Letter

An Isomerism Learning Model to Solve Time-Varying Problems Through Intelligent Collaboration

Zhihao Hao, Guancheng Wang, Bob Zhang, *Senior Member, IEEE*, Leyuan Fang, *Senior Member, IEEE*, and Haisheng Li

Dear Editor,

This letter deals with a solution for time-varying problems using an intelligent computational (IC) algorithm driven by a novel decentralized machine learning approach called isomerism learning. In order to meet the challenges of the model's privacy and security brought by traditional centralized learning models, a private permissioned blockchain is utilized to decentralize the model in order to achieve an effective coordination, thereby ensuring the credibility of the overall model without exposing the specific parameters and solution process. Moreover, nodes in the network are equipped with different models to meet many challenges caused by the model silos. Furthermore, an integration scheme is introduced to efficiently obtain the global solutions of time-varying problems. In this letter, the convergence of the proposed model is theoretically proven, where its efficiency is validated via experiments, which shows that it outperforms many state-of-the-art models using centralized processing.

Introduction: In recent years, solving time-varying problems has become a research hot spot, and many IC algorithms have been proposed. In particular, recurrent neural network (RNN) models have been greatly developed due to their outstanding performance in solving time-varying problems [1]–[3]. Many gradient-based models and numerical algorithms, e.g., the gradient neural network (GNN) model and the Newton Raphson iterative (NRI) algorithm, are designed to solve static problems, which may have lagging errors when solving time-varying problems due to the lack of velocity compensation of the time-varying parameters [4], [5]. To solve this limitation, the zeroing neural network (ZNN) model is reported and can solve time-varying problems without residual error as it utilizes the time-derivative information of the parameters. In recent years, many brilliant works have been designed, including the integration-implemented NRI (IINRI) algorithm [6], distributed RNN model [7], and finite-time neural dynamic model [8]. However, limited by the traditional centralized processing, these IC algorithms are greatly affected by many factors such as model scale, parameter size, and computing power. In addition, problems such as isolation and lack of interactions among centralized models can create model silos, which can lead to many challenges for transparency, efficiency, credibility and so on. All these challenges will reduce the stability of centralized models and become obstacles to the implementation of centralized IC algorithms in reality: It not only has higher requirements for ethics, privacy, and confidentiality, but also needs security and fault tolerance in design.

Distributed learning represented by federated learning (FL) is developed and can solve these problems partially by using a parameter server responsible for aggregation and distribution to update the local models, with the dynamic aggregation process also accurately grasping the dynamic changes of the parameters in time-varying problems. That being said, this approach does not take into account the model heterogeneity, which is caused by differences among models with various structures. Besides this, there may be competition in a distributed learning framework consisting of isomerous models. Influenced by some works such as the Matthew effect discussed in [9], the validity of the overall model may be controlled by several models with high performance, resulting in multi-center occurrence. This creates a challenge for security: the corruption of these models will lead to the failure of the overall model.

Motivated by these issues, we introduce isomerism learning (IL), which combines various distributed machine learning models and decentralized hardware infrastructures with a private permissioned blockchain. IL can be divided into three layers as Fig. 1 shows. Time-varying problems such as the dynamic Stein equation can be processed by an Isomerism Learning model based on various IC (IL-IC) algorithms that meet the challenges caused by model heterogeneity and model silos. The contributions of this letter can be summarized as follows: 1) We propose a novel distributed machine learning method to realize effective collaboration between isomerous models. 2) The proposed method solves a problem through cooperation between different nodes equipped with heterogeneous models. All models are involved to ensure the decentralization. 3) A distributed integration algorithm is designed to implement dynamic aggregation of the parameters and obtain the global solution for time-varying problems. 4) Extensive experiments demonstrate that the proposed method outperforms many centralized state-of-the-art models.

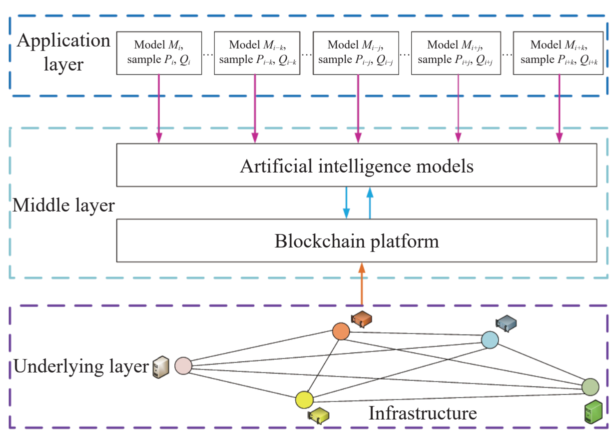

Fig. 1. Architecture of IL.

