryption (CLSC) algorithm, a data correctness verification mechanism, an IP anti-tracking measure, and a decentralized reliable federated learning (FL) framework. Network3 already has a mature vpn product with millions of users. network3 will build a bandwidth sharing network based on the existing vpn product, and then build a distributed arithmetic sharing platform on the mature decentralized network. network3's economic model is designed to validate the contribution of the participants in the system and to distribute the rewards.

1.2 Market Overview and Challenges

1.2.1 Arithmetic Power Demand for AI

As AI technology advances, especially driven by generative AI models such as GPT-3, the demand for arithmetic power is growing exponentially.For example, the training cost of GPT-3 may be as high as $12 million, far exceeding that of previous-generation models.This surge in cost is not limited to the training phase, but also includes the actual application phase of the model's reasoning.For example, according to data from earlier this year, the required arithmetic demand at 13 million unique users is a staggering 30,000+ A100 GPUs, with an initial investment cost of 800*millionandanestimated*700,000 in daily model inference costs.

1.2.2 Monopolization of the arithmetic power economy

Currently, the arithmetic power economy in AI is monopolized by a handful of giants.For example, NVIDIA dominates the AI GPU market, and its products are expensive and usually snapped up by large companies in Silicon Valley. In addition, cloud computing platforms such as AWS and Microsoft Azure hold a large amount of server and GPU resources.This monopoly not only limits the ability of small businesses and independent developers to access arithmetic resources, but also results in high utilization costs.

1.2.3 The Potential of a Decentralized Arithmetic Market

Against this backdrop, a decentralized arithmetic marketplace is particularly important.This marketplace aims to create an open arithmetic marketplace that enables anyone with idle arithmetic resources to offer their resources on this marketplace through a token incentive mechanism, thus meeting the growing demand for arithmetic.This model not only provides services to the B-end user and developer community, but also helps to break the existing arithmetic monopoly, reduce costs, and improve resource utilization efficiency.

1.3 Decentralized Bandwidth and Arithmetic Market

In the foreseeable future where AI is widely popularized and applied, model training requires a lot of data and arithmetic power. Current model training requires huge costs and has a significant lag. In the face of the huge demand for data and computing power, the vast majority of AI companies do not have the ability to do so, so the underlying infrastructure of AI applications is actually monopolized by a few companies. Such a future is unreasonable. Human data and the resulting models should not be in the hands of a few companies. There are also some arithmetic shortages due to national competition and legal policies that limit the growth of most AI companies. Therefore mechanisms to utilize unused bandwidth and arithmetic power are essential. Decentralized bandwidth and arithmetic market is a typical Depin network that utilizes the restricted resources of edge nodes to serve application scenarios that require resources such as data, routing, arithmetic, etc., most typically decentralized AI. in this market, any individual or organization with idle resources can provide their own resources to serve the B-end users and developer community. This model has the potential to revolutionize the way existing data and arithmetic power is distributed and used, break the computing monopoly of traditional companies, and provide a fairer and more cost-effective way to utilize data and arithmetic resources while protecting privacy.

1.4 Network3's Vision and Strategy

Against this market backdrop, Network3 aims to solve the arithmetic challenges in the AI and blockchain industries through the establishment of a decentralized arithmetic marketplace.Network3 plans to utilize a token economy model to incentivize individuals and organizations globally to share their arithmetic resources.This model is not only expected to reduce the cost of arithmetic resources and improve efficiency, but also provide a new revenue stream and business model for participants in the decentralized network.

Our vision is to build a fairer, more efficient and sustainable distributed computing ecosystem through innovative technology and economic models.Network3 aims to break the traditional monopoly of arithmetic power and promote the fair distribution and efficient utilization of arithmetic resources, thereby promoting the wider application and development of AI and blockchain technologies.

2 Network3's Solution

Network3 revolves around key technologies, including an efficient anonymous certificateless signcryption (CLSC) algo- rithm, a decentralized reputation mechanism, and an IP antitracking measure. The anonymous CLSC algorithm offers a unique blend of identity authentication and secure data sharing under anonymous conditions, backed by thorough security and performance analysis. The decentralized data correctness verification mechanism, infused with homomorphic encryp- tion, secret sharing, and Reed-Solomon coding, provides a solution to potential inaccuracies in received data. The IP antitracking measure ensures a fully anonymous data transmission experience.

Based on these underlying secure anonymous data transfer protocols, Network3 will build decentralized bandwidth and arithmetic markets and design a matching contribution verification scheme. For the arithmetic market, we build a decentralized AI system based on federated learning implementation to complete the decentralized model training process and provide a decentralized scheme for the model inference process afterwards.

The main designs include:

• Firstly, we propose a highly efficient anonymous CLSC algorithm for authenticated transmission and offer a detailed algorithm construction with security and perfor- mance analysis, underscoring the algorithm's effectiveness and practicality.

• Secondly, we devise a decentralized rating-based data correctness verification mechanism to address the issue of data inaccuracies at the receiving end.

• Thirdly, we design a novel IP antitracking mechanism to achieve completely anonymous data transmission and protect the inherent freedom for the ubiquitous interaction and cooperation.

• Finally, we present a groundbreaking decentralized federated learning framework, thoughtfully designed to tackle the core challenges associated with realizing practical and dependable decentralized AI capabilities.

2.1 Mechanisms for the bandwidth market

For the starting arithmetic market, we have designed a perfect rating scheme. The system evaluates the contribution based on the amount of data transmitted by the participants' nodes. Beyond the initial period, to prevent nodes from committing evil, we design a data correctness verification mechanism and penalize the evil doers with a fraud proof-like scheme. Based on the correctness verification mechanism, we constructed a rating system that scores and rates the reputation of the nodes, and the reputation score is directly linked to the rewards received by the nodes.

