**A Novel Virtual Network Fault Diagnosis Method Based on Long Short-Term Memory Neural Networks**

Lei Zhang, Xiaorong Zhu, Su Zhao and Ding Xu

Nanjing University of Posts and Telecommunications

Nanjing, China

zhentoufei@163.com, xrzhu@njupt.edu.cn, zhaos@njupt.edu.cn, xuding@njupt.edu.cn

*Abstract*—Network virtualization has emerged as a significant trend to solve the issue caused by ossification of traditional network. Under the circumstance of network virtualization, substrate network and virtual networks are inextricably interdepending to each other. The substrate network serves many virtual networks. Substrate network faults lead to different virtual network faults. A service’s failure may introduce additional influence on other services. Therefore, it has become a big challenge to predict when and where a fault happens in the network. In this paper, we propose a fault diagnosis method by deep learning to predict the failure of virtual networks. Our deep learning model enables the earlier failure predictions by the Long Short-Term Memory (LSTM) network which discovers the long-term features of network history data. Simulation results show that the proposed method performs well on faults prediction.

Keywords—network virtualization; fault analysis; deep learning; LSTM

# Introduction

Network virtualization is the evolution direction of future network, which is also an import approach to solve the rigid management of present network. The core idea of network virtualization is to realize the refinement and division of traditional network by flexible abstraction, allocation and isolation of the substrate physical infrastructure resources.

The infrastructure providers (InPs) provide the substrate network resources for service provider (SPs) to build up virtual networks for different services, and SPs deploy services and applications on virtual networks. With the continuous emergence of various types of SPs, more and more applications running on the virtual networks. Mapping mechanism is becoming increasingly complex. Therefore, in the network virtualization environment, when a virtual network has a fault symptom, the underlying network will be affected, and then indirectly affect the virtual parts which are mapped to the components with a fault.

Fault diagnosis is important in virtual network fault management. Fault diagnosis models are designed to discover suspicious symptoms that may lead to network faults and provide fault warnings for the redistribution of virtual network resources and ensure the QoS(Quality of service).

There are some diagnostic methods in the area of network virtualization. Jian [1] proposes a method of multi-dimensional fuzzy association mining rule for fault diagnosis of communication network based on relationship between business symptoms and network faults. Lin [2] proposed a fast convergence algorithm based on fuzzy reasoning to sense the law of network signal, which solves the network fault caused by transmission signal error. Matrix algorithm [3], neural network algorithm [4], Bayesian algorithm [5] are also used in the distributed network fault location, which have some reference value for the virtual network faults diagnosis. The fault diagnosis technology research literatures are relatively few about network virtualization. Pan [6] proposed a new fault diagnosis method based on feature reasoning against the challenges of network virtualization. Liu [7] presents an algorithm for internal evaluation of virtual network and substrate network faults based on trust evaluation. In [8], a service fault propagation model is proposed, which reduces the error probability of the diagnosis algorithm.

Furthermore, the present researches on network failure mainly focus on ensuring the rapid recovery of services when failures happen. Backup strategy of the protection policy leads to the low utilization of network resources. The recovery strategy can degrade the service quality if the recovery is not timely.

In order to solve the above problems, we present a virtual network fault diagnosis model based on LSTM. To analysis the network faults, time dimension information is considered. According to the changes of the network parameters in the past period of time, LSTM neural network is capable of learning long-term dependencies information of the network states. This method focuses on the parameter information of previous periods and the fault labels of the next, reducing the complexity of modeling process. In the network virtualization environment, we tries to ensure the QoS of the network through faults alarms before the fault happens so that SPs have enough time to be prepared for the network faults.

The rest of the paper is organized as follows. Section II describes the vitalization network fault diagnosis problem and the background knowledge of LSTM. In section III, we propose deep learning model for fault diagnosis. Section IV presents the experiments conducted on the LSTM with the use of OPNET simulation data, analyzes the experiment results and makes a discussion about the deep learning model performance. Section V concludes the paper.

# Problem Description and Background Knowledge

## Network Virtualization Fault Diagnosis

This paper is to use the physical network and virtual network’s history normal and fault data to predict virtual network failure level probability in the future. The network structure is shown in Fig. 1. The servers have the ability to handle some faults, when a crash occurs, the servers can ensure the normal operation of the systems with the help of security procedures. For routers, due to limited computing resources, some small failures may lead to system paralysis. Different faults have different influence on different types of network components.

Luo [9] points that assuming there are failures in a system, when an error occurs in the system, the possibility of the system running normally for time is , where is an error parameter. When an error occurs, with time passing by, the longer the system’s error-free running time is, the more possibly another error happens.

Based on the above analysis, the parameters in the network virtualization structure can be divided into topological , performance , function and statistical as is shown below:

where:

is the location information of network components, usually represented by the components’ identifiers. represents the dependency of network components. is the adjacency information of the network components. indicates connectivity for the network components. For a virtual node, it is the number of the links connected to the node; the virtual link is two, indicating that it always connects with two nodes. is the bandwidth. indicates network delay. indicates the packet loss rate. indicates the network transmission rates. indicates the types of network components, like routers, switches, servers, links and so on. indicates the network components’ functions like forwarding, data processing, etc. indicates the set of network services provided by the network components. indicates the security level required. indicates the counts of network component failures. indicates the historical numbers of different types of failures group by components. indicates the collection of historical errors types of network components. indicates



1. Network Structure

the error-free time of the system from the last fault recover or from the start of the system if there are no faults occurs.

