



# Design a Multi-layer Supply Chain Optimization Model Based on Federated Learning and Blockchain

Qi Ma<sup>1</sup> (✉), Shaoyong Guo<sup>1</sup>, Tao Shen<sup>2</sup>, Yan Feng<sup>3</sup>, Jian Gao<sup>2</sup>, and Bei Liu<sup>4</sup>

<sup>1</sup> State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

{maqi, syguo}@bupt.edu.cn

<sup>2</sup> Kunming University of Science and Technology, Kunming, China

shentao@kust.edu.cn

<sup>3</sup> Yunnan Provincial Academy of Science and Technology, Kunming, China

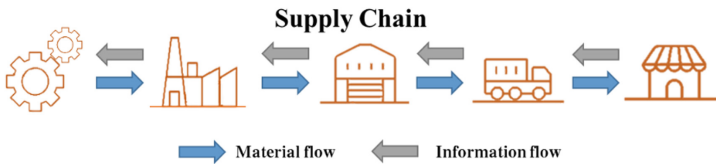
<sup>4</sup> 54th Research Institute of Electronic Technology Corporation, Shijiazhuang, China

**Abstract.** As enterprises have different data security and requirements, data governance between enterprises has become a problem. The traditional relationship between supply and demand can easily lead to the “bullwhip effect” in the supply chain (SC). In order to optimize the interests of most enterprises in the SC, this study designs a multi-layer SC network optimization model. Specifically, first of all, using blockchain technology, a supply and sales framework that conforms to the SCN is constructed, and the chain information flow is transformed into a decentralized information interaction method. Subsequently, the network is optimized from two aspects: on the basis of ensuring the privacy of enterprise data, a time-series federated learning scheme is proposed to predict the needs of multiple parties; relying on the framework of this paper, the enterprise inventory is stabilized. The supply and demand optimization framework proposed in this paper provides information forecasting and sharing services for enterprises from a new perspective. Under the circumstance of ensuring the self-interest of the enterprise, the overall cost reduction and efficiency increase of the SC can be realized.

**Keywords:** Supply chain · Blockchain · Federated learning · Data governance · Demand forecasting

## 1 Introduction

More and more businesses are sharing data with other businesses and using it to change their own demand response strategies. A supply chain (SC) of the entire production process from suppliers to manufacturers to customers has gradually formed between enterprises. Enterprises use SC to accurately understand the information of production, sales and warehousing, to update their own inventory. At the same time, enterprises also update their own production plans and logistics, to improve management level, reduce costs and improve external competitiveness (Fig. 1).



**Fig. 1.** Supply chain information and logistics relations

However, due to the existence of information barriers between SC nodes and the reliability of information transmitted in the chain, it is easy to produce a more serious “bullwhip effect”. Because enterprises keep their own data confidential, there is information conflict between SC nodes, which leads to low communication efficiency when all parties formulate production and procurement requirements, uneven SC data interaction, and complicated information verification methods.

In traditional SC discussions, we assume capacity, demand, and cost are known parameters. But in practice, the uncertainty of changes in factors such as customer demand, transportation and supply cycles can have a huge impact on SC. Therefore, timely and accurate forecasting of demand is necessary. In this paper, we construct a new supply and demand forecasting framework, design and implement the forecasting of SC enterprise demand.

The traditional method, which can only predict based on enterprise data, cannot take into account the overall existence of SC, which is likely to cause information gaps between SC enterprises. In this paper, we use blockchain technology to build a new information sharing model, changing the original chain or mesh information transmission method to decentralized information sharing, to expand the information transparency between SC enterprises. In order to maximize the use of enterprise data for prediction and ensure the privacy and security of core data of each enterprise, we use the Federated Learning (FL) method to predict the supply/demand problems in the SC, improve the prediction accuracy, and improve the prediction accuracy. To achieve the purpose of optimizing the SC.

