Hierarchy Bayesian model based services awareness of high-speed optical access networks*

BAI Hui-feng (白晖峰)**
Beijing SmartChip Microelectronics Technology Company Limited, Beijing 102200, China

(Received 14 November 2017; Revised 3 January 2018)

As the speed of optical access networks soars with ever increasing multiple services, the service-supporting ability of optical access networks suffers greatly from the shortage of service awareness. Aiming to solve this problem, a hierarchy Bayesian model based services awareness mechanism is proposed for high-speed optical access networks. This approach builds a so-called hierarchy Bayesian model, according to the structure of typical optical access networks. Moreover, the proposed scheme is able to conduct simple services awareness operation in each optical network unit (ONU) and to perform complex services awareness from the whole view of system in optical line terminal (OLT). Simulation results show that the proposed scheme is able to achieve better quality of services (QoS), in terms of packet loss rate and time delay.
Document code: A **Article ID:** 1673-1905(2018)02-0114-5
DOI https://doi.org/10.1007/s11801-018-7244-5

The rapid progress of passive optical networks (PON) technologies has already provided strong support to diversified multiple services with ever-increasing capacity[1-3]. And great effort has been made to develop the high-speed optical access network under the combination of research institutes and carriers. Especially, the recent breakthrough on 10-gigabyte (10G) and even 40-gigabyte (40G) Ethernet passive optical network (EPON) makes it one strong candidate to future access networks[4-6].

In current circumstances of ever-increasing services traffic load, high-speed optical access networks still fail to actively be aware of services characteristics and to provide active support on quality of service (QoS), especially due to the ever-enlarged scale and the ultra high speed of X-gigabyte (XG) EPON and so on[7]. Thus, the services awareness ability is necessity to be fully aware of characteristics and requirements of multiple services and to actively match these services with outstanding performance[7]. In general, the services awareness contains three aspects, including awareness objects, awareness methods and awareness results[7,8].

It has been proved that there exist several key parameters having strong relations with services traffics characteristics, including the packet size, inter-arrival time and duration time, etc. Those parameters of service characteristics can be gained through independent different methods. By using some proper classification technologies, these parameters of traffics charactristics are allowed to perform services-awareness function in 10G-EPON. Great effort has been done in the field of services awareness. Deep researches on the relation between statistic characteristics and service protocols have been done in Refs.[6,7]. A series of researches on data-traffic identification and classification mechanism have been reported in Refs.[8-12], which are based on the packet length, arrival time interval and arrival sequence. A Bayesian classifier based services awareness (BC-SA) has been presented in Ref.[12]. This also implies that more pattern recognition algorithms also have great potential to realize service awareness.

Due to the master/slave structure of PON system, the optical line terminal (OLT) has strong processing ability, while each optical network unit (ONU) is just equipped with limited ability and fails to support all functions of services awareness. Additionally, there exists a problem for all ONUs to keep the consistency of services awareness results, since each ONU works independently under the control of OLT.

Aiming to overcome these problems mentioned above, this paper proposes a hierarchy Bayesian model based services awareness mechanism in high speed PON system. Different from traditional Bayesian model, this proposed hierarchy Bayesian model can be divided into three parts: the rule layer and the process layer running in the OLT and the execution layer embedded in all ONUs. And this model is suitable for the PON structure. Thus, the ONU is able to realize limited services awareness function

* This work has been supported by the Science and Technology Project of State Grid Corporation of China: "Research on the Power-Grid Services Oriented "IP + Optics" Coordination Choreography Technology".
** E-mail: lancer101@163.com

with consistency results, working together with the process layer Bayesian module in OLT. Simulation results show the proposed mechanism is able to achieve better quality of services (QoS) in 10G-EPON.

In general, the Bayesian classification is one typical classification model based on statistical method. The Bayesian theory combines both the priori probability and the posteriori probability together, and makes full use of the priori information and the sample data information to determine the posteriori probability.