2.2 Mechanisms for the arithmetic power market

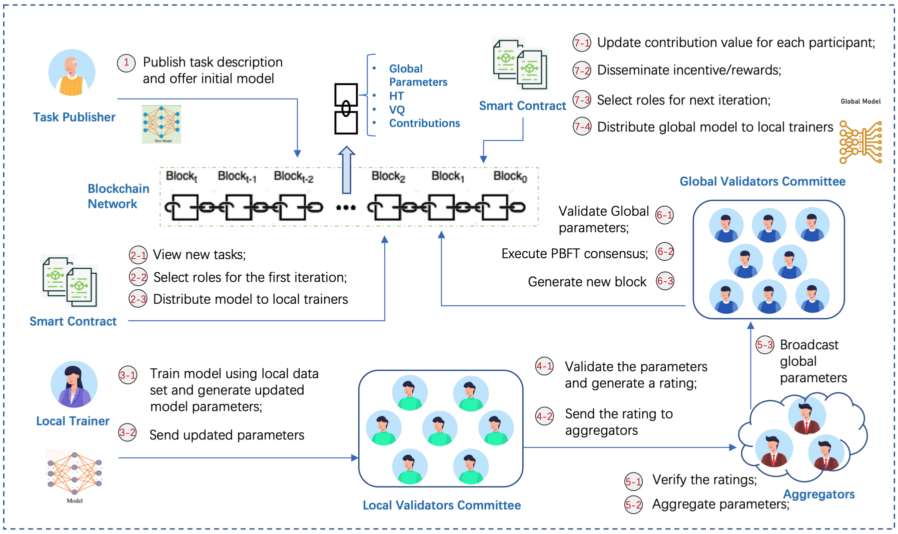
Task Publisher
1 Publish task description and offer initial model
Global Parameters
HT
VQ
Contributions
Smart Contract
7-1 Update contribution value for each participant;
7-2 Disseminate incentive/rewards;
7-3 Select roles for next iteration;
7-4 Distribute global model to local trainers
Global Model
Blockchain Network
Block$_t$
Block$_{t-1}$
Block$_{t-2}$
Block$_2$
Block$_1$
Block$_0$
Global Validators Committee
Validate Global parameters; 6-1
Execute PBFT consensus; 6-2
Generate new block 6-3
2-1 View new tasks;
2-2 Select roles for the first iteration;
2-3 Distribute model to local trainers
Smart Contract
Local Trainer
3-1 Train model using local data set and generate updated model parameters;
3-2 Send updated parameters
Model
Local Validators Committee
4-1 Validate the parameters and generate a rating;
4-2 Send the rating to aggregators
5-3 Broadcast global parameters
5-1 Verify the ratings;
5-2 Aggregate parameters;
Aggregators

Figure 1: Processes of the proposed decentralized AI framework

We have completed the design of a complete set of decentralized AI systems, including mechanisms for data annotation and local training of training nodes, model aggregation, validation of the training process, model consensus and contribution validation. For the specific process, please check Network3's whitepaper. The basic process is that the task publisher releases the model training task on the platform, the aggregation node receives the task and starts to do task analysis and scheduling, the training node receives the task and starts to do local training, and directly uses the local data for data annotation and pre-training, and the more complex tasks may use the data from other data-providing nodes, and during training, the training results will be released to the validation node for verification, which is set up with an anti Collusion attack data validation mechanism is set up, the training results are aggregated in the aggregation node, and the results are transmitted to the global validator for validation, the validation is completed to reach a consensus on the training results can be delivered to the task publisher, or directly open to the community. The task publisher can be the actual company or individual, or the community grants.

3 Network3 Token ($N3) Economic Modeling

3.1 Token Issuance and Distribution

Network3's tokens ($N3) are central to the flow of value in its ecosystem and the operation of the network.Tokens can be acquired in the following ways:

3.1.1 Participation in training models

Participation in training models is one of the main ways in which $N3 tokens are acquired.In this process, participants support the training of AI models by contributing their computational resources (CPU/GPU, etc.).This involves not only the direct contribution of arithmetic power, but also activities such as optimizing the model and improving training efficiency.The greater the contribution, the greater the $N3 token reward.This approach encourages and rewards individuals or groups who actively participate in AI model training and optimization.

3.1.2 Shared resources (arithmetic, data, bandwidth)

Sharing resources is another important way to acquire tokens.Users can participate in the Network3 ecosystem by sharing their computing power, data, or bandwidth resources.For example, users can offer idle computing power (e.g., a personal computer's CPU/GPU) to the network, or provide high-quality datasets to support more accurate AI model training. In addition, the sharing of bandwidth is crucial to maintain the efficient operation of the network. In this way, resource providers are able to earn $N3 as a reward while contributing to the development of the Network3 ecosystem.

3.1.3 Participate in Governance

Participating in the governance of the Network3 network is also a way to earn $N3 tokens.Token holders can participate in the decision-making process, such as voting on updates to network protocols and adjustments to reward policies.This participation not only enhances the democracy and transparency of the network, but also allows holders to directly influence the direction of Network3.Governance participants are rewarded according to their level of contribution to the network's decision making are rewarded with $N3 tokens.This approach is designed to incentivize and empower members who actively participate in Network3's governance and decision-making process, ensuring that the network evolves for the common good of the community.

3.1.4 Token Release Allocation

Reflecting Network3's emphasis on active contributors to the ecosystem, 90% of the token release will be allocated to activities that involve training models and sharing resources.This portion of the token allocation is intended to maximize incentives for activities that directly contribute to the growth and effectiveness of the network, such as model training and resource sharing.The remaining 10% of tokens are allocated to the team that operates and maintains the Network3 network as a reward for their work in building and maintaining the network.This includes contributions to development, technical support, and network maintenance to ensure the stable operation and continued growth of the Network3 platform.

3.2 Token circulation mechanism

Network3's economic cycle is the dynamic process in which money flows, tokens are distributed, and incentives and penalties are cycled throughout the token system.This cycle ensures the vitality of the network, incentivizes user participation, and maintains the health of the network.