Due to the construction of the Yunnan Genetic Materials Project, there are more and more demands for sharing between databases. In this study, we adopted a two-layer demand allocation framework, which includes a blockchain layer that helps SC enterprise transactions and a time-series FL layer that completes demand forecasting.

The main contributions of this paper are summarized as follows:

1. We built an information sharing model that conforms to the SC scenario to ensure that enterprises communicate while optimizing the original SC information transfer relationship.
2. We designed a time-series demand forecasting algorithm based on FL, which can provide more accurate forecasting results for multi-party enterprises on the basis of ensuring the data privacy and security of each enterprise.
3. We implemented a learning control scheme. With the assistance of blockchain, we can quickly converge to find the best inventory, to optimize the SC and reduce the “bullwhip effect”.

4. The experimental results show that our method are superior to the traditional model, which can effectively suppress the “bullwhip effect” and achieve the ultimate goal of optimizing the SC.

The rest of the paper is structured as follows. Section 2 reviews related work on SC demand forecasting. In Sect. 3, we introduce the system architecture and model. Section 4 designs and studies a FL time series prediction model. Section 5 presents the specific research content, Sect. 5.2 presents simulation results, and Sect. 6 summarizes the paper and presents future work.

## 2 Related Work

In this section, we review recent work in various fields from three aspects: SC development, demand forecasting, and FL.

### 2.1 Supply Chain

At present, the theoretical research of SC management is more and more in-depth. Xu et al. [1] studied how the SC management platform of the enterprise makes the information flow, logistics and capital flow better flow in the chain. The key of SC management is to spend the lowest cost on the premise of controlling the inventory of node enterprises [2, 3]. Ma et al. [4] studied the coordination problem of agricultural product SC inventory system, to reduce inventory costs and management costs while improving the level of SC management. As the focus of new SC research, the overall cost, benefit and profit distribution of the SC should be considered through new technologies.

### 2.2 Demand Forecast

Demand forecasting is the basis for the overall planning of the SC. Traditional forecasting methods usually utilize linear functions of historical time series data to forecast future sales. However, they only apply to data with stable or seasonal sales trends.

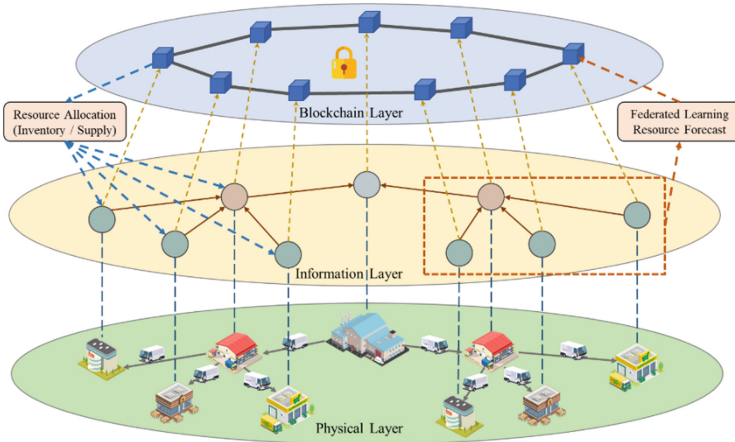
At present, Recurrent Neural Network (RNN) is very important in the field of deep learning. It is suitable for time series with strong randomness and poor stationarity. The LSTM neural network was proposed for the RNN neural network that cannot solve the long-order dependency of time series [5]. Sutskever [6] et al. (2014) proposed a Sequence to Sequence (Seq2Seq) learning model, which solves the problem of variable length of input and output. In this paper, we intend to use a recurrent neural network to automatically learn effective features from historical data for demand forecasting.

### 2.3 Federated Learning

Federated Learning (FL) [7] is a distributed machine learning (ML) approach that supports training ML models on distributed data (on edge devices). The development of FL is a research direction that is accompanied by industrial needs and the requirements of data privacy protection, it can alleviate privacy leakage to a certain extent [8, 9, 10, 11, 12]. In this paper, we try to design a new idea of FL that meets the requirements of SC scenarios to meet the requirements of SC enterprises.

