Pharmaceutical Cold Chain Management Based on Blockchain and Deep Learning

Ping Zhang[1], Xuchong Liu[2], Wenjun Li[3], Xiaofeng Yu[4]

[1] College of Computer Science, Hunan University of Technology and Business, China
[2] Hunan Provincial Key Laboratory of Network Investigational Technology, Hunan Police Academy, China
[3] School of Computer and Communication Engineering, Changsha University of Science and Technology, China
[4] School of Business, Nanjing University, China
pingzp@hotmail.com,liuxuchong@163.com, lwjcsust@csust.edu.cn, xiaofengyu@nju.edu.cn

Abstract

Pharmaceutical cold chain is a special branch of the logistics industry, which has strict requirements for warehousing and transportation in the supply chain. In order to improve the credibility of pharmaceutical products and ensure the life safety of the people, it is necessary to realize the credible traceability of the entire life cycle of pharmaceutical products. The cost of pharmaceutical cold chain is high. Warehousing and transportation cost per unit of cold chain products is much higher than those of ordinary supply chain products. Intelligent prediction of cold chain product demand is an important way to reduce warehousing and transportation cost and improve market competitiveness of cold chain enterprises. Firstly, this paper proposes a pharmaceutical cold chain supervision scheme based on blockchain, cloud storage and Internet of things to realize the trusted traceability of the whole life cycle of pharmaceutical products and ensure the product safety. Then, based on the high-quality and large-scale data generated in the proposed cold chain supervision system, a cold chain product demand forecasting scheme based on deep learning is constructed to assist the cold chain inventory management decision-making, so as to reduce the warehousing cost of cold chain products.

Keywords: Internet of things, Blockchain, Deep learning, Pharmaceutical cold chain, Cloud storage

1 Introduction

Pharmaceutical cold chain is a branch of the logistics industry [1], which is a systematic project between the producer and the user of refrigerated medicine for the purpose of disease prevention, diagnosis and treatment. The entire pharmaceutical cold chain includes a series of links such as production, transportation, storage and use. The refrigerated medicine circulated in the pharmaceutical cold chain includes preventive biological products, therapeutic biological products and diagnostic biological products. Representative biological products include various vaccines, antisera (immune serum), antitoxin, toxoid, immune preparations (such as thymosin, immune nucleic acid, etc.), diagnostic reagents, etc.

Pharmaceutical cold chain management is facing unprecedented opportunities and challenges. With the development of medicine and biotechnology, the market of refrigerated pharmaceutical products is expanding, and the scale of pharmaceutical cold chain is increasing. The particularity of refrigerated pharmaceutical products makes the pharmaceutical cold chain have more stringent requirements in warehousing and transportation. The cold chain needs physical means to ensure the proper temperature conditions along the supply chain. It is necessary to realize the temperature control by using special storage, loading and unloading facilities and refrigeration devices. At present, the global pharmaceutical cold chain industry is still in the primary stage of development. The coverage of pharmaceutical cold chain is very low. Most of the quality problems of drugs are related to cold chain logistics, and the problem of cold chain breaking still occurs from time to time.

Drug safety is directly related to people's livelihood and social stability. Safety is the most important service capacity in pharmaceutical cold chain logistics. This is determined by the particularity of logistics objects. Pharmaceutical cold chain distribution needs to go through multiple logistics links, and different links use different transportation resources and information systems. These technologies and systems are independent to each other. The data and information are not integrated, and information control does not work. Therefore, a comprehensive tracking system is urgently needed to monitor the status of cold chain products in real-time and improve the effect and efficiency of cold chain logistics management.

In order to ensure the life safety of the people, it is

*Corresponding Author: Xiaofeng Yu; E-mail: xiaofengyu@nju.edu.cn
DOI: 10.53106/160792642021122207007

necessary to realize the credible traceability of the whole life cycle of pharmaceutical products and improve the credibility of pharmaceutical products. At present, the pharmaceutical cold chain logistics system is not perfect. Cold chain logistics technology and information level is not high, lack of overall planning and coordination of upstream and downstream. Each entity in the cold chain is independent of each other and maintains its own business data. Data may be tampered with, which will inevitably reduce the credibility of information.

Warehousing and transportation cost of pharmaceutical cold chain is very high. Drugs and biological products on the cold chain are sensitive to the environment temperature, so they need special materials and temperature control facilities. Compared with the general supply chain, the cost of cold chain warehousing and transportation is significantly increased. According to supply chain theory, improving the ability of cold chain inventory management is an important way to reduce the cost of cold chain enterprises and avoid overstocking of products. How to improve the cold chain inventory management ability through technical means becomes realistic and urgent [2].

This paper will focus on solving the above two challenges of safety and cost. First of all, we build a pharmaceutical cold chain supervision scheme based on blockchain, cloud storage and Internet of things (IoT) [3-4] to realize the reliable traceability of the whole life cycle of products and ensure the safety of products. Through the pharmaceutical cold chain supervision scheme, the overall planning and coordination of upstream and downstream are realized. In the process of supervision, a large number of high-quality, globally related data will be produced. Therefore, the second research topic of this paper is based on high-quality business data. We construct a cold chain product demand forecasting scheme based on deep learning to assist cold chain inventory management decision-making and reduce the warehousing cost of cold chain products [5].

The main contributions of this paper are summarized as follows:

- A pharmaceutical cold chain supervision scheme based on blockchain, cloud storage and IoT is proposed to realize the reliable traceability of products in the whole life cycle and ensure the safety of products.

Based on the high-quality regulatory data on blockchain and cloud storage, a cold chain product demand prediction scheme based on deep learning is proposed to assist cold chain inventory management decision-making and reduce the storage management cost of cold chain products.

We evaluate the proposed scheme theoretically and experimentally, where the evaluation results demonstrate its feasibility and effectiveness in both reliable traceability and demand prediction of pharmaceutical cold chain management.