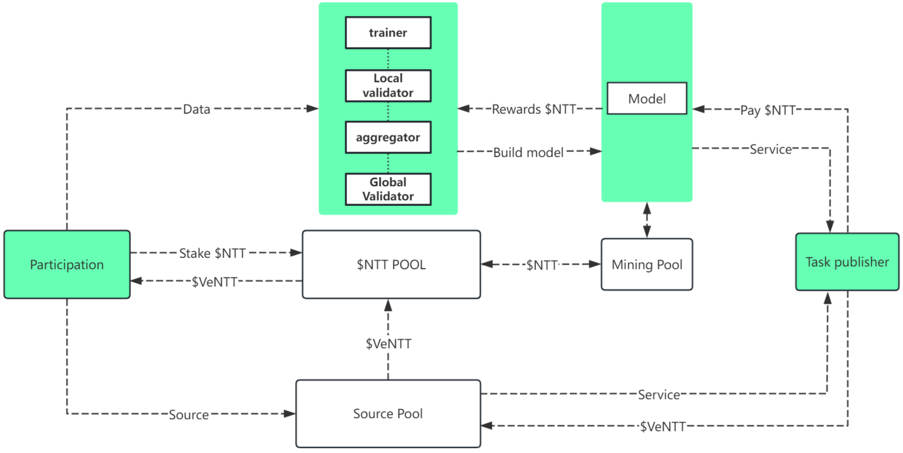

Figure 2: Network3's economic cycle

$N3 is Network3 Network's incentive and governance token. It has the following main functions primary functions: · Initial Token Distribution The Initial Token Distribution provides start-up capital for the network, including incentives for the development team, early investors and the Foundation, as well as funds raised through public sales.

· Pledging and Participation Participants participate in network activities, including model training, parameter validation, and governance voting, by pledging $N3 tokens.The act of pledging increases the participant's voting power in the network and potential network revenue sharing.

· Token Reward Disbursement Network event participants are rewarded with $N3 tokens based on their contributions, such as data provision, arithmetic contributions, model validation, and governance decisions.

· Token Reinvestment and Circulation Rewarded tokens can be used for further pledges, increasing voting power and incentive earnings for participants.Tokens can also be circulated on the market to pay for services, purchase data or computing resources.

· Consumption and Network Fees Network operations such as model training and validation consume $N3 as a cost of computing and storage resources.Some tokens are withdrawn from circulation through network fees such as transaction fees and service fees (reflows) to support development and maintenance funding.

· Network Inflation and New Token Issuance In order to maintain incentives for network participants and to support the addition of new users and expansion of the network, new \$N3 tokens can be issued through a controlled inflation strategy.

· Long-Term Growth and Value Capture As the network grows and expands, the value of \$N3 tokens is expected to reflect the overall growth of the Network3 platform.The value capture is reflected in the market value of the tokens, the utilization rate, and the practical application value of the network.

3.3 $N3 Staking and $veN3 Mechanism

3.3.1 Basic Staking Rewards

Holders of \$N3 are entitled to a basic Annual Percentage Rate (APR) by staking their tokens. The foundation of this APR is derived from 75% of the network's Gas fee revenue. This incentivizes users to participate in the network's security and consensus mechanisms.

Revenue Allocation:

· 75% of the Gas fee income is allocated to \$N3 stakers, rewarding them for their contribution to the network's liquidity and stability.

· The remaining 25% is allocated to developers who actively contribute to the Network3 ecosystem.

Before any rewards are distributed, operational costs are deducted from the total revenue. These costs include, but are not limited to, resource transfer fees, node operation costs, and infrastructure provider fees.

3.3.2 $veN3 Acquisition Through $N3 Staking

Stakers can boost their APR by locking their \$N3 to obtain \$veN3. The \$veN3 acquired is proportional to the duration of the stake, encouraging longer-term investments within the network.

Lock-In Periods and Corresponding APRs:

· Stakeholders can choose to lock their \$N3 for periods ranging from 7 days up to a maximum of 365 days.

· The lock-in period corresponds to a multiplier effect on the APR, rewarding users for longer commitment. The Boost factor increases linearly with the lock-in duration, with a maximum boost of 1.75x for a 365-day lock-in.

3.3.3 Rights and Rewards of $veN3 Holders

Holding \$veN3 grants users periodic rewards released from the network and increases their stake in governance decisions. The amount of \$veN3 held not only influences the holder's voting power but also correlates with the distribution of network profits.

Governance Participation:

· \$veN3 holders are vested with voting rights to participate in pivotal decisions regarding the network's future, such as upgrades and fee structure adjustments.

Profit Distribution:

· Profits are distributed after operating costs are deducted, with 25% automatically allocated to active developers and the remaining 75% to the reward contract for stakers.

3.3.4 Reward Distribution and $veN3 Unlocking Mechanism

Gas Fee Reward Distribution:

· Gas fee rewards are regularly distributed to $veN3 holders, aligning their interests with the network's profitability.

Weekly Settlement:

· The $N3 protocol settles Gas fee revenue weekly, distributing profits after operational costs have been accounted for.

Unlocking Mechanism:

· Upon the conclusion of the staking period, $N3 is automatically unlocked and the corresponding $veN34 expires, resetting the APR to the base level.

Early Unstake Penalties:

· Early unstaking results in a loss of $veN3 balance, reducing future rewards and voting power. The penalty is calculated based on the remaining lock-in time.

3.4 Token Distribution

· Seed Sale 3%
· Private Sale 2%
· Public Sale 1%
· Mining 40%
· Computing Power 35%
· Airdrop 2%
· Liquidity Provision 3%
· Ecosystem Dev 2%
· Team 12%

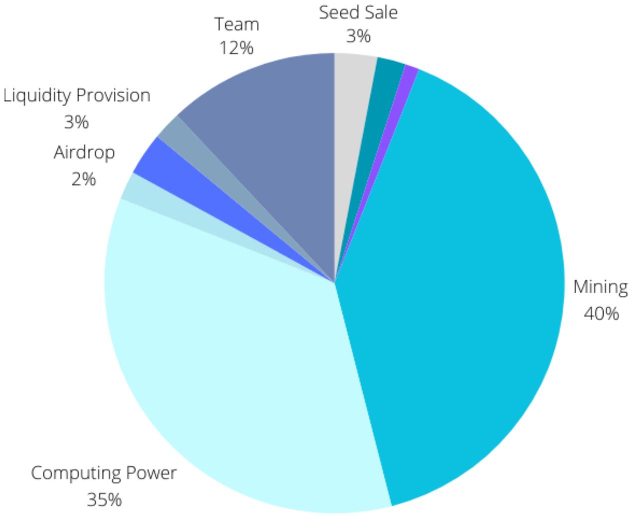

Figure 3: Token Distribution

4 Governance structure

$N3 token holders participate in the network's governance decisions, including protocol updates, reward policy adjustments, and so on.This decentralized governance structure ensures the democratic and transparent nature of the network. Governance decisions are made through a weighted voting mechanism where $N3 holdings determine the voting weights.This encourages long-term holdings and active participation in network governance.

