Naive-LSTM based services awareness of edge computing elastic optical networks*

HUO Chao[1], BAI Huifeng[1], YIN Zhibin[1], and YAN Bo[2]**

1. Beijing SmartChip Microelectronics Technology Company Limited, Beijing 102200, China

2. Huakeyinnuo (Tianjin) Energy Technology Co., Ltd., Tianjin 300143, China

(Received 8 November 2022; Revised 19 January 2023)

Great challenges and demands are presented by increasing edge computing services for current elastic optical networks (EONs) to deal with serious diversity and complexity of these services. To improve the match degree between edge computing and optical network, the services awareness function is necessary for EON. This article proposes a Naive long short-term memory (Naive-LSTM) based services awareness strategy of the EON, where the Naive-LSTM model makes full use of the most simplified structure and conducts discretization of the LSTM model. Moreover, the proposed algorithm can generate the probability output result to determine the quality of service (QoS) policy of EONs. After well learning operation, these Naive-LSTM classification agents in edge nodes of EONs are able to perform services awareness by obtaining data traffic characteristics from services traffics. Test results show that the proposed approach is feasible and efficient to improve edge computing ability of EONs.

Document code: A **Article ID:** 1673-1905(2023)05-0279-5

DOI https://doi.org/10.1007/s11801-023-2187-x

With great development of Internet of things (IoTs) and edge computing technologies, traditional optical networks in cloud computing architecture is undergoing great change and evolution[1,2]. Under this trend, the edge computing optical network is allowed to support newly emerging services locally with better real-time ability and lower cost[3,4]. Benefiting from high efficiency and flexibility, the elastic optical network (EON) has become strong candidate to large-scale edge computing system and plays an increasingly important role to support edge computing services[5].

Great effort has been made in the field of edge computing optical network to improve service respond speed and network efficiency, by connecting great number of edge computing devices with EON[6]. However, the challenge caused by special data traffic feature of edge computing is still there. The mismatching between edge computing services and EONs will make it hard to guarantee the quality of service (QoS).

Therefore, it is of great importance to be aware of services traffic and to supports diversified edge computing services[7,8]. This services awareness has become an increasingly important ability for the EON to achieve strong performances of network control, traffic schedule and resources allocation. To realize this goal, there exist several categories of awareness methods[9-12]. In Ref.[9], the Bayes network model was proposed to realize services awareness. And echo-state-network (ESN) algorithm was also discussed in Ref.[10] as one of strong method to achieve service classification function with great efficiency and accuracy. Moreover, the long short-term memory (LSTM) model as a typical neural network has the hidden layer elements matrix inside, with the well-trained output layer results[11]. With advantages of great reliability and accuracy, the LSTM model has been deeply introduced into the field of forecasting or state prediction, and it is also a strong candidate to perform services classification[12]. However, some necessary improvement of original LSTM is still need to be done to make it well prepared for services awareness.

With the aim to overcome this problem, this paper proposes a Naive long short-term memory (Naive-LSTM) model. And the Naive-LSTM based services awareness algorithm is also presented for the edge computing driven EON. Moreover, both Naive-LSTM engine and agents are also embedded in the controller and each optical node in the EON to perform services awareness. Thus, system performances of the network blocking rate and the delay time can be greatly improved in the EON, which shows great feasibility and efficiency of the proposed algorithm.

Different from cloud computing architecture, the edge computing allows data to be processed locally by distributed edge computing devices with light-weighted computing and storing ability. In the edge computing network, original data can be distributed and be storied in

* This work has been supported by the National Key Research and Development Program of China (No.2021YFB2401304).

** E-mail: baihuifeng1984@163.com

several edge servers. When the service request arrives, the nearest edge node is chosen as edge data center. And original data is processed locally in these edge servers. After that, data processing results are updated to the cloud data center.

When combining the edge computing technology with EONs, the edge computing EON is composed by distributed edge computing services and the EON which connects these edge computing each servicer together[13,14]. By making full use of advantages of software defined network (SDN) technology, the edge computing EON can be realized[15]. The typical architecture of edge computing EON is depicted as Fig.1.