Fig. 2 shows how we define the network train data set and fault labels. At time , we can get a matrix which contain data of time length. For history data at time , the network state at time is seen as the data label which means that at time , the network’s fault is . Assuming our model is , we have the following formulation:

The data is collected at time and is the network fault states to be diagnosis in next period time .

## Long Short-Term Memory netwwork

Recurrent neural networks (RNNs) is the network containing a loop in internal networks for ensuring information persistence. RNNs have full connections between adjacent layers as well as among the nodes within the same layers. In addition, hidden units in RNNs receive a feedback from the previous states to current states. These features are suitable to handle with the temporal–spatial data [10].

LSTM network is a kind of RNNs. LSTM neural network can deal with the correlation problems of time series in both short and long term, using the hidden layer as a memory unit. Especially, LSTM performs well when handling with long sequence dependency problems avoiding the vanishing



1. Network data and faults label



1. Structure of LSTM cell

gradient problem in RNNs. Fig. 3 shows the structure of LSTM cell. The memory block contains a self-connected statecell, an input gate, an output gate and a forget gate, and , , is *sigmoid*, *tanh*, *tanh*[11].

We have the following definition. is the weight from unit to unit . The subscripts , , refer respectively to the input gate, forget get and output gate. The subscripts refers to one of the memory cells while is the -th cell in a layer.

In the memory structure, the activation vector of the input gate is:

is input vector. is the output from the previous cells. is the cell states in the previous time step. is the bias. The activation vector of the forget gate can be expressed as follows:

The status values of the memory cells are updated as follows:

The output of the memory cell is ultimately controlled by the activation values of the output gates:

The final LSTM unit output value is follow:

# Deep Learning Model For Fault Diagnosis

The whole network of virtualization is a distributed and parallel system with complex correlations among the links and nodes. Service arrival is stochastic. The variance ratio is nonlinear. Complicated correlations of data of time dimension



1. Architecture of LSTM network fault prediction model

need considering for better prediction of the network faults. High level data features describe the network better; the deep learning model can discover deep features and complex correlations between links and nodes. Therefore, compared with traditional methods, the deep learning model is a more powerful tool for mining deep features from the network data of network virtualization.

In the paper, we propose a deep learning model consists of LSTM layers and a softmax layer. LSTM layers can extract the high-level features from the raw network data with time information; the softmax layer performs the multiclass classification task with the help of deep features from LSTM layers. The architecture of LSTM network fault prediction model for this problem is shown in Fig. 4. The input layer data is the data vector (1). After receiving the data, the deep learning model performs features learning in LSTM. The last layer is softmax layer. This layer receives the processed feature data, calculates the possibility of every category and make multi-classification. The output of softmax layer is:

where is the number of faults’ type, is the feature vector of -th network sample data, is the corresponding fault label and is the parameters of softmax layer which is need to be trained. The cost function of the softmax layer is:

During the learning period, the loss function for multiple classes is used:

where is the network data, is the -th output classes while is the -th data label. In order to minimize the loss function in the learning period, we calculate the partial derivatives of loss function with respect to the weights:

where is the input of the network to unit at time , is the activation of unit at time and is the state of cell at time .

For output gates:

where  is the activation function.

For cells:

where

For forget gates:

For input gates:

After data sequence is given, the equation for gradient descent can be calculated as follows:

where is the -th weight update. is the learning rate and is the momentum parameter for speeding up convergence and helping to escape from local minima. When the validation set error begins to rise, the network weights update progress stops and we obtain the optimal LSTM model.

# Experiment and results

## Simulation and settings

In this paper, we build a network fault simulation model based on OPNET14.5 simulation software. The network parameters settings are in table II.

The services we used are default service models in OPNET module including Database Access (Heavy), Database Access (Light), Email (Heavy), Email (Light), File Transfer (Heavy), File Transfer (Light), File Print (Heavy), File Transfer (Light),Telnet Session (Heavy), Telnet Session (Light), Video Conferencing (Heavy), Video Conferencing (Light), Voice over IP Call (PCM Quality), Voice over IP Call (GSM Quality), Web Browsing (Heavy HTTP1.1), Web Browsing (Light HTTP1.1). The service types of the user nodes support are randomly allocated. The start time of the service is randomly set. The duration time of different service obey uniform distribution while the parameters of the distribution are different to diverse services. The occurrence of faults obeys Poisson distribution and when a fault is in occurrence, the fault thread stops choosing fault from the faults set we give in advance according to random number.