4.1 Actors

Local trainners: Responsible for data annotation and local training using local datasets or aggregated datasets;

Local Validators: Responsible for evaluating the model parameters uploaded by the trainers;

Aggregators: Responsible for task scheduling and aggregating model parameters;

Global Validators: Responsible for verifying the validity of the global model and reaching a global consensus, packaging and uploading the transaction process to Layer1

4.2 Node establishment and its responsibilities

Super node (global node): The threshold for the establishment of super nodes, the highest decision-making level in the network, is quite high, requiring a pledge equal to 0.5% of the total circulation of the network. This requirement not only shows a long-term commitment to the network, but also reflects how powerful these nodes are in terms of resources and capabilities. The main responsibility of supernodes is to make key decisions, such as protocol changes and major policy adjustments. In terms of network governance and oversight, SuperNodes play a leadership role in ensuring the overall health and sustainability of the Network3 ecosystem. Supernodes have decisive power in voting on key issues, and their decisions have far-reaching effects on changes in the overall network policy

Tier 1 Node:

Tier 1 nodes, or alternative nodes, are the middle tier in the network governance structure. In order to become a Level 1 node, participants need to pledge 50% of the amount of a super node. This threshold is intended to attract participants who have an in-depth understanding of how the network works. Tier 1 nodes play a key role in executing the decisions of the super nodes, as well as guaranteeing the smooth operation of the network's key functions. These nodes may also be granted the ability to make critical decisions under certain circumstances, especially if the supernodes are unable to respond instantly.

Ordinary nodes: Ordinary nodes are the basic layer of Network3's governance structure, with a threshold of 50% of the level 1 node, designed to allow more community members to participate in the network's operations. The responsibilities of ordinary nodes include participation in daily network maintenance and basic governance tasks, such as data validation and security maintenance. Despite their basic position in the governance hierarchy, common nodes play a crucial role in maintaining the stability and security of the network.

Through the extensive participation of ordinary nodes, Network3 is able to ensure the decentralized nature of the network and a high degree of community participation.

4.3 Node Staking and Roles

In Network3, different types of nodes participate in the network by pledging $N3 tokens, and each type of node plays a unique role and corresponds to a specific pledging mechanism.

4.3.1 Trainers

Trainers are key participants in the Network3 platform, responsible for contributing to the development and optimization of machine learning models.

Pledge mechanism for trainers:

·In order to become a trainer, participants need to pledge $N3 tokens, which serves as a guarantee of their participation in the contribution.

·The amount of pledge is proportional to the computational resources (e.g., CPU/GPU) provided by the participant, as well as their expected model training throughput.

·Depending on the accuracy and usefulness of the model, the trainer will be rewarded with a corresponding $N3. This motivates them to actively participate and optimize the AI model.

4.3.2 Validators

Validators are responsible for ensuring the integrity and accuracy of transactions and training results in Network3.

Validator's pledge mechanism:

Validators are required to pledge a relatively high amount of $N3 to ensure their integrity and accuracy in processing transactions. This pledge is equivalent to a bond for their honest participation in the network. The role of the Validator is crucial to ensure the stability and security of the network.

Become a Local Validator:

To become a local validator, participants must pledge $N3 and run a full node. Their participation in the consensus and validation process at the local level is crucial to maintaining the decentralization and security of the network.

Become a Global Validator:

Based on the amount of $N3 they pledge, their reputation and historical contributions, Global Validators are selected to oversee the broader network operations. They play a more important role in the network and have more responsibility.

4.3.3 Aggregators

Aggregators are specialized nodes that are responsible for collecting data from a variety of sources and compiling it into a consistent dataset for model training.

Become an aggregator:

Aggregators require a pledge of $N3 as a guarantee of the integrity and quality of the data they provide. Their contribution to the network is not only in the amount of data, but also in its quality and reliability.

4.4 Management, Exit and Participation Rights of Nodes

In order to maintain the stability and governance continuity of Network3, we have set up a series of mechanisms to ensure the effective management of nodes, reasonable exit, and participation rights of pledges.

4.4.1 Dynamic Management and Responsibilities of Nodes

When a node fails to fulfill its responsibilities or chooses to withdraw voluntarily, its pledged $N3 will be released according to the established procedures, a process automatically handled by the smart contract, ensuring transparency and fairness. This mechanism aims to maintain uninterrupted and stable network governance, ensuring that even in the event of a node's withdrawal, qualified candidates are replenished in a timely manner to maintain the normal operation of the network.

4.4.2 Node Withdrawal and Replacement

Network3's node election cycle is set to be quarterly and lasts for one week to keep the network alive and responsive to changes in the community. $N3 token holders gain voting rights by pledging tokens, which are proportional to the amount pledged. To incentivize long-term and active participation, an enhanced voting pool was created to encourage users to gain more voting power through additional pledges.

4.4.3 Pledgee rights and participation

Pledgers not only have the right to vote, but can also increase their pledges during the election period to gain more votes and potential revenue. All voting results and records are open and transparent for community members to monitor and verify, ensuring a fair and traceable election process.

4.5 Punishment and Prevention of Misconduct

In order to maintain the security and integrity of Network3, we have implemented a set of strict penalties for misbehavior and established effective preventive measures.

4.5.1 Penalty for Misbehavior

If a node misbehaves, such as submitting false transactions or manipulating data, part of its pledged $N3 may be confiscated. The severity of the penalty is proportional to the severity of the violation, which is intended to deter misconduct and protect the health of the network.

4.5.2 Prevention of Misconduct

A robust reporting and dispute resolution system is in place to identify and penalize misbehaving nodes. Honest nodes that successfully identify misbehavior are rewarded, encouraging nodes to monitor each other to ensure the integrity of the network.