Assume that $U=\{X_1, X_2,...X_n, C\}$ is the finite set of discrete random variables, in which X_i is attribute variable with value x_i and C is class variable from $\{c_1, c_2,...c_n\}$. Thus, $X_i=(x_1, x_2,...x_n)$ is the probability that it belongs to c_i, as defined by

$$\begin{aligned} P\left(c_j \mid x_1, x_2, ..., x_n\right) &= \frac{P\left(x_1, x_2, ..., x_n \mid c_j\right) \cdot P\left(c_j\right)}{P\left(x_1, x_2, ..., x_n\right)} \\ &= \alpha \cdot P\left(c_j\right) \cdot P\left(x_1, x_2, ..., x_n \mid c_j\right) \end{aligned} ,\quad(1)$$

where α is the regularization factor, and $P(c_j)$ is the prior probability of c_j, while $P(x_1, x_2,...x_n|c_j)$ is the corresponding posterior probability. According to the probability chain rule, Eq.(1) could be changed into

$$\begin{aligned} &P\left(c_j \mid x_1, x_2, ..., x_n\right) = \\ &\alpha \cdot P\left(c_j\right) \cdot \prod_{i=1}^{n} P\left(x_i \mid x_1, x_2, ..., x_{i-1}, c_j\right) \end{aligned} . \quad(2)$$

Given the training data set $D=\{x_1, x_2,\ldots x_N\}$, the goal of classification is to make analysis of data set D and to use a mapping function $f()$, in order to give the right tag C^* of $X_i=(x_1, x_2,...x_n)$. Thus, the Bayesian classification model can be defined as

$$\begin{cases} P\left(C^*\right) = \max P\left(C_i\right) \\ P\left(C_j\right) = \alpha \cdot P\left(c_j\right) \cdot \prod_{i=1}^{n} P\left(x_i \mid x_1, x_2, ..., x_{i-1}, c_j\right) \end{cases} . \quad(3)$$

Genarally, it is assumed that all attribute variables are independent, so that the Bayesian classification has high efficiency for its simple structure.

In general, the Bayesian classification is divided into two stages: the Bayesian training and the Bayesian inference. And the classification ability of Bayesian model is limited by completeness and comprehensiveness of the Bayesian training. Under the condition that statistical knowledge is full enough, the Bayesian classification is able to achieve high accuracy with minimum average risk. Therefore, it is the key of Bayesian classification to obtain the probability density function of each sample characteristics X_i.

On the basis of the Bayesian classification theory, this paper proposes a hierarchy Bayesian model, which can be devided into three layers: the rule layer, the process layer and the execution layer.

As depicted in Fig.1, the rule layer is responsible for initiation of Bayesian classification, including main factor values.

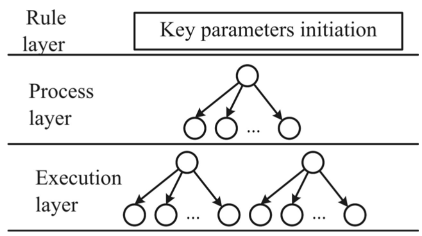

Fig.1 Architecture of the hierarchy Bayesian model

The task of process layer is to train the Bayesian classificer until it is fully ready for classification operation, using the training sample data set. Moreover, the process layer also provides all information to the execution layer, after the Bayesian classifier is fully trained.

The execution layer is the right one to conduct the services classification or identification operation, using the well-trained Bayesian classifier initiated by the process layer. This layer consists of a number of independent agent Bayesian classifiers that are all the copies of main classifier from process layer, where their task is only to conduct classification function. That means all agent Bayesian classifiers in execution layer just have limited part function of the complete Baysian classifier, which are only able to conduct simple classification according to Eq.(3).

Aiming to avoid the over fitting phenomenon of the hierarchy Bayesian model, the pre-processing of parameters data is necessary, before the execution layer conducts classification operation. And these parameters of service i include packet size $P_{\mathrm{SIZE}}(i)$, inter-arrival time $P_{\mathrm{INTERVAL}}(i)$ and service-duration time $P_{\mathrm{DUR}}(i)$. Therefore, the pre-processing can be conducted using

$$\begin{cases} U(i) = \left(u_{i,1}, u_{i,2}, u_{i,3}\right) \\ u_{i,1} = \dfrac{P_{\mathrm{SIZE}}(i)}{P_{\mathrm{SIZE\_MAX}}} \\ u_{i,2} = \dfrac{P_{\mathrm{INTERVAL}}(i)}{P_{\mathrm{INTERVAL\_MAX}}} \\ u_{i,3} = \dfrac{P_{\mathrm{DUR}}(i)}{P_{\mathrm{DUR\_MAX}}} \end{cases} , \quad(4)$$

where $P_{\mathrm{SIZE\text{-}MAX}}$ is the maximum of packet size, $P_{\mathrm{INTERVAL\text{-}MAX}}$ is the maximum of inter-arrival time and $P_{\mathrm{DUR\text{-}MAX}}$ is the maximum service-duration time among all kinds of services. To make sure the accuracy from the whole view of hierarchy Bayesian model, these parameters, including $P_{\mathrm{SIZE\text{-}MAX}}$, $P_{\mathrm{INTERVAL\text{-}MAX}}$ and $P_{\mathrm{DUR\text{-}MAX}}$ of each agent Bayesian classifier, must keep the consistency with the same values under control of process layer in the OLT:

$$\begin{cases} P_{\mathrm{SIZE\_MAX}} = MAX\left[P_{\mathrm{SIZE\_MAX}}(j)\right] \\ P_{\mathrm{INTERVAL\_MAX}} = MAX\left[P_{\mathrm{INTERVAL\_MAX}}(j)\right] \\ P_{\mathrm{DUR\_MAX}} = MAX\left[P_{\mathrm{DUR\_MAX}}(j)\right] \end{cases} , \quad(5)$$

where $P_{\mathrm{SIZE\text{-}MAX}}(j)$ is the maximum of packet size in the agent Bayesian classifier j, $P_{\text{INTERVAL-MAX}}(j)$ is the maximum of inter-arrival time and $P_{\text{DUR-MAX}}(j)$ is the maximum service-duration time in the Bayesian classifier j. Thus, the consistency of classification result can be kept.

As one important type of high-speed optical access networks, the 10G-EPON is chosen to realize services awareness scheme in this paper and the 10G-EPON is also the typical combination of shared channel and point-to-point communication.

In this proposed scheme, these parameters of service-traffics characteristics are selected for services awareness: the packet size x_1, the inter-arrival time x_2, the traffic duration time x_3, and the number of packets x_4.

Generally, the 10G-EPON architecture consists of one OLT and a number of ONUs, where the OLT is in charge of these control, management and data transmission functions of all ONUs, and each ONU is the simple one to provide access point for services packets under the control of OLT device. Thus, the proposed hierarchy Bayesian model based services awareness system in 10G-EPON must also adopt the "master and agents" structure due to the typical OLT/ONUs architecture of 10G-EPON, in which both functions of the rule layer and process layer are centrilized in the OLT device and the execution layer function is realized by each ONU via embedded software or hardware. And the structure of the proposed scheme is depicted by Fig.2.

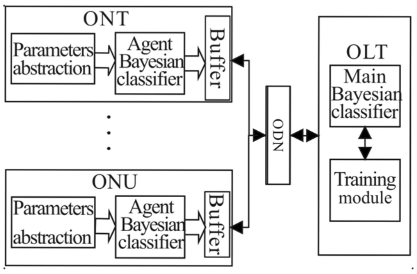

Fig.2 Hierarchy Bayesian based services awareness system structure in 10G-EPON

In this proposed hierarchy Bayesian model based services awareness system, the OLT is responsible for generating and training the Bayesian classifior template for the whole system to make sure the consistency of the services awareness result. After the configuration is completed by OLT, each ONU is allowed to independently execute the services awareness.

The rule layer and process layer work in the OLT and play the role of generating Bayesian classifier and training this Bayesian classifier. Moreover, both layers together consist of the Bayesian generation module, the Bayesian classifier and the Bayesian training module, all of which are embedded in the OLT.

In order to achieve better accuracy of Bayesian classification, a so-called accuracy-weighted training method of classifier is also introduced into the process layer in the hierarchy Bayesian model. The main idea of this cycled training is that the process layer repeats the Bayesian training for N cycles, in which those samples with wrong training result in the former cycle will be given higher weight value in this round of Bayesian training.

And the procedure of accuracy-weighted training method improved Bayesian classification is dipected as follows and repeated for T cycles.

Step 1: Set the initial training parameters, including the set of samples (X_i, c_i) and weight vector $\boldsymbol{w}$ of X_i. And the distribution of those samples is D:

$$\begin{cases}(X_i,c_i)\in\langle(X_1,c_1),(X_2,c_2),...,(X_N,c_N)\rangle \\ w_i=1/N\end{cases}. \tag{6}$$

Step 2: Calculate the number of all errors X_{error} of this round of cycle:

$$X_{\text{error}}=\sum_{h_t(X_i)\neq c_i} D_{\text{t}}(X_i,c_i). \tag{7}$$

Step 3: Compute the error factor α_{t} of this round of training, according to

$$\alpha_t=\frac{X_{\text{error}}}{1-X_{\text{error}}}. \tag{8}$$

Step 4: Conduct the normalization operation to the $D_{\text{t}}(X_i, c_i)$.

Step 5: Go to the next cycle of training, until T cycles of training are finished.

After T cycles of training are finished, the Beyasian classifier in the process layer can be ready to copy the structure and corresponding parameters of the well-trained Bayesian classifier to the execution layer that works in each ONU.