Architecture of IL-IC: IL-IC provides security measures to support confidentiality and sovereignty enabled by a private permissioned blockchain, which consists of one permissioned node and multiple permissionless nodes like [10]. Therefore, it can perfectly

Corresponding author: Bob Zhang.

Citation: Z. H. Hao, G. C. Wang, B. Zhang, L. Y. Fang, and H. S. Li, "An isomerism learning model to solve time-varying problems through intelligent collaboration," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 8, pp. 1772–1774, Aug. 2023.

Z. H. Hao is with the Department of Computer and Information Science, University of Macau, Macau 999078, the School of Data Science, the Chinese University of Hong Kong, Shenzhen 518172, Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518000, and China Industrial Control Systems Cyber Emergency Response Team, Beijing 100040, China (e-mail: hao.zhihao@connect.um.edu.mo).

G. C. Wang and B. Zhang are with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: yc07455@um.edu.mo; bobzhang@um.edu.mo).

L. Y. Fang is with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, and also with the Peng Cheng Laboratory, Shenzhen 518000, China (e-mail: fangleyuan@gmail.com).

H. S. Li is with the Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, and also with the School of Computer Science and Engineering, Beijing Technology and Business University, Beijing 100048, China (e-mail: lihsh@btbu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.123360

cater to the master-worker architecture of traditional FL [11]. Each node is clearly defined, and only pre-authorized nodes can participate in the overall calculation process. The enrollment process of new nodes is dynamic, and a new node needs to use a smart contract to be involved. Also, the dedicated server can be dispensed because the permissioned node can be elected dynamically. Permissionless nodes equipped with the IC algorithms can work together to solve the overall problem with the help of a permissioned node. Here, the relevant information generated by each iteration process is also recorded in a shared ledger for all parties to review. Through this, the system can increase trust between unfamiliar anonymous nodes without the involvement of a third party. Also, security is improved because the recorded information only contains the results of each iteration and does not disclose specific parameters of the local models. The overall process can be divided into several steps:

First, the permissionless nodes initialize their local models before computing the result locally and sending it to the permissioned node. Second, after the permissioned node receives all the results from the permissionless nodes, it performs the result aggregation operation. Afterwards, the permissioned node sends the aggregated result back to the permissionless nodes, which will update the local model for the next iteration process after that (refer to Step 4 in Fig. 2). The whole process is shown in Fig. 2 and continues until reaching the maximum number of iterations.

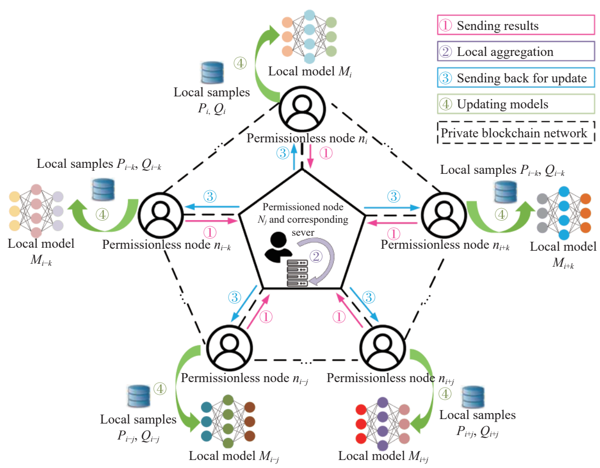

Fig. 2. Implementation process of IL-IC.

Methodology: 1) Distributed aggregation: We propose a distributed integration algorithm (DIA) based on the private blockchain network. DIA is applicable to the finite sum of the update step and the residual error of different IC algorithms. In this letter, DIA is exploited to solve the time-varying Stein equation [12], which is expressed as

$$P(t)X(t)Q(t)+X(t)=O(t) \tag{1}$$

where $X(t)\in\mathbb{R}^{n\times m}$ denotes the unknown matrix, $P(t)\in\mathbb{R}^{n\times n}$, $Q(t)\in\mathbb{R}^{m\times m}$, and $O(t)\in\mathbb{R}^{n\times m}$ are time-varying parameter matrices. For simplification, (1) can be rewritten as