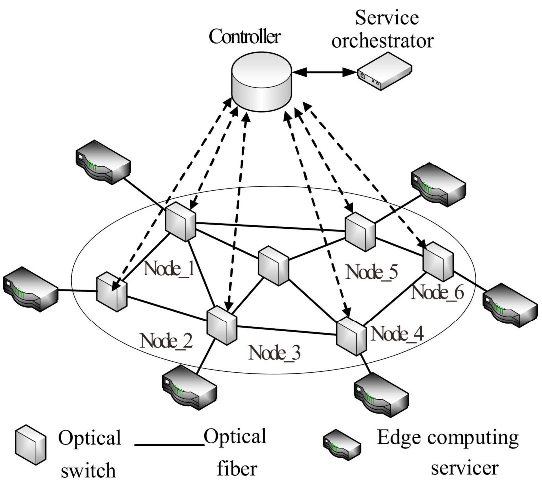

Fig.1 Typical architecture of edge computing EON

As the EON technology is adopted to realize the backbone network of IoT system, diversified services requests must be satisfied in this edge computing EON[16,17].

This article defines and proposes a deeply simplified LSTM model with discretization process, which will be used as classifier with lower complexity. The original LSTM structure is composed by one state element and three gates, which includes input gate, forget gate and output gate. The state element is to record current state and these gates are used to control forget or memory action. In this article, the structure of original LSTM is greatly simplified by merging the input gate and forget gate, which include only one state element and two gates. The simplification of original LSTM neural network is depicted by Fig.2.

And this simplified structure of LSTM element is described in details as follows. The input gate is to control the input information into the LSTM. The state element $\boldsymbol{c}_k$ is to combine input signal $\boldsymbol{z}_k$ and former state $\boldsymbol{c}_{k\text{-}1}$. And the output gate is to control the current output degree of state element signal $\boldsymbol{c}_k$.

Going further, key parameters of equation in each gate structure must also be reduced to improve the training speed of LSTM model. Moreover, the discretization must also be done to conduct classification function. Through both simplification and discretization of the original LSTM, this so-called Naive-LSTM model is defined as

$$\begin{cases} \boldsymbol{z}_t = g\left(\boldsymbol{W}_z \boldsymbol{x}_k + \boldsymbol{U}_z \boldsymbol{h}_{k-1} + \boldsymbol{b}_z\right) \\ \boldsymbol{i}_k = \sigma\left(\boldsymbol{U}_i \boldsymbol{h}_{k-1}\right) \\ \boldsymbol{c}_k = \left(1-\boldsymbol{i}_k\right) \odot \boldsymbol{c}_{k-1} + \boldsymbol{z}_k \\ \boldsymbol{o}_k = \sigma\left(\boldsymbol{U}_o \boldsymbol{h}_{k-1}\right) \\ \boldsymbol{h}_k = \boldsymbol{o}_k \odot g\left(\boldsymbol{c}_k\right) \end{cases}, \tag{1}$$

where $\boldsymbol{x}_k$ is the current input vector and $\boldsymbol{h}_k$ is the current output vector. And $\boldsymbol{z}_k$, $\boldsymbol{i}_k$, $\boldsymbol{c}_k$, $\boldsymbol{o}_k$ and $\boldsymbol{h}_k$ represent input signal, input gate, state element, output gate and output signal. $\boldsymbol{W}_z$ is the input weight matrix, and $\boldsymbol{U}_z$, $\boldsymbol{U}_i$ and $\boldsymbol{U}_o$ are recursive weight matrices of these gates. And σ is the Sigmod activation function as Eq.(2) and g is the Tanh activation function.

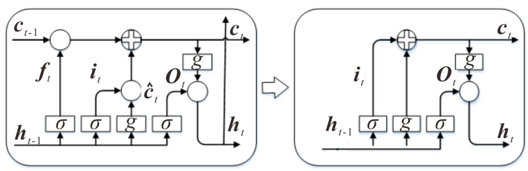

Fig.2 Simplification of LSTM element

$$\sigma_{\text{Sigmoid}}\left(x\right) = \frac{1}{\mathrm{e}^{-x}+1}. \tag{2}$$

The main principle of Naive-LSTM based identification is given as

$$\begin{cases} \boldsymbol{z}_{k+1}^{(n)} = g\left(\boldsymbol{W}_z \boldsymbol{x}_{k+1}^{(n-1)} + \boldsymbol{U}_z \boldsymbol{h}_{k+1}^{(n-1)} + \boldsymbol{b}_z\right) \\ \boldsymbol{z}_{k+1}^{(0)} = 0 \end{cases}. \tag{3}$$

This model not only has the advantage of original LSTM with fast and simple training performance, but also has strong potential to conduct classification.