## Results and analysis

In the experiment, six different types of faults are given. The simulation time is 60 days; data are collected at every 2 second. The number of parameter vectors we get is 2549698 for each component. Each network component data has 119 dimensions. The performance of the method are evaluated based on the following metrics:

Fig. 5 shows the fluctuation of the training accuracy during training process with different network scale. That scale is 1 means only one component’s information is included at each network component point; that scale is 2 includes the component information and its neighbor components’ information while 3 includes the information of neighbors’ neighbors and so on. It means that with the increase of the parameter scale, more network components connected to the present one directly or indirectly will be taken into consideration and more data will be fed to the LSTM network. As is shown in the figure, when network scale is 1 and 2, sample is 2549698 and time step is 400, there is little information to describe the component in the network. The stability of training accuracy performs poorly.

As the scale increases, more information about the network is available and training accuracy changes more stably with higher accuracy.

Fig. 6 represents the changes of training and validation accuracy of our model compared with the Naïve Bayesian(NB) classifier. As is shown in the figure, LSTM models is more

1. Network Settings

|  |  |
| --- | --- |
| Parameters | Values |
| Number of users | 500 |
| Number of services | 16 |
| Number of routers | 20 |
| Number of links | 47 |
| Number of Servers | 10 |
| Network Span | 500km\*500km |



1. Training accuracy of LSTM during train process

accurate than NB classifier when the scale becomes larger. When scale is less than 1, our model overfits the fault diagnosis task. With more network nodes and links are fed into the deep learning model, its performance improves a lot. In the environment of network virtualization, network scale is usually more than 2. The network component’s information and its neighbors’ information can be obtained and fed into the LSTM network for training.

We check the recall of different faults as is shown in table II. In the experiment, faults are divided into 6 categories: . is the least severe fault while is the most. The more severe the fault is, the more accurately the model can predict. This is good for faults diagnosis in network virtualization for the reason that with great accuracy of severe faults prediction, we can have better preparation for it and reduce the loss.

In the end, we do a response time test. With the model we train under the condition that sample number is 2549698, scale is 4 and time step is 400. The system that the model running on is Ubuntu14.04 with Intel(R) Core(TM) i7-6700 CPU and 8G memory. The average response time is 1.079s. In the experiment experiment, the length of prediction time is 800s calculated by . The spare time is enough for the faults preparation.



1. Training and validation accuracy of LSTM

# Conclusion

In this paper, we propose a network faults diagnosis method based on deep learning for the network virtualization environment. The contribution of our work is the analysis of the parameters in the faults diagnosis methods under the environment of network virtualization. Our deep learning model enables the earlier failure predictions by the LSTM method on discovering the long-term features of network history data. Simulation results show that the proposed method performs well on faults prediction.

# Acknowledgements

This work was supported by Natural Science Foundation of China (61372125), National Science and Technology Major Project of Ministry of Science and Technology of China ( 2017ZX030107001), 973 project (2013CB329104), and Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (16KJA510005).

1. Network Faults Recall

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Fault |  |  |  |  |  |  |
| Recall (%) | 89.87 | 92.12 | 94.23 | 97.98 | 98.16 | 97.91 |

##### References

1. J. Wu and X. M. Li. “Communication Network Alarm Correlation Based on Multi-dimensional Fuzzy Association Rules Mining”, Second International Conference on Electric Information and Control Engineering(ICEICE), vol. 01, pp. 439-443, April 2012
2. L. Yu, Y. Sun, K. J. Li and M. Xu, “An Improved Genetic Algorithm Based on FuzzyInference Theory and Its Application in DistributionNetwork Fault Location”, 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), June 2016
3. J. L. Huang and G. L. Yang, “Modified matrix algorithm for fault section location of distribution network”, Power System Protection and Control, vol.42, pp.41-45, June 2014
4. L. Ge, L. Y. Peng, R. H. Liu, F. Zhou and W. Xin, “Fault location in distribution network with distributed generation based on neural network”, China International Conference on Electricity Distribution(CICED), vol.10, pp.209-212, September 2014
5. R. Li, C. Y. Gao and L. Y. Zhang,“The distributed fault diagnosis of power networks based on Bayesian rough set method”, Journal of North China Electric Power University, vol. 37, pp.1-7, February 2010
6. Y. L. Pan, X. S. Qiu and S. L Zhang. “Fault diagnosis in network virtualization environment.” In 18th International conference on telecommunications (ICT), Ayia, pp. 517-522, May 2011.
7. N. Liu, S. L. Zhang, X. D. Wang, “Virtual network fault diagnosis using trust evaluation”, Video engineering，2016
8. S. L. Zhang, X. S. Qiu and L. M. Meng. “Service fault diagnosis algorithm in network virtualization environment.” Journal of software, vol.23, pp.2772-2782, October 2012
9. Z. Luo, P Cao, G Tang and L Wu. “A Modification to the Jelinski-Moranda Software Reliability Growth Model Based on Cloud Model Theory”. International Conference on Computational Intelligence and Security, Cis 2011, Sanya, Hainan, China, pp. 195-198, December. 2011.
10. K. Yao, T. Cohn, K. Vylomova, K. Duh and C. Dyer, “Depth-gated recurrent neural networks” arXiv preprint arXiv:1508.03790, 2015.
11. A. Graves, A. R. Mohamed and G. Hinton, “Speech recognitionwith deep recurrent neural networks,” in Acoustics,Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013, pp. 6645–6649.