4.5.3 Node Reputation System

Network3 also introduces a node reputation system for tracking and evaluating nodes' performance and behavior over time. Nodes with higher reputations are more likely to be selected for higher level roles and receive a larger percentage of the network rewards, incentivizing consistent, high quality participation.

5 Incentive Model

Network3 allows participants to mine by sharing bandwidth and arithmetic. Bandwidth mining utilizes participants' local idle bandwidth to participate in data transmission in the network. Arithmetic mining includes complex tasks such as task scheduling, local training, model validation, and parameter aggregation. In Network3's incentive model, it includes the design of three main frameworks: the integral system for bandwidth mining, the proof of participants' contribution, and the game-theoretic in distributed arithmetic.

Under our framework, we provide two examples of FL incentive mechanisms, one on the demand side and the other on the supply side. The proposed Crémer-McLean mechanism and Procurement-VCG (PVCG) mechanism encourage FL participants to truthfully report their type parameters and offer their best datasets to the federation. These mechanisms also provide theoretical guarantees for incentive compatibility, allocative efficiency, individual rationality, and weak budget balancedness.

5.1 Points system for bandwidth market

The bandwidth sharing network is the foundation of the whole Layer2 network, and the distributed off-chain computing on Network3 relies on the bandwidth sharing network to accelerate the data transmission and the privacy and security of the transmission process.Network3 will prioritize the launch of the bandwidth sharing network application, so that the miners can use the local idle bandwidth to participate in the Layer2 mining activities first. At this stage, N3 tokens have not yet been issued, and the process of bandwidth sharing is relatively simple, we will first design a point system to calculate the income of miners in the bandwidth market, and then carry out token mapping after the launch of the arithmetic market test network, and the ratio of mapping will be decided by the community of N3 holders.

5.2 Assessment of contributions

5.2.1 In the bandwidth market

The contribution determination in the bandwidth market will be determined by monitoring the bandwidth usage of the nodes, whose miner's contribution ρ can be determined by the parameters of bandwidth sharing online hours T, data transfer volume D, network delay t_{delay} and node reputation δ.

$$\rho = (\omega_1 \cdot T - \omega_2 \cdot t_{delay} + \omega_3 \cdot \delta) \cdot D \tag{1}$$

Where $\omega_1 \cdot T - \omega_2 \cdot t_{delay} + \omega_3 \cdot \delta$ is always greater than 0. If its value is less than 0, the contribution ρ is taken directly to 0, and the node will be moved to the blacklist within the reputation system, and not allowed to join the network. Node reputation $\delta <= 0$, is determined by the data correctness verification mechanism. The node reputation of Network3's bandwidth sharing system only considers whether the node is malicious or not, so normally normal nodes $\delta = 0$. When a participant node transmits data incorrectly, the node reputation scoring mechanism is triggered. In the early days of Network3 system, the reputation system is maintained by the official, and points are deducted from the evil nodes directly, non-human information distortion will not harm the data transmission, but it can be considered as unreliable corresponding to the node's operating environment, which will deduct fewer reputation points and will be accumulated continuously. In this case, a small number of reputation points will be deducted and accumulated. In the case of human-induced evil behaviors, such as counterfeiting, tampering, interception, or even poisoning attacks, the node's reputation points will be deducted to a systematic infinity value, and the node will be blacklisted. In fact, this kind of situation almost does not happen, Network3 has designed a perfect security anonymous transmission protocol to avoid the above attacks on the data transmission process, the blacklisting operation in the reputation score is only to increase the cost of malicious nodes.

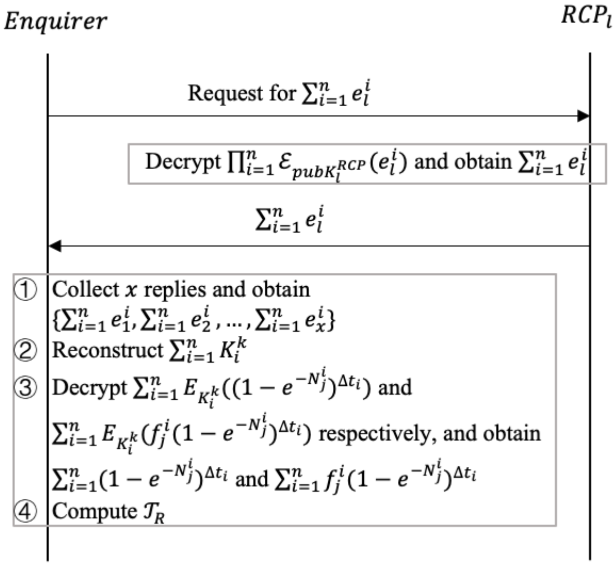
$Enquirer$
RCP_l
Request for $\sum_{i=1}^{n} e_l^i$
Decrypt $\prod_{i=1}^{n} \mathcal{E}_{pubK_l^{RCP}}(e_l^i)$ and obtain $\sum_{i=1}^{n} e_l^i$
$\sum_{i=1}^{n} e_l^i$
① Collect x replies and obtain $\{\sum_{i=1}^{n} e_1^i, \sum_{i=1}^{n} e_2^i, \ldots, \sum_{i=1}^{n} e_x^i\}$
② Reconstruct $\sum_{i=1}^{n} K_l^k$
③ Decrypt $\sum_{i=1}^{n} E_{K_i^k}((1 - e^{-N_j^i})^{\Delta t_i})$ and $\sum_{i=1}^{n} E_{K_i^k}(f_j^i(1 - e^{-N_j^i})^{\Delta t_i})$ respectively, and obtain $\sum_{i=1}^{n}(1 - e^{-N_j^i})^{\Delta t_i}$ and $\sum_{i=1}^{n} f_j^i(1 - e^{-N_j^i})^{\Delta t_i}$
④ Compute $\mathcal{T}_R$

Figure 4: Process of key reconstruction and average rating calculation.

Once the computing market is completed, the management of the reputation scoring system will be transferred from Network3 to the community. We use homomorphic encryption to build a decentralized fair and impartial scoring calculation system. We describe the details of this rating system in a whitepaper and verify its security against various types of attacks.

5.2.2 In the arithmetic market

For distributed arithmetic sharing systems, in which contribution evaluation is relatively complex, we have completed the design and elaborated the reasoning process in the whitepaper, and listed here the calculation of rewards for each node role.