In the aspect of algorithm complexity, the original LSTM needs to update a number of parameters up to $4(mn+n^2+n)$, since each gate has $(mn+n^2+n)$ parameters to be updated. While the Naive-LSTM reduces this number to $3(mn+n^2+n-2mn-2n)$ and its computation complexity is further simplified when compared with the original one.

Aimed to improve the edge computing ability of EONs, this article proposes the Naive-LSTM based services identification in the edge computing EON, together with the detailed procedure.

The proposed Naive-LSTM based services awareness scheme is divided into two parts, and they are the Naive-LSTM engine module in controller and agent modules in edge optical node. Moreover, the Naive-LSTM engine model is responsible of the Naive-LSTM model training for services classification. While the agent one works inside edge optical node to abstract service characteristic parameters from newly arrived packets sub-traffic and to conduct services awareness. The core model of Naive-LSTM based services awareness that works in controller and edge optical nodes is depicted in

Fig.3.

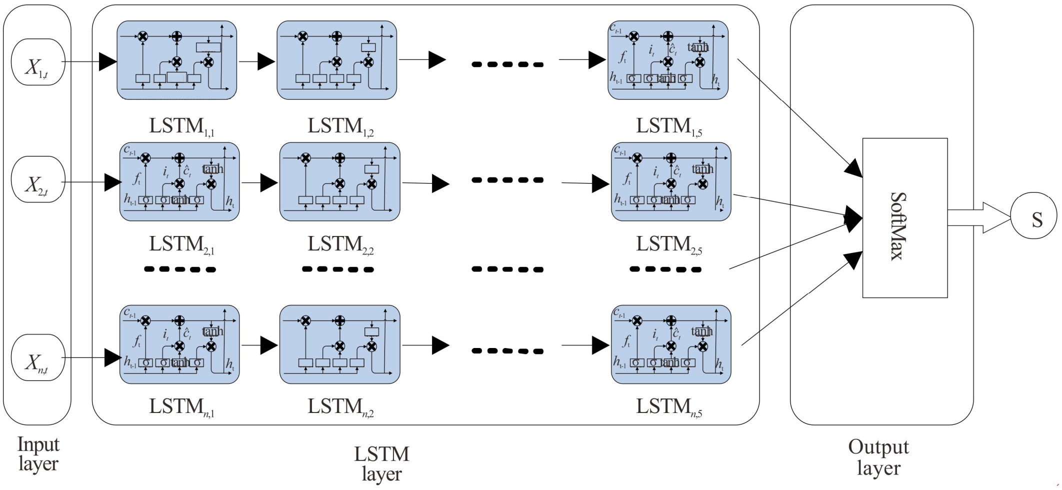

Fig.3 Naive-LSTM based services awareness model

In this naive-LIST core model, the structure is consisted by input layer, LSTM layer and output layer. Moreover, the LSTM layer is an $m*n$ matrix of Naive-LSTM element.

In general, there exists two categories of methods to realize services identification: the packet-level identification and the traffic-level one. The main parameters of the first one mainly collects packet size, arrival interval and arrive time, while the later one generally focuses on traffic size, duration time and so on. In fact, five packets of the same service traffic is enough to get these service characteristics parameters. Combining both approaches mentioned above together, the sub-traffic is define as the set of firstly five packets. These parameters of sub-traffic for service classification are also defined as follows.

(1) Sub-traffic size: the sum value of lengths of the first five packets as a whole from the same edge computing service traffic;

(2) Average packet size: the average packet length of the same edge computing service;

(3) Maximum packet size: the maximum value of the packet length belonging to the same service traffic;

(4) Minimum packet size: the minimum value of the packet length belonging to the same service traffic;

(5) Average interval time: the average interval time between packets from the same sub-traffic;

(6) Duration time: the duration period of the sub-traffic from the same edge computing service.

In the whole EON network, both initiation and training of the Naive-LSTM model are unitedly conducted by the controller of EON based on SDN. In the control plan, there exist one training module that has a deal of training samples with enough number. And the detailed procedure of Naive-ESN model training in the control plan is depicted as follows, using the gradient descent algorithm.