2 Related Work

Blockchain is a secure distributed ledger, which can conduct trusted transactions across untrusted entities. Zhang et al. [6] analyzed blockchain transaction databases and proposes a storage optimization scheme. Philippe Jacquet and Bernard Mans [7] proposed a scheme to eliminate the burden of proof of work to solve the problem of energy waste in cryptocurrencies such as bitcoin. Quan et al. [8] reduces the amount of data storage of blockchain nodes by using down sampling technology. A high throughput, low latency and deterministic acknowledgement mechanism is proposed in [9] to accelerate the block acknowledgement mechanism of bitcoin. George Drosatos and Eleni Kaldoudi [10] reviewed the application of blockchain technology in biomedical field. Maxmen [11] discussed the possibility of applying blockchain to health research. Mathieu Boussard et al. [12] constructed a trust evaluation framework based on blockchain to solve the automated risk management of smart home network devices.

Supply chain refers to the network chain structure formed by upstream and downstream enterprises that provide products or services to end users in the process of production and circulation. Some researchers have studied the application of blockchain in supply chain. Wang et al. [13] proposed an optimized Merkle tree structure for efficient transaction verification in IIoT systems. Heinrich et al. [14] applied blockchain technology to the supply chain of high value botanical material. Rita Azzi et al. [15] studied the benefits of introducing blockchain into supply chain and the challenges encountered in the supply chain management ecosystem based on blockchain.

Internet of things is an ideal platform for monitoring the real-time status of goods in the cold chain. The IoT connects objects through the network and transmits information, so as to realize intelligent recognition of objects, positioning [16], etc. It provides real-time access to physical object information [17], when the large-scale data storage problem is handled properly [18]. V. Siris et al. [19] proposed a model to provide decentralized authorization for restricted IoT devices by using smart contract and interledger mechanism.

Artificial intelligence technology has been widely used in related fields. ARIMA [20] is a time series prediction model which combines moving average and autoregression. PSO-ELM [21] is a hybrid machine learning model based on Particle Swarm Optimization (PSO) and Extreme Learning Machine (ELM). XGBoost [22] is a scheme based on boosting tree. The deep learning model is complex, and thus needs to be trained based on big data [23].

3 Pharmaceutical Cold Chain Monitoring Based on Blockchain and IoT

In this section, we first give an overview of the proposed scheme, as well as some basic knowledge of pharmaceutical cold chain, and then introduces the pharmaceutical cold chain monitoring scheme based on blockchain and IoT.

3.1 Model Definition

Figure 1 is a simplified structure of the pharmaceutical cold chain. It mainly adopted in China. The management of vaccine distribution may not be the same in different countries. The entities in the pharmaceutical cold chain can be summarized as pharmaceutical manufacturers, distributors, hospitals and clinics, patients. In addition, there are government regulatory agencies, which are mainly responsible for supervising the production and consumption safety of pharmaceutical products. Each entity in the pharmaceutical cold chain is briefly described as following.

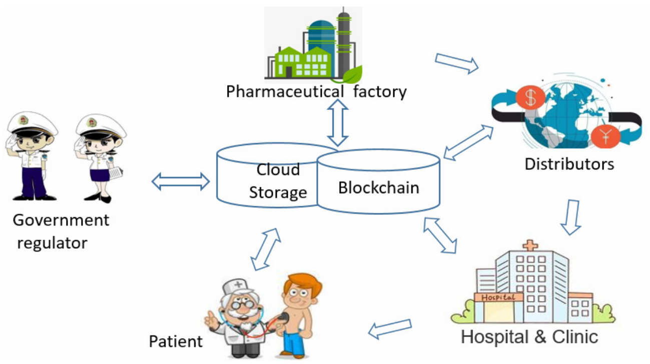


Figure 1. System overview

3.1.1 Pharmaceutical Manufacturers

Enterprises engaged in pharmaceutical production activities should obtain drug production licenses. Pharmaceutical manufacturers shall review and inspect the whole process of production and quality in accordance with regulations. The manufacturer shall establish a complete production quality management system, continuously strengthen the deviation management, and record all the data formed in the production and inspection process by means of information technology, so as to ensure the continuous compliance of the whole production process with legal requirements.

3.1.2 Distributors

A distributor is usually an intermediary who transfers goods from the manufacturer to the following entity. There are usually multi-level distributors in the market. The existence of distributors is conducive to promoting the circulation of pharmaceutical products and expanding the sales scope of products.

3.1.3 Hospitals and Clinics

Hospitals or clinics provide medical diagnosis services for patients, and prescribe and sell pharmaceutical products to patients according to the diagnosis results. Some types of drugs can also be sold in pharmacies. The proposed scheme can be extended to such cases.

3.1.4 Patients

Patients are individuals who ultimately purchase and use the product. For special groups, such as infants and young children, the legal guardian is usually responsible for handling related affairs.

3.1.5 Government Regulators

There are two types of government regulatory departments, one is pharmaceutical product quality supervision department. It mainly includes the drug regulatory department under the State Council and its authorized product approval and issuance agencies. The other is the health department of the State Council and its subordinate agencies.

3.2 Cold Chain Monitoring Based on Blockchain and IoT

Pharmaceutical cold chain monitoring needs to ensure the reliable traceability of pharmaceutical products in the whole life cycle. This includes three aspects. First, it is necessary to assign a trusted and unique ID for pharmaceutical products in the entire supply chain. By associating all kinds of data in the cold chain process with the unique ID, the relevant content can be traced. Secondly, it is necessary to guarantee trusted traceability of IoT monitoring data in the pharmaceutical cold chain. Third, we need to ensure the trusted traceability of business data in the process of pharmaceutical product circulation.

Blockchain has the characteristics of decentralization and non tampering, which can be used to solve the problem of data credibility in the process of pharmaceutical cold chain monitoring. However, the cost of data maintenance on the blockchain is too high. The amount of IoT monitoring data and business support data in the process of pharmaceutical product circulation is very large. Therefore, it is not suitable for direct storage in the blockchain. Therefore, we introduce a cloud storage module, which is responsible for data storage. The private key signature of the data before storage can prevent the data from being forged. By recording the data address and the hash value of these data in the blockchain, we can access the relevant data and verify its integrity.

3.2.1 Registration of Pharmaceutical Products

As shown in Figure 2, the registration of pharmaceutical products is initiated by pharmaceutical manufacturers. Pharmaceutical manufacturers use their private keys to sign the supporting business data related to the registration of pharmaceutical products (such as the approval data of the competent government department and the supporting data of the production process). Then the data is stored in the cloud storage module together with the signature. The cloud storage module will feed back the data storage address. Private key signature can prevent others from forging data.