### 3 Supply Chain Architecture Based on Consortium Chain and Federated Learning

The framework constructed in this paper is mainly divided into three parts. The physical layer is the actual SC scenario; the information layer is to abstract the actual enterprise as a node to build a SC network; the blockchain layer realizes information sharing and data governance (Fig. 2).



**Fig. 2.** Supply chain information sharing framework

In this paper, we use blockchain technology to build a suitable information sharing platform, and rely on FL technology to protect and predict the needs of enterprises, and comprehensively optimize the SC in terms of supply and demand.

Next, this chapter will elaborate the structure of this paper in more detail through the demand forecast stage and the supply feedback stage.

#### 3.1 Demand Forecast Stage

We ensure the efficiency, accuracy and reliability of forecasts by designing new supply and demand interaction models. In order to accurately describe the model training mode under the combined architecture of FL and alliance chain, as shown in the figure, the global training process can be divided into the following steps:

1. An enterprise  $E_i$  provides the initial training model and publishes it on the alliance chain for consensus
2. Other enterprises  $E_j (j \in \mathbb{R} \text{ and } j \neq i)$  on the alliance chain respond to enterprise  $E_i$  and start to use FL for demand forecasting;
3. Each enterprise uses its own data to forecast, and finally publishes the demand forecast result to the consensus on the alliance chain;

4. Enterprise  $E_i$  proposes to the manufacturer to make up the position through the prediction result, and publishes it on the alliance chain, and then supplies the demand.
5. Upstream producers get the news of replenishment from the alliance chain and allocate relevant resources for them.
6. The producer sends the relevant supply to the consortium chain for consensus.

The FL process involved is shown in the figure below (Fig. 3).

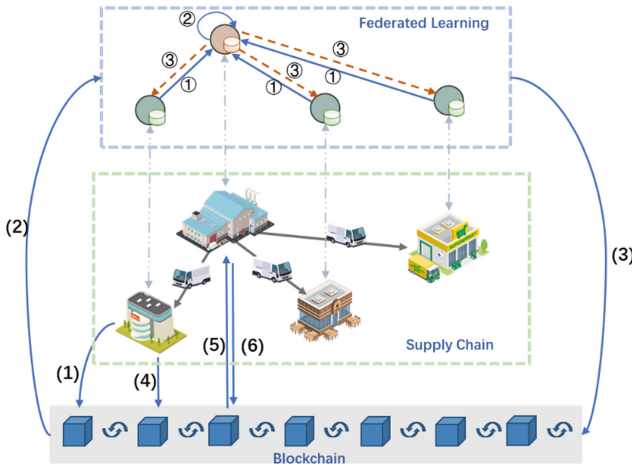


Fig. 3. The whole process of data prediction and sharing

### 3.2 Supply Stage

The information flow in the SC is a chain structure, which is prone to distortion in the process of data transmission, resulting the bullwhip effect. By enabling SC enterprises to on the chain, they can get a clearer information flow and respond more efficiently to the supply/demand of enterprises.

Different positions and relationships of enterprises in the SC make different enterprises have different willingness to share information. After the private information is accessed by other companies, the party that provides the private information will receive the agreed number of points as compensation to improve its initiative in information sharing.

In this paper, the process for enterprises to share supply information through blockchain is as follows:

1. According to the demand forecast results, each enterprise determines the warehouse replenishment demand order to the upstream enterprise, submits the order demand on the alliance chain, and the enterprise signs it and sends it to all endorsement nodes.

2. The production/manufacture summarizes the warehouse replenishment requirements put forward by each downstream enterprise;
3. Broadcast the calculated supply proposal to the consortium chain;
4. After receiving the delivery proposal, the logistics transporter will arrange the transportation plan reasonably and publish the transportation proposal on the alliance chain;
5. After consensus, the logistics party completes the transportation of the goods and ends the supply.