Rewards for Local Trainers: In Formula 2, s_x represents the local dataset size of t_x, *epochs* represents the number of training rounds of t_x, and the reward unit $\hat{r}$ is a hyperparameter. HT is a hash table maintained by the aggregation node for determining the accuracy of the model parameters of the training node. The parameters are considered successfully verified if $HT(P_x^j) \geq 0.5$. The expected reward obtained by the trainer t_x in the j-th iteration is

$$r_x^j = \begin{cases} |s_x| \cdot |epochs| \cdot \hat{r}, & \text{if } HT(P_x^j) \geq 0.5, \\ 0, & \text{if } HT(P_x^j) < 0.5. \end{cases} \tag{2}$$

Rewards for Local Validators: VQ is a queue that stores the voting sequence of all local evaluators, maintained by the aggregator and recorded on the blockchain. $v^j(y,i)$ is the evaluation value of evaluator e_y on trainer t_i in the j-th iteration. Based on the index of the vote, $r_y(j)$ is calculated using Formula 3.

$$r_y^j = \sum_{i}^{|T|} \frac{|VQ| - VQ.FetchIndex(v^j(y,i))}{VQ} \cdot |s_y| \cdot \hat{r} \tag{3}$$

Rewards for Aggregators: The computational and storage costs for aggregators mainly come from maintaining the parameter verification hash table HT and the evaluator voting queue VQ, their rewards r_a^j are calculated using Formula 4.

$$r_a^j = |HT| \cdot |VQ| \cdot \hat{r} \tag{4}$$

Rewards for Global Validators: Global validators play a crucial role in the block creation process by verifying the global models generated by aggregators and determining the final result by executing PBFT consensus. Therefore, they will get equitable rewards derived from the tokens generated during the creation of a new block in the blockchain network. The concrete amount relies on the tokenomic design and the number of validators in the PBFT committee, as depicted in Formula 5.

$$r_g^j = R/m, \tag{5}$$

where R denotes the corresponding value of the rewarded token and m represents the number of global validators.

5.3 The Game-Theoretic Framework in Arithmetic Markets

In the role benefit distribution model of federated learning, the benefits of the global verifier and local verifier are almost fixed like the miners in the blockchain, and will only change with the activity of the system, the aggregator is the task scheduler and coordinator, and its benefits are positively correlated with the trainer.The only more complicated thing is that during the local training process of the model, the trainer may use the data from other data sources, and we need to give additional consideration to the gaming problem in the local training scenario of federated learning.Trainer's operating environment will have some thresholds, while other network participants who do not fulfill the hardware requirements can participate in the training process by providing data. A game model is needed between the data providers and the real trainers to distribute their benefits fairly.

In order to incentivize data providers to contribute the best data sets to the training process, we need to pay sufficient rewards to data providers to cover their costs. The marginal monetary reward for contributing more data should be no less than the resulting marginal cost. In addition, our goal is to maintain a balanced budget and optimize social welfare. At least three sources of information asymmetry are intertwined in this problem: 1) the dataset owned by each data provider; 2) the cost of each data provider; and 3) the valuation of trained FL models by model users. To overcome these information asymmetries and achieve the above objectives, we design rational incentives, i.e., a function that calculates participants' payoffs[1].

There exists a set of n data providers, denoted by $N = \{0, \dots, n-1\}$, and another set of m model users, denoted by $M = \{n, \dots, n+m-1\}$. Each data provider $i \in N$ owns a dataset $\bar{d}_i$. It claims it owns a dataset $\hat{d}_i$. The federation accepts a dataset $d_i \leq \hat{d}_i$ from this data provider. We call $\eta_i = d_i \oslash \hat{d}_i$ the acceptance ratio, where $\oslash$ denotes element-wise division. Trained on datasets $d = (d_0, \dots, d_{n-1})$ from all data providers, the usefulness of the federated model is $Q(d)$. Model users may be granted limited access to the federated model such that the usefulness of the federated model to model user j is $\kappa_j Q(d)$, where κ_j is called the access permission. Each data provider $i \in N$ has a cost type $\gamma_i \in \Gamma_i$. Its cost of contributing data d_i is $c(d_i, \gamma_i)$. The collection of cost types of all data providers forms the cost type profile $\gamma = (\gamma_0, \dots, \gamma_n)$. Data provider i may report a different cost type $\hat{\gamma}_i$[2].

Each model user $j \in M$ has a valuation type $\theta_j \in \Theta_j$. Its valuation on the trained federated model is

$$w(\kappa_j Q(d), \theta_j) = v(d, \kappa_j, \theta_j) \tag{6}$$

.

The collection of valuation types of all model users forms the valuation type profile $\theta = (\theta_n, \dots, \theta_{n+m-1})$. Model user j may report a different valuation type $\hat{\theta}_j$. The payment to data provider $i \in N$ is $p_i \geq 0$. The payment to model user $j \in M$ is $p_j \leq 0$. We denote $p_s = (p_0, \dots, p_{n-1})$ and $p_d = (p_n, \dots, p_{n+m-1})$. The federation income is $I = -\sum_{j=n}^{n+m-1} p_j$; the federation expenditure is $E = \sum_{i=0}^{n-1} p_i$; the federation profit is $P = \sum_{l=0}^{n+m-1} p_l$.Participants' preferences are represented by quasi-linear utility functions:

$$\begin{gathered} u_i(\cdot) = p_i(\cdot) - c_i(\cdot) \text{ for } i \in N; \\ u_j(\cdot) = p_j(\cdot) + v_j(\cdot) \text{for } j \in M. \end{gathered} \tag{7}$$

The social effect of federated learning is measured by social surplus, defined as

$$S(\cdot) = \sum_{j=n}^{n+m-1} v_j(\cdot) - \sum_{i=0}^{n-1} c_i(\cdot) \tag{8}$$

, which includes consumer surplus $S_d = \sum_{j=n}^{n+m-1} v_j(\cdot)$ and producer surplus $S_d = -\sum_{i=0}^{n-1} c_i(\cdot)$. There are user-defined unfairness functions $\bar{w}_s(p_s, c)$ and $\bar{w}_d(p_s, v)$ that measure the unfairness among data providers and model users.