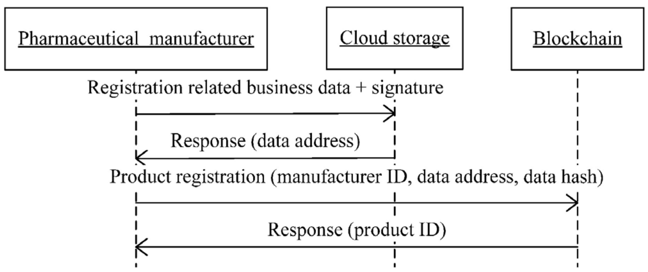


Figure 2. Registration of pharmaceutical products

Pharmaceutical manufacturers send pharmaceutical product registration requests to the blockchain platform, and send their own ID, supporting data storage address and data hash value. After verifying the data, the blockchain platform generates ID for pharmaceutical products. The received address and other data are written into the blockchain together with the product ID. By writing the data hash value into the blockchain record, we can prevent pharmaceutical manufacturers from tampering with the supporting business data.

3.2.2 Trusted Storage of IoT Monitoring Data

As shown in Figure 3, IoT monitoring mainly involves cold chain warehousing and cold chain transportation. IoT monitoring module is responsible for data storage to cloud storage module. When IoT monitoring module stores data, it needs to use its private key to sign the data to confirm the data source. The cloud storage module will feed back the storage address of the data.

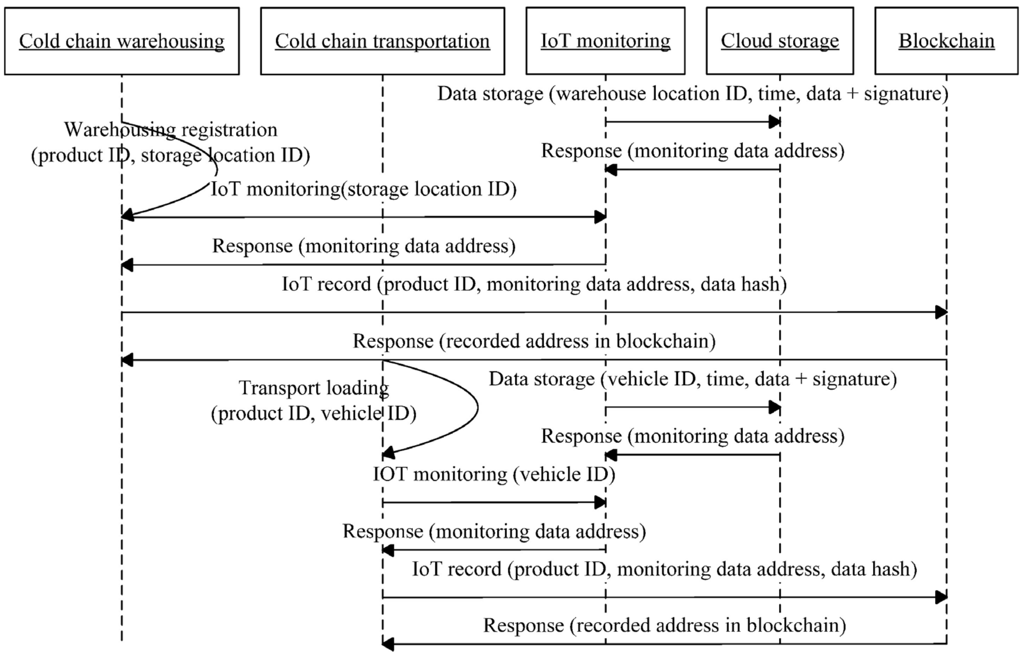


Figure 3. Trusted storage of IoT monitoring data

The main difference between warehouse monitoring and transportation monitoring lies in monitoring granularity and data storage organization mode. Warehouse monitoring is mainly organized according to storage location or shelf. Logistics transportation monitoring is mainly organized in the form of containers or vehicles. For example, in Figure 3, the storage location ID is uploaded when the data is storing. While the logistics transportation monitoring, we upload the vehicle ID when the data is storing.

When the cold chain warehouse registers the pharmaceutical products, it will assign the specific storage location for the pharmaceutical products in the warehouse, and thus establish the mapping between the pharmaceutical product ID and the storage location ID. This kind of data belongs to the business support data of pharmaceutical cold chain logistics, which also needs to be uploaded to the cloud storage module. The storage of business process data is described in the next section.

Cold chain warehouse requests the access address of IoT monitoring data from the IoT monitoring module based on the warehouse location ID and the storage time interval of pharmaceutical products. IoT monitoring module will feed back the corresponding monitoring data storage address according to the information provided by cold chain warehouse. Cold chain warehouse sends the pharmaceutical product ID, IoT monitoring data storage address and data hash value to the blockchain platform, and requests to write them to the blockchain platform. After verifying the data, the blockchain platform writes the data to the blockchain and feeds back the record address on the chain. The record address will be sent to stakeholders as part of the business data. Each stakeholder can find the record according to the record address of the blockchain, so as to obtain the access address and hash value of the monitoring data.

3.2.3 Trusted Storage of Cold Chain Business Data

In the monitoring process of cold chain logistics, a large number of business processes are involved, and a large number of business data will be produced. If these data are directly stored in the blockchain, the cost is too high. Therefore, we store these business data into cloud storage module. The storage address of business data is written into the blockchain, which will significantly reduce the amount of data in the blockchain.

As shown in Figure 4, there are three types of business entities involved in the cold chain monitoring sub process. They are the supplier, the cold chain transportation, and the receiver. All three of them generate different types and quantities of business data. The process of business data storage for different business entities is basically similar. They need to provide their own ID, business data, and use their private key to sign the business data. The storage address of business data and its hash value will be written to the blockchain platform. After the cloud storage platform completes data storage, it will feed back the storage address of the business data. The blockchain platform is responsible for feeding back the record address on the blockchain.