## 4 Federated Learning Time Series Prediction Model Based on Seq2seq Algorithm

### 4.1 FL Aggregation Algorithm

Assume that the edge node  $n$  represented by each enterprise has a local data sample  $S_n$ . The total size of data samples for  $n$  edge nodes is  $\sum_{n=1}^N S_n = S$ . Then, the goal of FL is to reduce the global loss function  $l(\Phi)$ , which is minimized by minimizing the weighted average of the local loss functions  $l_n(\Phi)$  trained by each edge node using its local dataset. Therefore, the local loss function  $l_n(\Phi)$  and the global loss function  $l(\Phi)$  are calculated as follows:

$$l_n(\Phi) = \frac{1}{s_n} \sum_{i \in s_n} f_i(\Phi) \quad (1)$$

$$\min_{\Phi} l(\Phi) = \sum_{n=1}^N \frac{s_n}{S} l_n(\Phi) \quad (2)$$

where  $f_i(\Phi)$  is the loss function of sample data  $i$  in the local dataset of edge node  $n$ .

The following is the construction of the FL process algorithm.

### 4.2 Local Algorithm Design

LSTM is the best choice for time series demand forecasting [14]. Data instability, data changes or irregularities, and fluctuating demand can be easily overcome using LSTM models. In 2014, Kyunghyun Cho [15] and others proposed a Sequence to Sequence (Seq2Seq) learning model structure, which consists of two parts: an encoder (Encoder) and a decoder (Decoder): an encoder as the input sequence. A fixed-length vector feature is generated, and another decoder that acts as this vector outputs the corresponding sequence (Fig. 4).

The framework consists of two parts: Encoder and Decoder. In the encoding phase, a value is input at each moment, and the hidden state changes according to the formula:

$$h_t = f(h_{t-1}, x_t) \quad (3)$$

$f(\cdot)$  is activation function. After reading each value of the sequence, a fixed-length semantic vector  $C$  is obtained:

$$C = q(h_1, h_2, h_3, \dots, h_t) \quad (4)$$

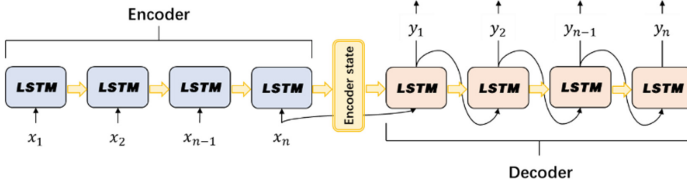


Fig. 4. Seq2Seq basic structure

In the decoding stage, the value  $y_t$  of the next output is predicted based on the given  $C$  and the output sequence  $y_1, y_2, y_3, \dots, y_{y-1}$ , and the probability of the target sequence is maximized as:

$$y_t = \operatorname{argmax} P(y_t) = \prod_{t=1}^T P(y_t | y_1, \dots, y_{t-1}, C) \tag{5}$$

Based on the Encoder-Decoder model, in order to eliminate the long-term dependency problem of the RNN model and inherit the memory function of the RNN, four LSTM networks are used to build a sequence-to-sequence model framework.

In this paper, we add the Attention mechanism to the original structure to break through the bottleneck limiting the performance of the model by inputting different semantic features  $c$  at each time.

Each  $c$  will automatically select the most appropriate context information for the current  $y$  to be output. Specifically, we use  $a_{i,j}$  to measure the correlation between  $h_j$  in the  $j$  stage in the Encoder and the  $i$  stage during decoding, and finally the input context information  $c_i$  of the  $i$  stage in the Decoder comes from all  $h_j$  pairs of  $a_{i,j}$  weighted sum.