We set the data providers as the supply side and the trainers as the demand side, and use the PVCG model, which maximizes social welfare, as the supply-side auction offer model, and use the Cremer-McLean mechanism to maximize the demand-side utility under the condition of satisfying the five target properties in Appendix A.

Given that the federation Income $I(Q)$ and the model quality $Q(\hat{d}, \eta)$ are exogenous functions, the supply-side FL incentive mechanism design is to design the optimal. Specifically on the supply side letting the data provisions provide the maximum data efficiency $\eta_i(\hat{d}, \hat{\gamma})$ and provide the optimal reward $p_i(\hat{d}, \hat{\gamma})$ for them, and on the demand side letting the trainers produce the optimal model results so that the trainer outputs the maximum model validity $\kappa_j(\hat{\theta})$ and provides the optimal reward $p_j(\hat{\theta})$ for them.

Crémer-McLean mechanism:The detailed process of the Crémer-McLean mechanism for the demand side involves the following steps:

*Step*1 : Consumer Valuation and Preferences : - Each consumer i privately holds a valuation θ_i for the product or service being offered.

*Step*2 : Decision Rule : - Consumers submit their valuations through a decision rule $\kappa(\hat{\theta})$, where $\hat{\theta}$ represents the reported valuations.

*Step*3 : Payment Rule Design : - An interim incentive compatible and interim individually rational payment rule $p(\hat{\theta})$ is designed to extract the full consumer surplus.

Crémer-McLean Theorem Formulation : - The Crémer-McLean Theorem states that for any decision rule $\kappa(\hat{\theta})$ satisfying the Crémer-McLean condition and identifiability condition, there exists a payment rule $p(\hat{\theta})$ that extracts full consumer surplus:

$$-\sum_{j=n}^{n+m-1} p_j(\hat{\theta}) = \sum_{j=n}^{n+m-1} w(\kappa_j(\hat{\theta})Q, \theta_j) \tag{9}$$

Optimization Process : The payment rule $p(\hat{\theta})$ can be found by minimizing a loss function to ensure interim incentive compatibility, individual rationality, and full consumer surplus extraction.The optimization process for the Crémer-McLean mechanism is shown in the appendixB.1.

Procurement-VCG (PVCG) mechanism: As a counterpart of the Crémer-McLean mechanism, we create the PVCG on the supply side. The PVCG mechanism is designed to incentivize FL participants to truthfully report their type parameters and offer their best datasets to the federation. This mechanism provides theoretical guarantees for incentive compatibility, allocative efficiency, individual rationality, and weak budget balancedness. The PVCG mechanism, along with the Crémer-McLean mechanism, aims to address the challenges of information asymmetry and free-riding in federated learning by providing appropriate incentives to participants.

When designing the supply-side mechanism, we assume the federation income $I(Q)$ to be an exogenous function that depends on the quality of the federated model Q. This assumption allows us to focus on optimizing the supply-side FL incentive mechanism without directly considering the intricacies of how the federation income is determined. By assuming $I(Q)$ as an exogenous parameter, we can streamline the design process and concentrate on factors such as dataset contributions, cost types, and acceptance ratios to achieve the desired objectives in federated learning. The supply-side use of the PVCG mechanism gives us the revenue of the federated learning training process as:

$$I(Q) = -\sum_{j=n}^{n+m-1} p_j(\theta) = \sum_{j=n}^{n+m-1} w(\kappa_j(\theta)Q, \theta_j) \tag{10}$$

Procurement auction process[3] of PVCG and payment calculation for data providers are as follows:

*Step*1. Data providers claim datasets to offer and bid on cost types: As the first step, each data provider submits a sealed bid for their claimed datasets and cost types. The claimed dataset $\hat{d}_i$ is the dataset that data provider i claims to offer for federated learning. It may differ from the actual dataset $\bar{d}_i$ owned by the data provider. Similarly, the reported cost type $\hat{\gamma}_i$ may differ from the true cost type γ_i.

*Step*2. The coordinator chooses the optimal acceptance ratios: The coordinator determines the optimal acceptance ratios η_i for each data provider by maximizing the social surplus:

$$\eta^* = \arg \max_{\eta \in [0,1]^{dim(d_i) \times n}} \left\{ I(\hat{d} \cdot \eta) - \sum_{i=0}^{n-1} c_i(\hat{d}_i \cdot \eta_i, \hat{\gamma}_i) \right\} \tag{11}$$

*Step*3. Data providers contribute accepted datasets to federated learning: Data providers contribute their accepted datasets $\hat{d} \cdot \eta^*$ to federated learning. If a data provider cannot contribute $\hat{d}_i \cdot \eta_i^* \leq \hat{d}_i$, a high punishment is imposed. The income to the federation is $I(\hat{d} \cdot \eta^*)$.

*Step*4. The coordinator makes transfer payments to data providers according to the PVCG sharing rule: The PVCG payment $p_i(\cdot)$ consists of the VCG payment τ_i and the optimal adjustment payment h_i^*: $p_i(\cdot) = \tau_i(\cdot) + h_i^*(\cdot)$

The VCG payment τ_i to data provider i is calculated as: $\tau_i = S^*(\hat{d}, \hat{\gamma}) - S_{-i}^*(\hat{d}_{-i}, \hat{\gamma}_{-i}) + c(\hat{d}_i \cdot \eta_i^*)$

where S^* represents the maximum producer surplus and S_{-i}^* is the surplus without data provider i. $\hat{d}_{-i}$ and $\hat{\gamma}_{-i}$ denote the claimed datasets and reported cost types excluding data provider i.

Appendix

A Desirable properties of FL incentive mechanism design

A.0.1 Incentive Compatibility (IC)

IC is attained if in equilibrium, all participants report their types truthfully, i.e., $\hat{\theta} = \theta$. Different types of equilibriums correspond to different IC conditions, which can be one of Nash Incentive Compatibility (NIC), Dominant Incentive Compatibility (DIC), Bayesian Incentive Compatibility (BIC), or Perfect Bayesian Incentive Compatibility (PBIC).