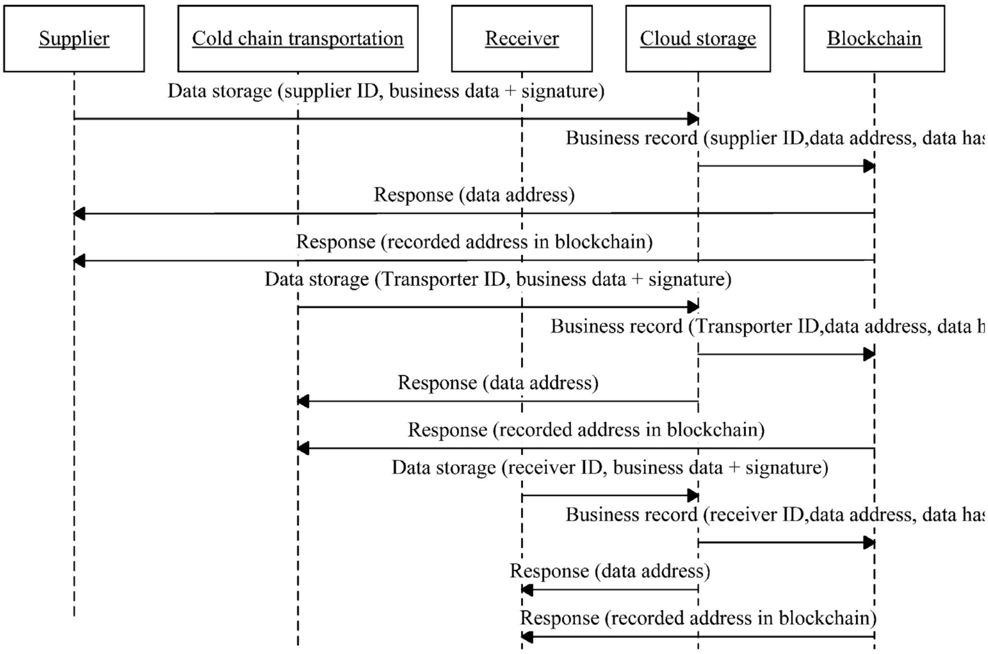


Figure 4. Trusted storage of cold chain business data

4 Inventory Management of Cold Chain Storage Based on Blockchain Data

This section first gives an overview of cold chain warehousing, and gives the framework of intelligent inventory management based on blockchain data. Then the inventory management problem is modelled and the solution is given. Finally, the feasibility of introducing deep learning model is discussed. A more detailed cold chain inventory management scheme based on deep learning will be introduced in the next section.

4.1 Overview of Cold Chain Inventory Management

Cold chain inventory management is very important in the pharmaceutical cold chain. According to modern supply chain theory, warehouse is the core of supply chain. Warehouse is the inventory control center in the supply chain. Warehouse is the dispatching center in logistics and supply chain. Warehousing is directly related to the efficiency and response speed of the supply chain.

Warehousing cost of cold chain is very high. On the one hand, if the inventory is insufficient, cross regional deployment of pharmaceutical products or urgent procurement will be triggered, which will lead to additional procurement costs and time delays, and the cost of cold chain transportation will also increase sharply. Especially when the upper level supplier also has the problem of insufficient inventory, the procurement cost and cycle will be greatly extended. As the pharmaceutical consumption demand can not be met in time, it will reduce the quality of service, lost existing customers or potential customers, and even have a negative impact on patients' health. On the other hand, if excessive procurement, it will increase the operating cost of cold chain warehousing. In serious cases, it will even lead to medicine overstock, resulting in a large number of expired medicine.

There are a large number of consumer groups at the end of the pharmaceutical cold chain, and the differences between individuals are large, which leads to the randomness of pharmaceutical demand. Pharmaceutical cold chain storage needs to be optimized according to demand changing of pharmaceutical consumption market, so as to improve the operation efficiency.

4.1.1 Intelligent Inventory Management Framework based on Blockchain Data

The intelligent inventory management process based on blockchain and cloud storage data is shown in Figure 5.

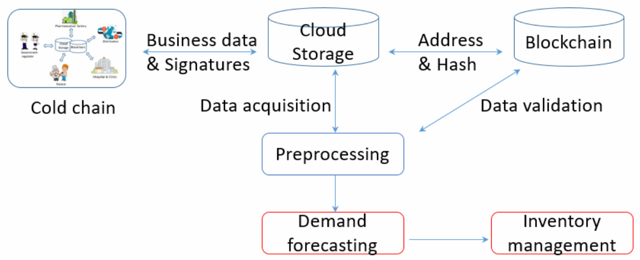


Figure 5. Intelligent inventory management process on blockchain data

Through the blockchain based cold chain monitoring module, each entity node in the supply chain uploads all kinds of business data on the supply chain to cloud storage. The data is signed by each entity node before uploading. The hash value and cloud storage address of the data are uploaded to the blockchain platform. Data signature mechanism makes other nodes unable to forge data. Data hash and blockchain technology make it impossible for entity nodes to tamper with data. Thus, credibility of cloud storage data is realized.

The blockchain based cold chain monitoring realizes the seamless docking of the whole process business data. This not only ensures the trusted traceability of the entire life cycle of the product, but also brings convenience for the extraction and preprocessing of the training data. High quality and mass data stored in cloud storage platform and blockchain provides the possibility for intelligent application.

Data preprocessing module obtains data from cloud storage and blockchain respectively. The data on the blockchain is mainly used for data verification. After data cleaning and sorting, the data can be used as training data to train various prediction models.

The complete solution of the inventory management problem mainly consists of two parts: one is demand forecasting, and the other is inventory management. As mentioned above, the performance of demand forecasting directly determines the effect of inventory management. The rest of this section will present the model and solution of inventory management problems, and the next section mainly introduces the solution of demand forecasting problem.

4.2 Modeling and Solving of Inventory Management Problem

The basic problems of warehouse inventory management are as follows: in the context of uncertain demand, an optimal decision is made on the storage inventory to minimize the expected cost in a single order cycle.