Step 1: Compute the value of $\boldsymbol{h}_k$ and these error terms of $\boldsymbol{z}_k$, $\boldsymbol{i}_k$ and $\boldsymbol{o}_k$, according to

$$\begin{cases}\Delta\boldsymbol{h}_k=\Delta\boldsymbol{z}_{k+1}\boldsymbol{U}_z+\Delta\boldsymbol{i}_{k+1}\boldsymbol{U}_i+\Delta\boldsymbol{o}_{k+1}\boldsymbol{U}_o\\ \Delta\boldsymbol{z}_k=\Delta\boldsymbol{h}_k\times\boldsymbol{o}_k\times g\left(\boldsymbol{c}_k\right)\times\boldsymbol{i}_k\times\boldsymbol{z}_k\\ \Delta\boldsymbol{i}_k=\Delta\boldsymbol{h}_k\times\boldsymbol{o}_k\times g\left(\boldsymbol{c}_k\right)\times\boldsymbol{i}_k\times\boldsymbol{z}_k\\ \Delta\boldsymbol{o}_k=\Delta\boldsymbol{h}_k\times g\left(\boldsymbol{c}_k\right)\times\boldsymbol{o}_k\end{cases}. \tag{4}$$

Step 2: Compute the updating Eq.(5) of input weight matrix and so on.

$$\begin{cases}\Delta\boldsymbol{W}_{z,k}=\Delta\boldsymbol{z}_k\times\boldsymbol{x}_k\\ \Delta\boldsymbol{U}_{\Omega,k}=\Delta\boldsymbol{\Omega}_k\times\boldsymbol{h}_{k-1}\\ \Delta\boldsymbol{b}_{z,k}=\Delta\boldsymbol{z}_k\end{cases}, \tag{5}$$

where $\boldsymbol{\Omega}$ represents any one of $\boldsymbol{z}$, $\boldsymbol{i}$ or $\boldsymbol{o}$.

Step 3: Compute these matrixes after k rounds of updating using

$$\begin{cases}\Delta\boldsymbol{W}_{z,k}=\boldsymbol{W}_{z,k+1}-\eta\times\Delta\boldsymbol{W}_{z,k}\\ \Delta\boldsymbol{U}_{\Omega,k}=\boldsymbol{U}_{\Omega,k+1}-\eta\times\Delta\boldsymbol{U}_{\Omega,k}\\ \Delta\boldsymbol{b}_{z,k}=\boldsymbol{b}_{z,k+1}-\eta\times\Delta\boldsymbol{b}_{z,k}\end{cases}. \tag{6}$$

Step 4: Compute the root mean squared error (*RMSE*) and judge weather the *RMSE* of training samples reaches the expected one to complete the training.

After the Naive-LSTM services awareness engine in the controller is well trained, the Naive-LSTM model information is distributed to all agent-ones in edge optical nodes in the whole EON.

Under the edge computing EON circumstance, the Naive-LSTM agent module is embedded in all nodes of the data plan. In this article, the service-awareness ability of edge computing EON is mainly realized by those distributed EON nodes.

The controller in control plan collects topology information of the whole EON. Later, some routing algorithms are adopted to compute new route and allocate proper links and nodes resources for newly arrived service, under current network states and limited conditions. Later, new flow-tables are generated by the controller and will be send to all nodes of new route. Thus, the service awareness procedure is given in detail as follows.

Step 1: On receiving new service traffic, the edge optical node searches for the record of this flow-table;

Step 2: If there exists related record in flow-table, this node directly transports packets of this newly arrived service traffic; Otherwise, tern to Step 3;

Step 3: The edge optical node abstracts sub-traffic of these newly arrived services, and obtains the parameter set of all related characteristics service traffic;

Step 4: The Naive-LSTM agent model in edge optical node conducts the classification computation and determine the type of this service traffic with the QoS requirements;

Step 5: Edge optical node generates the connection request message including QoS requirements parameters and reports them to the controller;

Step 6: The controller computes the route according to service request and using related routing algorithms;

Step 7: The controller modifies related flow-table of involved optical nodes so that network resources can be allocated to newly arrived service traffic;

Step 8: These involved nodes in EON send acknowledgment (ACK) messages to the controller and the connection for newly arrived services is successively complete.