The execution layer is embedded and works in all ONUs, which consists of the services traffic information abstraction module and the agent Bayesian classifier. For each ONU in this system, there are three main tasks: initiating the Bayesion classifer under control from the OLT, collecting these services traffics parameters and executing services awareness independently.

When each ONU receives the copy of fully trained Bayesian classifier from the OLT, the "agent Bayesian classifier" module in each ONU will immediately re-build the same Bayesian classifier with the one in OLT. After that, all ONUs are allowed to conduct the services awareness task independently. As shown in Fig.2, the key information of newly arrived services traffic is abstracted firstly and forms the services variable set X_i $(x_1, x_2, \ldots, x_n)$ to the agent Bayesian classifier. The later will judge which type of services X_i belongs to, and set the priority value of this service traffic. Then, the buffer is able to conduct the schedule according to the priority value given by the agent Bayesian classifier. Thus, the services awareness is realized in ONU level[13-15].

The complexity of this hierarchy Bayesian model based classification is another important issue, which must be low enough for implementation. According to the Bayesian classification theory, it just needs two times of multiply operation to deal with one kind of classification parameter, where the complexity is $O(1)$. If the number of classification parameters reaches m, the complexity is just $O(m)\times O(1)$ for the Bayesian classification computation in each ONU.

To evaluate the proposed scheme, a 10G-EPON system simulation platform is built, which is composed of one OLT and 32 ONUs. Key parameters of the 10G-EPON system in this simulation are given in Tab.1.

Tab.1 Key parameters of 10G-EPON system

Name	Value	Name	Value
ONU number	32	Protection bandwidth	250 bytes
Queue number	8	Buffer	1M
Upstream bandwidth	10G	Distance	5 km
DBA period	1 ms	Simulation times	1 000

In this simulation, comparison is made between the original Bayesian based services awareness and the proposed hierarchy Bayesian based services awareness, where the former is realized by software and the later can be directly realized by hardware.

In this test, four categaries of services are set to produce traffic load, and the requirements of services to network are represented by three parameters: the average packet delay, the packet loss rate and the jitter. Moreover, each kind of service is divided into four levels in terms of priority: case_1 with the highest level, case_2 and case_3 with the mediate level, and case_4 with the lowest level. The traffic load of each class service follows the Passion distribution.

Tab.2 Simulation parameters of services traffics

Priority	Type	Proportion
Case_1	Control message	10%
Case_2	Voice services	20%
Case_3	Video services	30%
Case_4	Data services	40%

The comparison result in Fig.3 obviously suggests that the hierarchy Bayesian model has better accuracy than the original one. When the total number of sample data is not full enough, the performances of both are poor. As the data sample set for training increases, both approaches work well with better accuracy. Moreover, the proposed hierarchy Bayesian model can reach better performance than the original one. Benefitting from the accuracy-weighted training method to correct the wrong judgement of some sample data, the main Bayesian classifier in OLT is able to be generated with more accurate result, by paying more attention to those data samples with mistake classification.

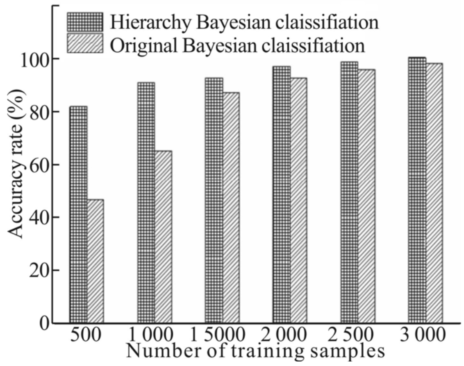

Fig.3 Classification accuracy comparison

The delay time of delay-sensitive service is given in Fig.4. Obviously, delay time of all services soars as the traffic load increases. Under the same traffic load condition, cass_1, case_2 and case_3 services with hierarchy Bayesian model show better value than their corresponding original ones, while case_4 service with hierarchy Bayesian model gets the worst value. Because the case_4 service has the lowest requirement on delay time, it is tolerant to this time-delay performance. Additionally, the case_1 service in hierarchy Bayesian model can achieve the best performance.