The expected cost C mainly consists of four parts: fixed cost c_1, variable cost c_2, various losses caused by insufficient inventory resources c_3, and cost increase caused by excessive inventory c_4. It can be described by the following formula.

$$C = c_1 + c_2 + c_3 + c_4$$

Fixed cost c_1 is the basic cost to maintain the normal operation of the business. The variable cost can be expressed as $c_2 = \alpha(s - s_0)$, where s_0 is the inventory at the beginning of the decision cycle, s is the inventory of the current decision cycle, and α is the unit inventory purchase cost. All kinds of losses caused by insufficient inventory resources can be expressed as $c_3 = \beta(\max(y - s, 0))$, where β is the loss value caused by insufficient inventory per unit, and y is a random variable used to describe uncertain demand. All kinds of costs caused by excessive inventory can be expressed as $c_4 = \gamma(\max(s - y, 0))$, where γ is the cost increase value caused by excessive inventory per unit. The optimal inventory management decision is an expected cost minimization problem.

$$s^* = \mathrm{argmin}\ C$$

This is a classical newsvendor problem, which can be solved by the standard first-order condition of the classical newsvendor problem. The optimal result is

$$s^* = \phi^{-1}(\frac{\beta-\alpha}{\beta+\gamma}),$$

where ϕ is the cumulative distribution function of y.

The key to solve the above problems is how to obtain the cumulative distribution function of y. There are some challenges in this issue. First of all, there are many factors that affect the demand, and there is a complex nonlinear relationship between these factors and demand, so it is difficult to construct a complex machine learning model. Secondly, to improve the prediction accuracy of complex model, a large number of high-quality training data are needed. Next, we will introduce the basic ideas to solve these challenges, and briefly analyze its feasibility.

4.3 The Feasibility of Introducing Deep Learning Model

4.3.1 Challenges of Deep Learning Model

Deep learning has the ability to solve complex nonlinear problems. There are many factors affecting cold chain storage, including demand side, supply side, logistics side, market share change and so on. The influence of these factors is highly nonlinear in nature, and deep learning realizes the nonlinear transformation from input to output, which is one of the important reasons for deep learning to make breakthroughs in many complex problems. Deep learning is more powerful in end-to-end learning, that is, it does not need to divide artificial steps or subproblems, but is completely handed over to the neural network to directly learn the mapping from the original input to the desired output. Compared with divide and conquer strategy, end-to-end learning has the advantage of synergy, and it is more likely to obtain better global solutions.

However, there are also challenges brought by the complexity of the model when using deep learning to predict demand. According to machine learning theory, more training data is needed for complex models. With the increase of the number of model variables, higher requirements are put forward for the amount of training data.

The increase demand of training data brought by the complexity of deep learning cannot be the reason for us to give up the deep learning model. The training process of machine learning uses the model to fit the data. However, the purpose of machine learning is not to fit the training data set correctly, but to correctly predict the samples that have not appeared in the training set. There are many factors that affect cold chain storage and are highly nonlinear, which means that the model used should have a certain degree of complexity, only in this way can these factors be expressed. However, the use of complex models requires increasing the amount of data for training. Therefore, the increase of data demand is caused by the complexity of the problem, not by the complexity of deep learning model.

4.3.2 Deal with Challenges

The historical data accumulated by the daily business process of the node to be studied is most closely related to the node, so it is the best to use these data to predict the demand of the node directly. However, the amount of business data of a single node is not large enough, and the short-term accumulation of historical data cannot meet the demand of complex model for the amount of training data.

Theoretically, there is a direct or indirect relationship between different regions in the same market. For example, there are similar consumption patterns among users in different regions. At the same time, the same brand may have similar market influence in different regions. In addition, the relationship between upstream and downstream entities of the same supply chain is very close. Therefore, at least in theory, it is feasible to increase the number of training data by using the same correlation contained in the data between different entities.

The blockchain based monitoring architecture can satisfy the demand of data quantity and data quality in the deep learning training process. On the one hand, it is reflected in the data quantity. Based on the traceability system constructed by blockchain, the relationship between the data of the whole supply chain is established. It makes the data of each node in the supply chain no longer isolated, and it is more convenient to obtain the data. We can increase the amount of training data by introducing data from other nodes in the supply chain. This will greatly increase the training performance. On the other hand, it is reflected in the data quality. The introduction of blockchain technology solves the problem of data credibility and ensures the quality of training data.

5 Cold Chain Storage Demand Forecasting Based on Deep Learning

In the previous section, we modeled the inventory management problem and gave the solution. To improve the accuracy of the solution, the key is to forecast and measure the demand and uncertainty of cold chain storage. This section will introduce the solutions of cold chain storage demand forecasting and demand uncertainty quantitative measurement.

5.1 Cold Chain Storage Demand Forecasting

As shown in Figure 6, the whole demand forecasting model consists of three modules: convolutional neural network (CNN) [24], Long Short-Term Memory (LSTM) [25] module and graph embedding module.

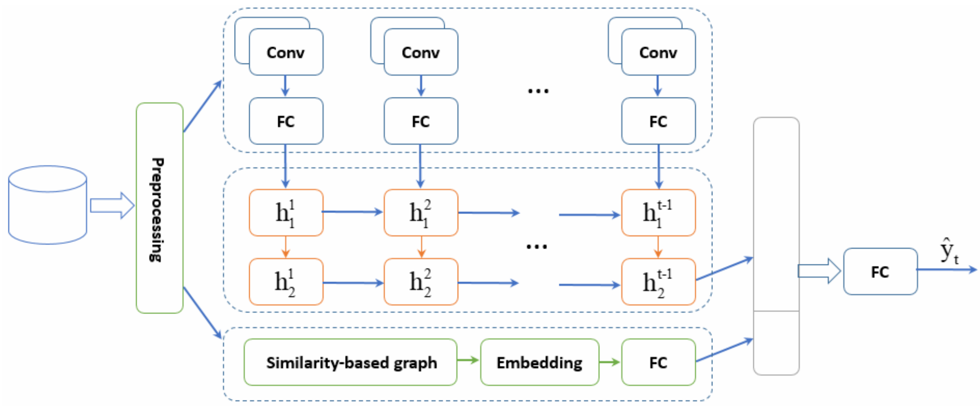


Figure 6. Cold chain storage demand forecasting based on deep leaning

CNN module is used to learn the business data of cold chain products from the perspective of spatial correlation. LSTM module is used to learn the business data of cold chain products from the perspective of time relevance. Graph embedding module is used to break through the constraint of spatial distance, and establish connection for nodes with long distance but have a similar business model.

5.1.1 CNN Module

First, data preprocessing. We divide the whole city into several disjoint rectangular regions, and integrate the data in each region to obtain the spatial dimension data set. In the spatial dimension, each point contains the business data aggregation of the regional location. It should be noted that since most business data is high-dimensional data, spatial dimension data is actually composed of multiple spatial dimension data pieces. The spatial dimension data is then organized in the form of time series, and the size of time slice is 1 day.