Thus, the Naive-LSTM model based services awareness algorithm is well prepared for the edge computing EON.

To evaluate the proposed Naive-LSTM model based services awareness scheme, the simulation is conducted in this section, where the simulation environment is constructed by NS3 network simulation software tools and the NSF-net topology is adopted. And this simulated platform of the EON mainly consists of 14 nodes and 21 fiber links with spectrum range of 4 THz, where each link contains 150 slots and each physical node has 750 units of computing capacity. And there is one controller with service awareness engine and 2 000 training samples inside the training module, in which all nodes of data plan are equipped with the services awareness agent.

In this simulation, several typical edge computing services are divided into three categories: the class_1 with the highest priority, the class_2 with medium priority and the class_3 with the lowest priority. Additionally, the proportion of each type of service is given in Tab.1.

Tab.1 Service types list

Service type	Characteristic	Proportion
Control	Poisson	10%
Flow media	Poisson	15%
VoIP	Poisson	20%
Alarm	Period	10%
IPv6	Constance	25%
Data service	Poisson	20%

Simulation analysis is conducted mainly in terms of services connection blocking rate and connection establishment delay time. The comparison is made between this proposed scheme and the ESN algorithm driven services awareness in the edge computing EON. And simulation results are given in Fig.4 and Fig.5.

The comparison of service blocking rate is made in Fig.4, as the communication success is one of key performances in the EON. Obviously, these blocking rate curves of class_1 services and class_2 services with the Naive-LSTM are better than those ones with the typical ESN algorithm, while the class_3 service using the Naive-LSTM algorithm has the worst performance, because it has lower demand on transporting succession rate. Thought comparison, the blocking rate of the Naive-LSTM based services awareness scheme is much more reasonable than the ESN based one. That is because the proposed Naive-LSTM algorithm has higher accuracy of classification than the ESN algorithm. With this proposed services awareness function, different results among those services are greatly improved. Thus, both successive rate and matching degree of connection establishment of all services can be further satisfied by using the Naive-LSTM based services awareness.

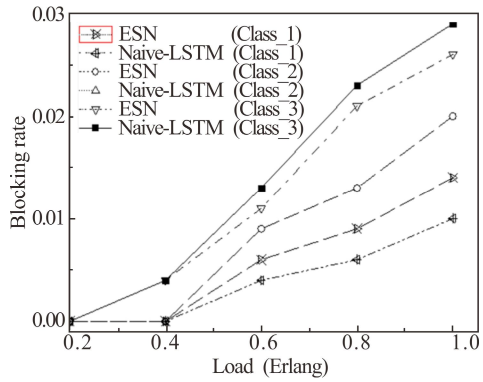

Fig.4 Comparison of block rate

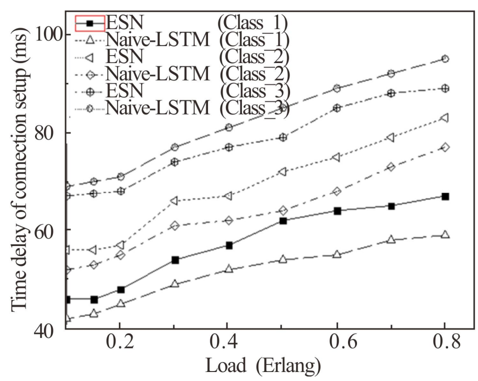

Fig.5 Comparison of time delay

As the delay time of service connection establishment is another key point of EON system, Fig.5 gives the comparison result of the delay time between the proposed algorithm and the ESN based one. Obviously, all services under each priority soar as the service burden increases. With the lowest priority on delay time, the class_3 service is more tolerant to the real-time performance. Services of class_1 with the Naive-LSTM algorithm achieve the best real-time result. It is suggested that the advantage of the proposed algorithm has faster processing speed.

Overall, it can be drawn that this proposed scheme is able to greatly improve the service awareness ability of the edge computing EON system. And it also suggests that the proposed approach can provide better support to industrial edge computing services with much more feasibility, greater efficiency and accuracy of performance.