$$A(t)\mathbf{x}(t)=\mathbf{b}(t) \tag{2}$$

where $A(t)=Q^T(t)\otimes P(t)+I^{m\times m}\otimes I^{n\times n}$, T represents the transpose of the matrix or vector, and $\otimes$ denotes the Kronecker product. $\mathbf{x}(t)$ and $\mathbf{b}(t)$ denote the vectorial form of $X(t)$ and $O(t)$, respectively [12]. DIA achieves model updating through results aggregation, and the solution is updated as

$$\mathbf{x}_{k+1}=\mathbf{x}_k+\delta_{IL-IC,k} \tag{3}$$

where $\delta_{IL-IC,k}$ is a linear combination of the update steps of different IC algorithms (e.g., the total number is G). Moreover, $\delta_{IL-IC,k}$ can be expressed as

$$\delta_{IL-IC}=\sum_{g=1}^{G}w_{g,k}\delta_{g,k} \tag{4}$$

where $w_{g,k}$ and $\delta_{g,k}$ are the update step and the corresponding weight of the gth algorithm, correspondingly. Furthermore, the weight of various IC algorithms are determined by their residual error $\epsilon_{g,k}$ in the last update step, it can be calculated as

$$w_{g,k}=\frac{1/\|\epsilon_{g,k}\|_2}{\sum_{g=1}^{G}1/\|\epsilon_{g,k}\|_2}. \tag{5}$$

The proposed model sends the result of the permissionless nodes to the permissioned node for aggregation after each iteration (the overall process is shown in Algorithm 1). According to (5), the smaller the residual error the IC algorithm has, the higher its weight is when calculating the update step of the IL-IC model. In addition, the sum of the weight among IC algorithms equals one.

Algorithm 1 Distributed Integration Algorithm

Permissioned nodes N_j executes:
for each model computation round $r=1,2,\ldots,$ **do**
  N_j determines the set s_r of permissionless nodes
  for $g\in s_r$ **in parallel do**
    PermissionlessNodesUPDATE(δ_g, w_g)
  end for
  N_j executes (4)
  N_j sends $\delta_{IL-IC,r}$ back to s_r for the model update of the next iteration
end for
Obtain the final the δ_{IL-IC} of global solution
PermissionlessNodes UPDATE
Initialize the local model $M_{g,(g=1,2,\ldots,G)}$ provided by N_j and randomly generate $\mathbf{x}_1$
for each iteration k=1,2,...,K **do**
  Samples P_k and Q_k
  Obtain the generation result $\delta_{g,k}$, $\epsilon_{g,k}$, and $\mathbf{x}_{k+1}$
end for
Send ($\delta_{g,k}$, $\epsilon_{g,k}$) to the permissioned node N_j

2) Convergence analysis: Without loss of generality, the IL-IC model exploited two prevalent types of models (the ZNN-type models and NRI-type models), there are investigated and the following theorem is provided and proven.

Theorem 1: When using the IL-IC model to solve the time-varying Stein equation, the estimated solution globally converges to the theoretical one.

Proof: According to (3), we have

$$\mathbf{x}_{k+1}-\mathbf{x}_k=\begin{bmatrix}\delta_{1,k}\\ \cdots\\ \delta_{g,k}\end{bmatrix}[w_{1,k}\quad\cdots\quad w_{g,k}] \tag{6}$$

which can be treated as a linear combination of IC models. According to the definition of the representative models [6] and [8], (6) can be reformulated as

$$\dot{\epsilon}_{IL-IC,k}=-(\mu_{1,k}f_1(\epsilon_{IL-IC_k})+\cdots+\mu_{g,k}f_g(\epsilon_{IL-IC_k})) \tag{7}$$

where $\epsilon_{IL-IC,k}$ denotes the residual error of the IL-IC model. $\mu_{1,k},\ldots,\mu_{g,k}$ are positive parameters and $f_1(\cdot),f_2(\cdot),\ldots,f_g(\cdot)$ are activation functions, which are related to the corresponding IC models. Moreover, defining a Lyapunov candidate $V=\epsilon^T_{IL-IC,k}\epsilon_{IL-IC,k}$, its time derivative can be expressed as

$$\begin{aligned}\dot{V}=\epsilon^T_{IL-IC,k}\dot{\epsilon}_{IL-IC,k}&=-(\mu_{1,k}\epsilon^T_{IL-IC,k}f_1(\epsilon_{IL-IC,k})\\&+\cdots+\mu_{g,k}\epsilon^T_{IL-IC,k}f_g(\epsilon_{IL-IC,k})).\end{aligned} \tag{8}$$