Second, the preprocessed data are input into CNN module in chronological order. CNN module (i.e., Conv in Figure 6) extracts spatial correlation features from data by convolution layers. The data processed by convolution layers will be converged through full connection layer (i.e., FC in Figure 6) to realize docking with LSTM module.

5.1.2 LSTM Module

LSTM module has good long-term memory ability. We use LSTM module to extract time-related features. In the previous step, we have extracted the spatial correlation features of data from different time nodes by CNN module. After feature extraction of spatial data at different times by CNN module, we can get a convolution layer feature sequence organized in time sequence. These convolution feature sequences are sent to LSTM module (as shown in Figure 6, h_i^j is the state tensor of LSTM module) in time order to extract the time-related features contained in the sequences.

5.1.3 Graph Embedding Module

According to common sense, we can judge that the spatial correlation between nodes is not entirely determined by the distance of geographical location. For example, readers can compare the following two situations: first, two areas far apart, but both belong to the area where residents are more concentrated; second, two areas close to each other, but one is residential area and the other is industrial area. Obviously, we have reason to believe that the similarity of demand patterns between the two regions in the first case is higher.

In this scheme, we will capture the similarity in distance by graph embedding. It includes the following steps:

First, the similarity graph is constructed (i.e., similarity-based graph in Figure 6). We construct an undirected graph with all regions as vertices and the similarity between regions as weights. According to the demand of each region at each time, a demand sequence can be obtained, and the similarity between different regions can be obtained by dynamic time warping (DTW) algorithm.

Second, graph embedding (i.e., embedding in Figure 6). Based on the similarity graph obtained in the previous step, we use graph embedding to transform it. This method can construct a vector for each region, which describes the spatial similarity between nodes.

Finally, through a full connection layer (i.e., FC in Figure 6), the vector is transformed and integrated with the output of LSTM module.

5.2 Demand Uncertainty Measurement

According to the inventory management problem model and solution given in the previous section, in order to achieve the optimal decision of inventory management, it is necessary to quantitatively measure the uncertainty of demand to obtain its cumulative probability distribution function. More specifically, under the assumption of Gaussian model, we need to calculate the mean and variance of uncertain demand quantitatively.

The mainstream deep learning technology cannot express the model uncertainty. These traditional neural networks usually use maximum likelihood estimation or maximum a posteriori to train and optimize the parameters, which ignores the uncertainty. Therefore, it often produces a point estimation instead of the probability distribution of uncertain value. In these schemes, the most common method of quantitative measurement uncertainty is using the softmax function to obtain the probability, but this often leads to great training deviation.

Bayesian model has a complete mathematical framework, which can derive model uncertainty. But the uncertainty estimation algorithm based on Bayesian neural networks is very complex.

In this paper, we adopted a dropout-based method to estimate uncertainty [26]. Dropout is a common regularization technique in the field of deep learning. It is often used to solve the over fitting problem in the neural networks training. The basic idea of dropout is to make the neurons in the hidden layer stop working with a certain probability in the forward propagation process. In this way, the model will not rely too much on some local features, so as to improve the generalization ability of the model.

Yarin Gal and Zoubin Ghahramani [26] regard dropout training in deep neural network as approximate Bayesian reasoning in depth Gaussian process, and use the dropout NN to model uncertainty. This method can reduce the computational complexity and does not affect the accuracy of the estimation results.

The prediction probability based on depth Gaussian model is described as follows:

$$\mathrm{p}(y_t \mid x_t, X, Y) = \int p(y_t \mid \omega) p(\omega \mid x_t, X, \mathrm{Y}) d\omega$$

where ω is the random variables for a model with L layers. X and Y are input and output datasets, respectively.

Based on the above prediction probability distribution. The mean and variance can be calculated as follows:

$$\mathrm{E}(\mathrm{y_t}) = \frac{1}{T}\sum_{t=1}^{T} \hat{y}_t(x_t)$$

$$Var(y_t) = \tau^{-1} I_d + \frac{1}{T}\sum_{t=1}^{T} \hat{y}_t(x_t)^{\mathrm{T}} \hat{y}_{\mathrm{t}}(x_t) - \mathrm{E}(\mathrm{y_t})^T \mathrm{E}(\mathrm{y_t})$$

where T is the number of samples and τ is the model precision. For more details, please refer to the original paper.

6 Theoretical Analysis and Experimental Evaluation

6.1 Analysis of Credible Traceability Function

6.1.1 Data Credibility

First, data forgery and denial. All supporting data are stored in cloud computing platform, and private key signature is required before data storage. On the one hand, it can prevent the malicious third party from forging the data. On the other hand, it can also avoid the data uploader from denying the data.

Second, data tampering. A malicious third party can't tamper with data without a private key. Once there is a drug safety incident, legitimate users have the motivation to tamper with data to avoid responsibility. However, the hash value of the data is stored on the blockchain. Once the data is tampered with, the hash value will change. Therefore, data tampering is easy to be detected by users.

6.1.2 Traceability Issues

In order to realize the traceability function, we should have the ability to access both the business data and the IoT monitoring data. Through the registration of pharmaceutical products on the blockchain, pharmaceutical products are assigned a unique ID. Both the business data and the IoT monitoring data are retrieved by using this product ID.

During the trusted storage of cold chain business data, three types of business entities (i.e., the supplier, the cold chain transportation, and the receiver) involve the product ID in their business data. As a result, we can retrieve business data of a specific product through its ID.

The warehouse monitoring data is organized according to the storage location ID. When products are registered in a cold chain warehouse, they will be stored in a specific storage location, and thus establish the mapping relationship between the storage location ID and the produce ID. Through the produce ID and the in-warehouse time, we can obtain the corresponding storage location ID, and then we can access related monitoring data by using the storage location ID.

The cold chain transportation monitoring data is organized in the form of vehicle ID, and the mapping relationship between vehicle ID and produce ID will be established during transportation loading process. Through the produce ID and the transportation time, we can obtain the corresponding vehicle ID, and then we can access related monitoring data by using the vehicle ID.