The edge computing technology had already played increasingly important role in the industrial-IoT. Great challenges and demands were also presented by increasing edge computing services for current EONs to deal with serious diversity and complexity of these services. To better supporting ability for edge computing services, the services awareness ability was indeed necessary. This article proposed a Naive-LSTM based services awareness algorithm of the EON, where the Naive-LSTM model was defined by simplifying the inner structure and conducting discretization of the original LSTM model. The proposed algorithm could generate the probability output result to determine the QoS policy of edge computing EONs. After well training operation, these Naive-LSTM classification agents in edge computing node of EONs was able to perform services awareness, by obtaining data traffic features from services traffics. Test results show that the proposed approach was able to improve edge computing oriented supporting ability of EONs.

Statements and Declarations

The authors declare that there are no conflicts of interest related to this article.

References

[1] JI Y, ZHANG J, WANG X, et al. Towards converged, collaborative and co-automatic (3C) optical networks[J]. Science China information sciences, 2018, 61(12)：121301.

[2] PELLW I, PAOLUCCI F, SONKOLY B et al. Latency-sensitive edge/cloud serverless dynamic deployment over telemetry-based packet-optical network[J]. IEEE journal on selected areas in communications, 2021, 39(9)：2849-2863.

[3] ALGHAMDI K, BRAUN R. Deploying hand off-mechanism with the software defined network vs mobile IP for 5G network：a feasibility study[J]. Journal of advances in technology and engineering research, 2019, 5(2)：79-84.

[4] LI Y, ZENG Z, LI J, et al. Distributed model training based on data parallelism in edge computing-enabled elastic optical networks[J]. IEEE communications letters, 2021, 25(4)：1241-1244.

[5] YANG H, YAO Q, BAO B, et al. A 3-CS distributed federated transfer learning framework for intelligent edge optical networks[J]. Frontiers in communications and networks, 2021, 2(29)：700912.

[6] LI J, HUA N, ZHONG Z, et al. Flexible low-latency metro-access converged network architecture based on optical time slice switching[J]. Journal of optical communications and networking, 2019, 11(12)：624-635.

[7] VILALTA R, MANSO C, YOSHIKANE N, et al. Experimental evaluation of control and monitoring protocols for optical SDN networks and equipment[J]. Journal of optical communications and networking, 2021, 13(8)：D1-D12.

[8] BAGCI K, TEKALP A. SDN-enabled distributed open exchange ： dynamic QoS-path optimization in multi-operator services[J]. Computer networks, 2019, 162：106845.1-106845.10.

[9] BAI H, LI M, WANG D. Bayesian classifier based service-aware mechanism in 10G-EPON for smart power grid[J]. Acta photonica sinica, 2013, 42(6)：668-673.

[10] BAI H, CHEN W, WANG L, et al. Naive echo-state-network based services awareness algorithm of software defined optical networks[J]. China communications, 2020, 17(4)：11-18.

[11] CUI Z, HENRICKSON K, KE R, et al. Traffic graph convolutional recurrent neural network：a deep learning framework for network-scale traffic learning and forecasting[J]. IEEE transactions on intelligent transportation systems, 2019, 21(11)：4883-4894.

[12] GU R, ZHANG S, JI Y, et al. Network slicing and efficient ONU migration for reliable communications in converged vehicular and fixed access network[J]. Vehicular communications, 2018, 11(1)：57-67.

[13] HUI Y, JIE Z, ZHAO Y, et al. Service-aware resources integrated resilience for software defined data center networking based on IP over flexi-grid optical networks[J]. Optical fiber technology, 2015, 21(1)：93-102.

[14] ZHU R, ZHAO Y, HUI Y, et al. Dynamic time and spectrum fragmentation-aware service provisioning in elastic optical networks with multi-path routing[J]. Optical fiber technology, 2016, 32(12)：13-22.

[15] AGRAWAL A, VYAS U, BHATIA V, et al. SLA-aware differentiated QoS in elastic optical networks[J]. Optical fiber technology, 2017, 36(7)：41-50.

[16] AMARAL P, DINIS J, PINTO P, et al. Machine learning in software defined networks：data collection and traffic classification[C]//IEEE 24th International Conference on Network Protocols, November 8-11, 2016, Singapore. New York：IEEE, 2016：1-5.

[17] ASSIS K, ALMEIDA R, SANTOS A, et al. Channel-based RSA approaches for QoS protection of slices over elastic optical networks[J]. IEEE access, 2022, (10)：20714-20726.