## 5 Experiment and Analysis

### 5.1 Experimental Design

The M5 dataset, provided by Walmart, involves unit sales of various products sold in the United States, organized as a grouped time series. The products are sold in ten stores in three states (California, Texas and Wisconsin). The specific data set structure and the total amount of data for different commodities are shown in the figure (Fig. 5):

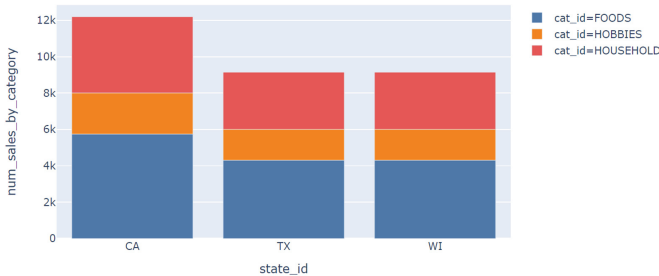


Fig. 5. Sales of goods by state in the dataset

We extracted multiple sets of sales data as learning tasks on the M5 dataset, evaluated our proposed FL time series algorithm, and compared the performance with other algorithms in the non-FL environment and in FL.

In the experiment, a single-machine pseudo-distribution is used to simulate two nodes, and the FedAvg algorithm is used. Each data set is stored in different local locations. The experimental environment is the Google colab platform. In this paper, we choose to optimize the parameters of the model through the Bayesian optimization algorithm.

In this paper, the LSTM model is built based on the Keras library, and the parameters are set as shown in the table through parameter optimization (Table 1).

**Table 1.** Parameter details

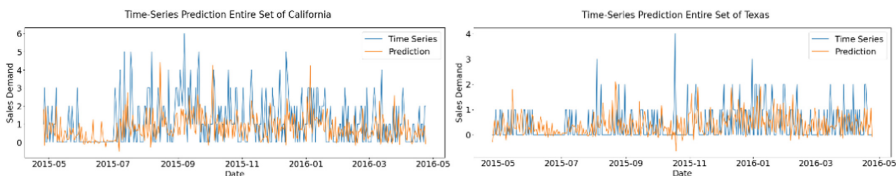
Property	Value
Learning_rate	0.005
Rounds	10
Batchsize	512
Local_epochs	2

This study uses FL to predict the data of each region for 30 days for the sales of multiple groups of the same product in different regions, that is, the model output length is 30. In this paper, we predict the research results through continuous training and empirical adjustment of parameters.

## 5.2 Simulation Analysis

The retail data in the M5 dataset is a stationary time series with seasonal characteristics from the time curve. For these commodities, customer buying needs are predictable, resulting in the most accurate short-term forecasts possible for expected top-up needs.

We use the Seq2Seq algorithm to predict the M5 data set, and in order to protect the commercial privacy of enterprises, we use FL to test data security (Fig. 6).



**Fig. 6.** Simulation of time series forecast results (California Texas)

The MAE, MSE and RMSE metrics were selected to evaluate the short-term sales forecast results of the Seq2Seq algorithm in California and Texas. The corresponding statistical index results are shown in the table:



At all prediction times, MAE results were generally small, and MSE was more biased in traditional Seq2Seq, but was well-received in other algorithms. At the same time, we can notice that under the FL method, the data in California has a better denial effect on the data in Texas, which reduces the prediction error (Tables 2 and 3).

**Table 2.** Algorithm error comparison (California-CA)

CA	Non-Fed		Fed	
	Seq2Seq	Seq2Seq-SCA	Seq2Seq	Seq2Seq-SCA
MAE	0.131025	-1.295403	0.008392	0.217017
MSE	1.428480	3.557133	1.491897	1.416363
RMSE	1.195190	1.886036	1.221432	1.190111

**Table 3.** Algorithm error comparison (Texas-TX)

TX	Non-Fed		Fed	
	Seq2Seq	Seq2Seq-SCA	Seq2Seq	Seq2Seq-SCA
mae	-0.231626	-0.321919	0.079820	0.288445
mse	0.916672	1.009948	1.579830	1.534100
rmse	0.957430	1.004962	1.256913	1.238588

It can be seen from the table that the MAE of extreme value prediction is much higher than that of the whole process, indicating that the overall prediction performance of the model is good, but the prediction ability of extreme value is slightly insufficient. This is mainly due to the fact that there are many factors affecting sales, such as seasonal changes, etc. Only considering historical sales data cannot learn the trend of sudden sales changes well.