Since V is positive definite and $\dot{V}$ is negative definite, the IL-IC model globally converges to the theoretical solution based on Lyapunov stability theory [13]. ■

Experimental results and analysis: In this section, we evaluated the performance of the IL-IC model. The experiment was run on MATLAB 2019b with 16 G memory, an Intel(R) Core(TM) i3-3110M CPU, and NVIDIA 1080Ti. Also, the test network consisted of a permissioned node and 5 permissionless nodes. What is more, the sampling time interval for all models is set to 0.039 s. In order to reduce the influence caused by the computation time for the different models, the communication time is set to 0.051 s. Based on this, we ran the IL-IC model and observed the gap between the model's generation results and the theoretical value in [0 s, 10 s]. The time-varying Stein equation is defined with

$$P(t)=\begin{bmatrix}2+\cos(2t) & 2\sin(2t)\\ \sin(2t) & 2-\cos(2t)\end{bmatrix}\in\mathbb{R}^{2\times 2}$$
$$Q(t)=\begin{bmatrix}\cos(2t)+2 & \sin(2t) & 0\\ \sin(2t) & 1 & \cos(2t)\\ \cos(2t) & \sin(2t) & 2\end{bmatrix}\in\mathbb{R}^{3\times 3}$$
$$O(t)=\begin{bmatrix}\cos(2t)-2 & \sin(2t)+1 & \cos(2t)\\ \sin(2t) & -2\sin(2t) & \cos(2t)+1\end{bmatrix}\in\mathbb{R}^{2\times 3}. \quad (9)$$

As can be seen from the Fig. 3, the curve of the theoretical value and the curve of the generation results fit well. Moreover, two different metrics, i.e., the mean absolute error (MAE) and the root mean square error (RMSE), are introduced to better evaluate the performance of these different models. The MAE of the residual error ϵ_{MAE} can be defined as $\epsilon_{\text{MAE}}=\frac{1}{b+1}\sum_{i=a}^{a+b}\|\epsilon_{g,i}\|_2$, while the RMSE of the residual error ϵ_{RMSE} can be expressed as $\epsilon_{\text{RMSE}}=\sqrt{\frac{1}{b+1}\sum_{i=a}^{a+b}(\epsilon_{g,i})^2}$, where the residual error becomes steady at $t=a\tau$. Based on these, the relevant values can be shown in Table 1. As illustrated in this table, the proposed model can outperform many models like NRI, IINRI, GNN, and bounded ZNN (BZNN) because it has a lower value in terms of ϵ_{MAE} and ϵ_{RMSE}. Moreover, the IL-IC model is effective without exposing the specific solutions and model parameters, thereby protecting privacy and security. However, there is still a slight gap compared to CZNN In the future, higher-accuracy models for integration and more weight assigned schemes can be designed to narrow the gap, such that the overall model performance is close to the highest-precision model.

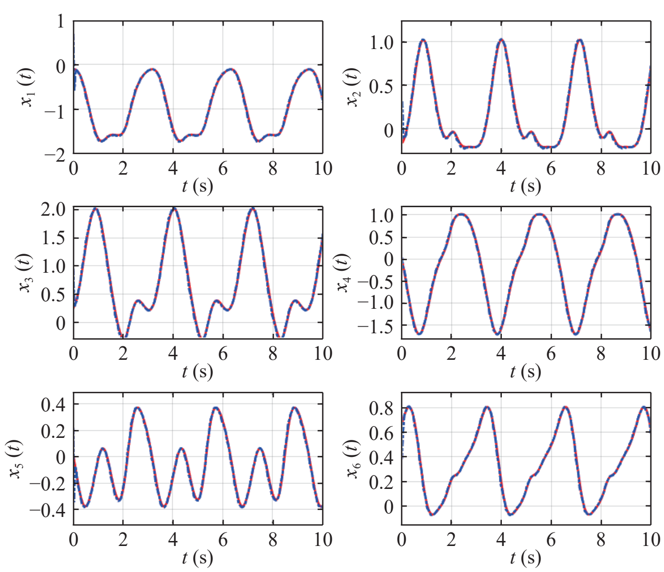

Fig. 3. The solution estimated by the IL-IC model (blue-dotted line) compared with its theoretical value (red-dash line).