A.0.2 Individual Rationality (IR)

A mechanism is individually rational (IR) if this mechanism does not make any player worse off than if he quits the federation, i.e.,

$$u_i(\hat{d}, \hat{\gamma}) \geq 0, \forall i \in N \text{ and } u_j(\hat{\theta}) \geq 0, \forall j \in M.$$

In games of incomplete information, IR can be ex-ante IR, interim IR or ex-post IR.

A.0.3 Budget Balancedness (BB)

BB requires that the total payments collected from participants equal the total cost of running the FL federation. Formally, BB is defined as

$$\sum_{i=1}^{n} p_i(\hat{d}, \hat{\gamma}) + \sum_{j=n+1}^{n+m-1} p_j(\hat{\theta}) = C(\hat{d}),$$

where $p_i(\hat{d}, \hat{\gamma})$ is the payment made to data provider i, $p_j(\hat{\theta})$ is the payment made to model user j, and $C(\hat{d})$ is the cost of running the FL federation.

A.0.4 Data Offering Rate (DOR)

DOR is defined as the total data offered by all data providers to the total data owned by all data providers, i.e.,

$$\text{DOR} = \frac{\sum_{i=1}^{n} \hat{d}_i}{\sum_{i=1}^{n} \bar{d}_i}.$$

The data offering rate varies from 0.0 to 1.0, with 1.0 indicating all data being offered. When a payment scheme is incentive-compatible, the data offering rate is 1.0.

A.0.5 Model Quality (MQ)

MQ requires that the FL incentive mechanism maximizes the quality of the federated model. Formally, MQ is defined as

$$\max_{\hat{d}} Q(\hat{d}),$$

where $Q(\hat{d})$ is the quality of the federated model trained on the dataset $\hat{d}$.

A.0.6 Fairness

Fairness requires that the FL incentive mechanism distributes the benefits of the federation fairly among participants. There are different definitions of fairness, such as envy-freeness, proportionality, and egalitarianism.

These objectives are often in tension with each other, and designing an FL incentive mechanism that satisfies all objectives simultaneously is challenging. In the next section, we propose a game-theoretic framework for FL incentive mechanism design that balances these objectives.

B Crémer-McLean

Theorem 1. *(Crémer-McLean Theorem) When the Crémer-McLean condition and the identifiability condition hold for Prior(θ), for any decision rule $\kappa(\hat{\theta})$, there exists an interim incentive compatible and interim individually rational payment rule $p(\hat{\theta})$ that extracts full consumer surplus, i.e., $-\sum_{j=n+m-1}^{n} p_j(\hat{\theta}) = \sum_{j=n+m-1}^{n} w(\kappa_j(\hat{\theta})Q, \theta_j)$.*

As an application of this theorem, we can set $\kappa_j(\hat{\theta}) \equiv 1$, i.e., every model user gets full access permission to the FL model. In this case, $w(\kappa_j(\hat{\theta})Q, \theta_j) = w(Q, \theta_j)$, and we can find an interim incentive compatible and interim individually rational payment rule $p(\hat{\theta})$ such that $-\sum_{j=n+m-1}^{n} p_j(\hat{\theta}) = \sum_{j=n+m-1}^{n} w(Q, \theta_j)$.

B.1 Training Crémer-McLean Mechanism

The Crémer-McLean payments can be calculated by automated mechanism design techniques, as discussed in [4].

The Crémer-McLean payments $p(\theta)$ should satisfy three constraints for expost full consumer surplus extraction, interim incentive compatibility, and expost individual rationality, respectively[4]:

$$\begin{cases} -\sum_{j=n}^{n+m-1} p_j(\theta) = \sum_{j=n}^{n+m-1} w(Q, \theta_j), & \forall \theta; \\ \sum_{\theta'_{-j}} [w(Q, \theta_j) + p_j(\theta_j, \theta'_{-j})]\text{Prior}(\theta'_{-j}|\theta_j) \geq 0, & \forall j \in M, \theta_j \in \Theta_j; \\ \sum_{\theta'_{-j}} [p_j(\theta_j, \theta'_{-j}) - p_j(\hat{\theta}_j, \theta'_{-j})]\text{Prior}(\theta'_{-j}|\theta_j) \geq 0, & \forall j \in M, \theta_j \in \Theta_j. \end{cases} \tag{12}$$

The Crémer-McLean Theorem guarantees the existence of a solution $p(\theta)$ to the above constraints. By minimizing the following loss function, the Crémer-McLean payments can be learned using standard backpropagation algorithms:

$$
\begin{aligned}
LOSS = &\left(\sum_{j=n}^{n+m-1}[w(Q,\theta_j)+p_j(\theta)]\right)^2 \\
&+\sum_{j=n}^{n+m-1}\mathrm{ReLu}\left\{-\sum_{\theta'_{-j}}[w(Q,\theta_j)+p_j(\theta_j,\theta'_{-j})]\mathrm{Prior}(\theta'_{-j}|\theta_j)\right\} \\
&+\sum_{j=n}^{n+m-1}\mathrm{ReLu}\left\{-\sum_{\theta'_{-j}}[p_j(\theta_j,\theta'_{-j})-p_j(\hat{\theta}_j,\theta'_{-j})]\mathrm{Prior}(\theta'_{-j}|\theta_j)\right\}
\end{aligned}
\tag{13}
$$

Here, θ, θ', $\hat{\theta}$and are drawn randomly from the prior distribution of θ.

References

[1] Mingshu Cong, Han Yu, Xi Weng, and Siu Ming Yiu. A game-theoretic framework for incentive mechanism design in federated learning. *Federated Learning: Privacy and Incentive*, pages 205–222, 2020.

[2] Mingshu Cong, Xi Weng, Han Yu, Jiabao Qu, and Siu Ming Yiu. Optimal procurement auction for cooperative production of virtual products: Vickrey-clarke-groves meet cremer-mclean. *arXiv preprint arXiv:2007.14780*, 2020.

[3] Hal R Varian and Christopher Harris. The vcg auction in theory and practice. *American Economic Review*, 104(5):442–445, 2014.

[4] Michael Albert, Vincent Conitzer, and Giuseppe Lopomo. Assessing the robustness of cremer-mclean with automated mechanism design. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 2015.