6.2 Experimental Setup

In order to verify the performance of the cold chain storage demand forecasting scheme, we conducted an experimental evaluation based on real data. The data set used is the business data of company T in Changsha. T company is mainly engaged in medical related cold chain logistics services. The time span of the data set is from January 1, 2016 to December 31, 2019, a total of three years. The minimum granularity of data is daily.

The benchmark schemes include ARIMA [20], PSO-ELM [21] and XGBoost [22]. The experimental evaluation metric include Mean Average Percentage Error (MAPE) and Rooted Mean Square Error (RMSE), which are defined as follows.

$$\mathrm{MAPE}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\frac{|\hat{y}_i - y_i|}{y_i}$$

$$RMSE=\sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}(\hat{y}_i - y_i)^2}$$

where N is the total number of samples. $\hat{y}_i$ and y_i are the predicted value and the real value, respectively.

6.3 Performance Comparison with Different Schemes

Table 1 shows the performance comparison results between the proposed scheme and the benchmark schemes. In this experiment, the proposed scheme has the best performance. ARIMA has the worst performance in these schemes. ARIMA completely relies on the historical demand value, and does not consider the nonlinear and spatial correlation. The performance of PSO-ELM and XGBoost are significantly improved compared with ARIMA. PSO-ELM adopts heuristic and population-based optimization techniques. XGBoost considers context features. Comparatively speaking, XGBoost performs better than the other two benchmark schemes. However, the performance of XGBoost is still worse than that of the proposed scheme in this experiment.

Table 1. Performance comparison with different schemes

Schemes	MAPE	RMSE
ARIMA	0.2531	12.352
PSO-ELM	0.2103	11.212
XGBoost	0.1986	10.926
The proposed scheme	0.1752	9.782

6.4 Performance under Different Settings

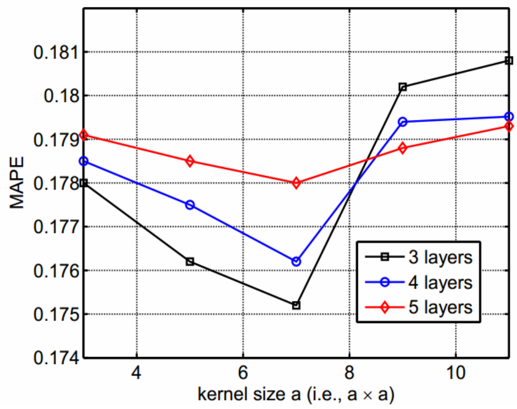
This section mainly considers the impact assessment of two types of settings, one is different parameter, the other is different structure.

6.4.1 Performance under Different Parameter

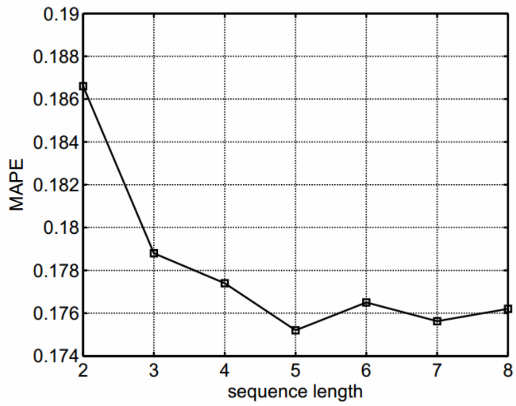
We adjust the convolution kernel size, the convolution layer number and the LSTM sequence length to study the effect of different parameter settings on the performance of the proposed scheme.

First of all, we study the influence of the convolution kernel size and the convolution layer number on the performance of the scheme. The experimental results are shown in Figure 7(a). The experimental results show that these two parameters have a significant impact on the performance of the scheme. In this experiment, when the size of convolution kernel is 7×7 and the number of convolution layers is 3, the proposed scheme achieves the best performance.

Then, we study the influence of LSTM sequence length on the performance of the scheme. The experimental results are shown in Figure 7(b). The experimental results show that the length of LSTM sequence also has a significant impact on the performance of the scheme. In this experiment, our method achieves the best performance when the length is 5 days. If the length value continues to increase, the performance change is relatively small.



(a) different convolution kernel size and layer number



(b) different LSTM sequence length

Figure 7. Performance comparison under different parameter

6.4.2 Performance under Different Structure

We compare the proposed scheme with the one without the graph embedding module. The comparison results are shown in Table 2. Without the embedding module, although it can still extract features from both temporal and spatial dimensions, the performance is significantly reduced. We think that the spatial correlation not only exists between adjacent regions, but also exists between distant regions. The similarity between regions extracted by graph embedding module is beyond the limit of region. It can effectively explore such correlation existed in distant regions, and thus has has a significant impact on the performance of the proposed scheme.

Table 2. Performance comparison under different structure

Schemes	MAPE	RMSE
The proposed scheme (without Embedding)	0.1835	10.093
The proposed scheme	0.1752	9.782

7 Conclusion

This paper studies the management of pharmaceutical cold chain, focusing on safety and cost. We introduce the blockchain technology to achieve reliable traceability of pharmaceutical products and improve the safety of products. In order to solve the data storage problem of blockchain technology, we introduce cloud storage module in the process of pharmaceutical cold chain supervision. In order to achieve high-precision demand prediction of pharmaceutical products and improve the quality of inventory management, we introduce deep learning technology to mine the spatiotemporal characteristics of the pharmaceutical cold chain data accumulated in blockchain and cloud storage.

Acknowledgements

This work is supported by the Humanities and Social Science Youth Fund of Ministry of Education of China (19YJCZH254), the Scientific Research Fund of Hunan Provincial Education Department (19B315, 20C0538), the Innovation driven plan project of Hunan University of Technology and Business in 2020, and the Open Research Fund of Hunan Provincial Key Laboratory of Network Investigational Technology (2018WLZC006), the Project of Hunan Social Science Achievement Appraisal Committee (XSP21YBZ113).

References

[1] M. Amini, C. C. Bienstock, M. Golias, Management of supply chains with attribute-sensitive products: A comprehensive literature review and future research agenda, *International Journal of Logistics Management*, Vol. 31, No. 4, pp. 885-903, November, 2020.

[2] S. Guo, T.-M. Choi, B. Shen, S. Jung, Inventory management in mass customization operations: A review, *IEEE Transactions on Engineering Management*, Vol. 66, No. 3, pp. 412-428, August, 2019.