Conclusions: In this letter, we proposed a distributed intelligent computing model, which aims to solve the problems like model silos, achieving effective collaboration between different models. For the current model security and trust issues, we use a private permissioned blockchain to achieve trusted interactions between unfamiliar nodes. The validity of the calculation results is guaranteed by the collaboration of different intelligent computing algorithms. Extensive experiments demonstrate that the proposed model's performance is not only close to the highest precision model, but also significantly outperforms most of the state-of-the-art models applying a centralized framework. In the future, we will introduce other blockchain platforms and try more state-of-the-art IC models and integration strategies to achieve swarm intelligence. In addition, this work can also provide a distributed scheme for the current development of artificial intelligence driven by big models and centralized ways. It provides a feasible opportunity for the implementation of efficient and trusted distributed artificial general intelligence.

Table 1. Comparison Between Our Method (IL-IC) and Various Centralized IC Algorithms

Models	ϵ_{MAE}	ϵ_{RMSE}
NRI [6]	3.7×10^{-1}	3.8×10^{-1}
IINRI [6]	2.5×10^{-1}	2.5×10^{-1}
GNN [4]	1.9×10^{-0}	2.0×10^{-0}
CZNN [8]	2.6×10^{-2}	2.7×10^{-2}
BZNN [1]	1.1×10^{-0}	1.1×10^{-0}
IL-IC	8.6×10^{-2}	9.2×10^{-2}

Acknowledgment: This work was supported in part by Shenzhen Science and Technology Program (ZDSYS20211021111415025), the Shenzhen Institute of Artificial Intelligence and Robotics for Society, the National Natural Science Foundation of China (62277001), the Scientific Research Program of Beijing Municipal Education Commission (KZ202110011017), the National Key Technology R&D Program of China (SQ2020YFB10027), and Major Science and Technology Special Project of Yunnan Province (202102AD08 0006).

References

[1] G. Wang, Z. Hao, B. Zhang, and L. Jin, "Convergence and robustness of bounded recurrent neural networks for solving dynamic Lyapunov equations," *Information Sciences*, vol. 588, pp. 106–123, Apr. 2022.

[2] M. Liu, X. Zhang, M. Shang, and L. Jin, "Gradient-based differential kWTA network with application to competitive coordination of multiple robots," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1452–1463, Aug. 2022.

[3] L. Jin, L. Wei, and S. Li, "Gradient-based differential neural-solution to time-dependent nonlinear optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 1, pp. 620–627, 2023.

[4] S. Liao, J. Liu, X. Xiao, D. Fu, G. Wang, and L. Jin, "Modified gradient neural networks for solving the time-varying Sylvester equation with adaptive coefficients and elimination of matrix inversion," *Neurocomputing*, vol. 379, pp. 1–11, Feb. 2020.

[5] G. Wang, Q. Li, S. Liu, H. Xiao, and B. Zhang, "New zeroing neural network with finite-time convergence for dynamic complex-value linear equation and its applications," *Chaos, Solitons & Fractals*, vol. 164, p. 112674, Nov. 2022.

[6] G. Wang, Z. Hao, H. Huang, and B. Zhang, "A proportional-integral iterative algorithm for time-variant equality-constrained quadratic programming problem with applications," *Artificial Intelligence Review*, vol. 56, pp. 4535–4556, 2023.

[7] L. Jin, X. Zheng and X. Luo, "Neural dynamics for distributed collaborative control of manipulators with time delays," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 5, pp. 854–863, May 2022.

[8] L. Jin, S. Li, B. Liao, and Z. Zhang, "Zeroing neural networks: A survey," *Neurocomputing*, vol. 267, pp. 597–604, Dec. 2017.

[9] M. Liu, S. Li, and L. Jin, "Modeling and analysis of Matthew effect under switching social networks via distributed competition," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 7, pp. 1311–1314, Jul. 2022.

[10] S. Warnat-Herresthal, H. Schultze, K. L. Shastry, S. Manamohan, S. Mukherjee, *et al*., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, pp. 265–270, May 2021.

[11] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intelligent Systems and Technology* (*TIST*), vol. 10, no. 2, pp. 1–19, Mar. 2019.

[12] J. Dai, L. Jia, and L. Xiao, "Design and analysis of two prescribed-time and robust ZNN models with application to time-variant Stein matrix equation," *IEEE Trans. Neural Networks Learning Systems*, vol. 32, no. 4, pp. 1668–1677, Apr. 2021.

[13] B. N. Datta, *Numerical Methods for Linear Control Systems*. Amsterdam, Netherlands: Elsevier Academic Press, 2004.