The Seq2Seq model did not change significantly when the prediction length increased from 30 to 60. At the same time, compared with the FL scenario, the Non-FL scenario has more accurate predictions when the amount of data is sufficient, and when the amount of data is insufficient, since FL can introduce external data for reference, the prediction error of products with insufficient data is smaller than that of Non-FL. Error in FL scenarios. Therefore, in practice, it is necessary to select an appropriate prediction model according to the requirements for accuracy and efficiency.

## 6 Conclusion

This study proposes a method for short-term demand forecasting using FL and the seq2seq algorithm. This further illustrates the effectiveness of building models with FL compared to local dataset predictions. The main conclusions are as follows: (1) The prediction effect of the Seq2Seq-SCA model is better. (2) For commodities with sufficient data, the prediction results under non-FL scenarios are better than FL results;

when the amount of data is insufficient, the prediction results under FL are better than those under non-FL. (3) Seq2Seq-SCA is more suitable for combining with FL.

The retail forecasting model based on FL is a news attempt applied to SC optimization. The next step will be to improve the consideration of the learning time factor to improve the applicability of the model.

**Acknowledgements.** This work was supported in Major science and technology special project of Science and Technology Department of Yunnan Province (202002AB080001-8).

## References

1. Xu, S.Y., Chen, J., Wu, M.G.: E-commerce supply chain process optimization based on whole-process sharing of Internet of Things identification technology. *Cmes Comput. Mod. Eng. Sci.* **126**(2), 843–854 (2021)
2. Yu, W.F., Hou, G.S., Xia, P.C.: Supply chain joint inventory management and cost optimization based on ant colony algorithm and fuzzy model. *Tehnicki Vjesnik-Techn. Gazette* **26**(6), 1729–1737 (2019)
3. Bieniek, M.: The ubiquitous nature of inventory: vendor managed consignment inventory in adverse market conditions. *J. Oper. Res.* **291**(2), 411–420 (2021)
4. Ma, X.L., Wang, S.Y.: Coordinating a three-echelon fresh agricultural products supply chain considering freshness-keeping effort with asymmetric information. *Appl. Math. Model.* **22**(67), 337–356 (2019)
5. Hochreiter, S., Schmidhuber, J.: Long Short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
6. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. In: *NIPS* (2014)
7. Konečný, J., McMahan, H.B., Ramage, D., et al.: Federated optimization: distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527* (2016)
8. Xu, G., Li, H., Liu, S., et al.: VerifyNet: secure and verifiable federated learning. *IEEE Trans. Inf. Forensics Secur.* **15**, 911–926 (2019)
9. Xu, G., Li, H., Ren, H., et al.: Data security issues in deep learning: attacks, countermeasures, and opportunities. *IEEE Commun. Mag.* **57**(11), 116–122 (2019)
10. Nasr, M., Shokri, R., Houmansadr, A.: Comprehensive privacy analysis of deep learning: stand-alone and federated learning under passive and active white-box inference attacks. In: *IEEE Symposium on Security and Privacy*, pp. 56–65 (2018)
11. Zhang, X., Zhao, J., Xu, C., et al.: CIPPPA: conditional identity privacy-preserving public auditing for cloud-based WBANs against malicious auditors. *IEEE Trans. Cloud Comput.*, 122–143 (2019)
12. Zhang, Y., Xu, C., Ni, J., et al.: Blockchain-assisted public-key encryption with keyword search against keyword guessing attacks for cloud storage. *IEEE Trans. Cloud Comput.*, 231–247 (2019)
13. Colahs Blog: Understanding LSTM networks (2015)
14. Cho, K., Van Merriënboer, B., Gulcehre, C., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078* (2014)