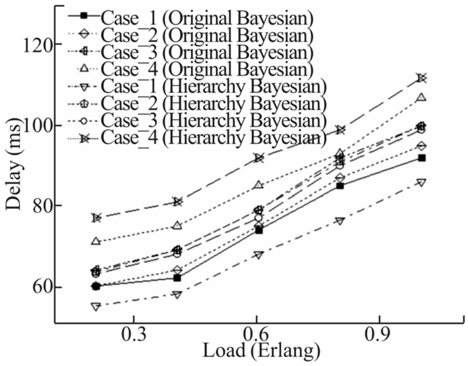

Fig.4 Comparison of time delay

Fig.5 shows the comparison of packet loss rate for packet-loss sensitive services, and these services are also divided into three levels (case_1, case_2, case_3 and case_4). Overall, the packet loss rate of all services soars as the traffic load becomes heavy. Similar to the comparison result of network delay time, case_1 service and case_2 service with hierarchy Bayesian model show lower packet loss rate than the original one. And the case_4 service with hierarchy Bayesian model still gets the worst result, since it has the lowest requirement on packet loss rate. Though comparison, the packet loss rate of hierarchy Bayesian model is much more reasonable than that of the original Bayesian model. With the hierarchy Bayesian model scheme, the differences among those services are more obvious. Thus, the requirements of different classes of services on packet loss rate can all be further satisfied by using the hierarchy Bayesian model mechanism.

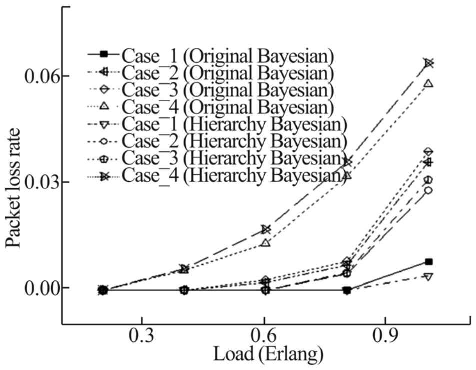

Fig.5 Comparison of packet loss rate

Combining these analysis results from Fig.4 to Fig.5, it can be drawn that this proposed hierarchy Bayesian model based services awareness scheme is able to realize distributed services awareness with outstanding performance and to keep the awareness consistency from the total view of the whole system.

A hierarchy Bayesian model based services awareness mechanism is proposed for high-speed optical access networks. This approach builds a so-called hierarchy Bayesian model according to the structure of typical optical access networks. Moreover, the proposed scheme is able to conduct simple services awareness operation in each ONU and to perform complex services awareness from the whole view of system in OLT for 10G-EPON system. Simulation results show that the proposed scheme is able to achieve better QoS in terms of packet loss rate and time delay.

References

[1] Van D P, Valcarenghi L and Chincoli M, Optical Switching & Networking **14**, 11 (2014).

[2] Li C, Guo W and Wang W, Optics Communications **370**, 43 (2016).

[3] Özgür Can Turna, Aydın M A and Zaim A H, Optical Switching & Networking **15**, 29 (2015).

[4] Houtsma V, Veen D V and Harstead E, Recent Progress on 25G EPON and Beyond, Proceedings of European Conference on Optical Communication, 2016.

[5] Tsutsumi T, Sakamoto T and Sakai Y, Journal of Lightwave Technology **33**, 1660 (2015).

[6] Roy R, Kramer G and Hajduczenia M, IEEE Communications Magazine **49**, 78 (2011).

[7] Garfias P, Gutiérrez L and Sallent S, IEEE/OSA Journal of Optical Communications & Networking **4**, 978 (2012).

[8] Khater N A and Overill R E, Network Traffic Classification Techniques and Challenges, IEEE Tenth International Conference on Digital Information Management, 3 (2016).

[9] He L, Xu C and Luo Y, vTC: Machine Learning Based Traffic Classification as a Virtual Network Function, ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization, 53 (2016).

[10] Lu C N, Huang C Y and Lin Y D, Journal of Network & Computer Applications **76**, 60 (2016).

[11] Tongaonkar A, Torres R and Iliofotou M, Computer Communications **56**, 35 (2015).

[12] BAI Hui-Feng, LI Ming-Wei and WANG Dong-Shan, Acta Photonica Sinica **42**, 668 (2013). (in Chinese)

[13] Van D P, Valcarenghi L and Dias M P I, Optics Express **23**, A1 (2015).

[14] Jun J H, Yu Z and Jung M S, Inter-ONU Scheduling Scheme for QoS Guarantee in 10G EPON, Proceedings of International Conference on Broadband and Wireless Computing, Communication and Applications, Fukuoka, Japan, 137 (2010).

[15] Park K S, Yu Z B and Cho J H, An Intra-ONU Scheduling Method in 10G EPON Supporting IEEE 802.1 AVB, Proceedings of the Fourth International Conference on Complex, Intelligent and Software Intensive Systems, Krakow, Poland, 631 (2010).