[3] W. Li, H. Xu, H. Li, Y. Yang, P. K. Sharma, J. Wang, S. Singh, Complexity and algorithms for superposed data uploading problem in networks with smart devices, *IEEE Internet of Things Journal*, Vol. 7, No. 7, pp. 5882-5891, July, 2020.

[4] P. Zhang, J. Wang, W. Li, A learning based joint compressive sensing for wireless sensing networks, *Computer Networks*, Vol. 168, Article No. 107030, February, 2020.

[5] J. Jemai, B. D. Chung, B. Sarkar, Environmental effect for a complex green supply-chain management to control waste: A sustainable approach, *Journal of Cleaner Production*, Vol. 277, Article No. 122919, December, 2020.

[6] J. Zhang, S. Zhong, J. Wang, X. Yu, O. Alfarraj, A Storage Optimization Scheme for Blockchain Transaction Databases, *Computer Systems Science and Engineering*, Vol. 36, No. 3, pp. 521-535, 2021.

[7] P. Jacquet, B. Mans, Green mining: Toward a less energetic impact of cryptocurrencies, *International Conference on Computer Communications*, Paris, France, 2019, pp. 210-215.

[8] L. Quan, Q. Huang, S. Zhang, Z. Wang, Downsampling blockchain algorithm, *International Conference on Computer Communications*, Paris, France, 2019, pp. 342-347.

[9] A. Hari, M. Kodialam, T. V. Lakshman, Accel: Accelerating the bitcoin blockchain for high-throughput, low-latency applications, *International Conference on Computer Communications*, Paris, France, 2019, pp. 2368-2376.

[10] G. Drosatos, E. Kaldoudi, Blockchain applications in the biomedical domain: A scoping review, *Computational and Structural Biotechnology Journal*, Vol. 17, pp. 229-240, 2019.

[11] A. Maxmen, Ai researchers embrace bitcoin technology to share medical data, *Nature*, Vol. 555, No. 7696, pp. 293-294, March, 2018.

[12] M. Boussard, S. Papillon, P. Peloso, M. Signorini, E. Waisbard, Steward:sdn and blockchain-based trust evaluation for automated risk management on iot devices, *International Conference on Computer Communications*, Paris, France, 2019, pp. 841-846.

[13] J. Wang, B. Wei, J. Zhang, X. Yu, P. K. Sharma, An Optimized Transaction Verification Method for Trustworthy Blockchain-enabled IIoT, *Ad Hoc Networks*, Vol. 119, Article No. 102526, August, 2021.

[14] M. Heinrich, F. Scotti, A. Booker, M. Fitzgerald, K. Y. Kum, K. Lobel, Unblocking high-value botanical value chains: Is there a role for blockchain systems? *Frontiers in Pharmacology*, Vol. 10, Article No. 396, pp. 1-10, April, 2019.

[15] R. Azzi, R. K. Chamoun, M. Sokhn, The power of a blockchain-based supply chain, *Computers and Industrial Engineering*, Vol. 135, pp. 582-592, September, 2019.

[16] W. Li, Z. Chen, X. Gao, W. Liu, J. Wang, Multimodel framework for indoor localization under mobile edge computing environment, *IEEE Internet of Things Journal*, Vol. 6, No. 3, pp. 4844-4853, June, 2019.

[17] P. Zhang, J. Wang, On enhancing network dynamic adaptability for compressive sensing in wsns, *IEEE Transactions on Communications*, Vol. 67, No. 12, pp. 8450-

8459, December, 2019.
[18] J. Wang, Y. Yang, J. Zhang, X. Yu, O. Alfarraj, A. Tolba, A Data-Aware Remote Procedure Call Method for Big Data Systems, *Computer Systems Science and Engineering*, Vol. 35, No. 6, pp. 523-532, 2020.
[19] V. A. Siris, D. Dimopoulos, N. Fotiou, S. Voulgaris, G. C. Polyzos, Interledger smart contracts for decentralized authorization to constrained things, *International Conference on Computer Communications*, Paris, France, 2019, pp. 336-341.
[20] T.-M. Choi, W.-K. Yeung, T. C. E. Cheng, Scheduling and co-ordination of multi-suppliers single-warehouse-operator single-manufacturer supply chains with variable production rates and storage costs, *International Journal of Production Research*, Vol. 51, No. 9, pp. 2593-2601, 2013.
[21] L. Lin, J. C. Handley, Y. Gu, L. Zhu, X. Wen, A. W. Sadek, Quantifying uncertainty in short-term traffic prediction and its application to optimal staffing plan development, *Transportation Research Part C: Emerging Technologies*, Vol. 92, pp. 323-348, July, 2018.
[22] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, *Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785-794.
[23] J. Wang, Y. Yang, T. Wang, R. S. Sherratt, J. Zhang, Big Data Service Architecture: A Survey, *Journal of Internet Technology*, Vol. 21, No. 2, pp. 393-405, March, 2020.
[24] J. Ke, H. Yang, H. Zheng, X. Chen, Y. Jia, P. Gong, J. Ye, Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 11, pp. 4160-4173, November, 2019.
[25] E. Choi, S. Cho, D. K. Kim, Power demand forecasting using long short-term memory (lstm) deep-learning model for monitoring energy sustainability, *Sustainability*, Vol. 12, No. 3, Article No. 1109, February, 2020.
[26] Y. Gal, Z. Ghahramani, Dropout as a bayesian approximation: Representing model uncertainty in deep learning, *International Conference on Machine Learning,* New York, NY, USA, 2016, pp. 1-10.

Biographies

Ping Zhang received his Ph.D. degree in computer science from Central South University, China. He is an Associate Professor with the College of Computer Science, Hunan University of Technology and Business. His current research interests include network, security, and machine learning.

Xuchong Liu received his Ph.D. degree from Central South University in 2010. He is vice-principal and professor in Hunan Police Academy, and his research interest includes network security, big data analysis, and electronic evidence.

Wenjun Li received his Ph.D. degree in Computer Science from Central South University in 2014. He is currently an associate professor in the School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, China. His research interests include algorithm analysis and optimization, and artificial intelligence.

Xiaofeng Yu received his Ph.D. degree from Nanjing University, in 2007. He is currently an associate professor in the School of Business, Nanjing University. His current research interests include strategy, technologies and applications of Electronic Business.