Abstract: Classification of scalable QoS systems requires more reasonable index, it needs comprehensive and deep analysis and evaluation of business requirement to realize the expansibility of the index and their mapping algorithm. In this paper, it proposes a set of extensible QoS index system on the existing basis that these mechanisms divide grade too simple, and the kinds of index are so few. Making good use of the advantages of cognitive networks on the dynamical re-configuring and communicating capabilities, meeting amounts of various requirements of QoS, designing interface of expandable QoS index, providing more precise and concrete QoS service and improving the efficiency of the networks resources are the main works of this paper.

1 INTRODUCTION

In the current information network, the network elements, such as nodes, protocol layers, strategies and behaviors, are restricted greatly in the aspects of status, scope of application, and response mechanism, therefore they cannot satisfy the QoS demands of clients with diversification and individuation. Some scholars have already proposed adaptive mechanism (Zhang L et al., 2007, Wang F et al., 2003, Partha P et al, 2000, Stroud R et al., 2004) in order to ensure the service qualities of sorts of business. These adaptive mechanisms which have been already existed, however, are all reactive modes that can only make adjustment after occurrences of problems. Also, the optimization objects are usually aiming at a layer of protocol stack, some key nodes of network, or a certain part of transmission link. Hence the end-to-end QoS performance of a network cannot be performed optimized. The information networks of new generation are with more complication and summarization. Therefore proactive adaptive algorithm needs more research to provide more suitable guarantee of QoS on business. Cognitive network (CN) (Thomas R W, 2005) can exactly meet this demand because of its unique ability of self-learning and re-configured. As the more intelligent, cognitive networks can help to provide users with richer types of business, more reliable information transmission and more optimized end-to-end QoS performances.

The principle of cognitive behavioral model (Feizi-Khankandi S et al., 2007) is: first, it detects the current state of the network, and then executes adjustment, adjudication, and implementation based on the observed network conditions and parameters. Cognitive technology makes communication entities possess a cognitive ability of the surrounding environment, and can make dynamical change which is intelligent, independent, and adaptive according to the change of surrounding environment.

The research of the cognitive networks on the goal of end-to-end QoS performance optimization in China and in abroad are both at an embryonic stage, with only a rough conceptual model and a small amount of cognitive network routing algorithm (Sahoo A, 2002). The implement of QoS mechanisms in Cognitive network includes: classification and definition of QoS metrics, admission control and consultation, resource reservation, resource scheduling and management issues which are still lack of practicable solutions nowadays.
Based on cognitive network which has the advantage of dynamical reconfiguration and information interaction capability as the background, this paper is to carry out extensible QoS classification mechanism which possess sorts of index options, it executes the reasonable QoS classification according to the needs of the business. The mechanism takes all factors into consideration such as the character of different users and the change of network state. It performs real-time end-to-end QoS classified negotiation and revision before the start and also during process of business, in order that it could provide more suitable service for users, and also raise the utilization rate of resources. This paper proposes many reasonable indexes and mapping algorithm, designs the expansion interface, based on the actual demand of the business to dynamically increase the new QoS index according to the sorts of business. To design the QoS negotiation and adaptive adjusting strategy for progressing the end to end QoS negotiation and re-negotiation, it fully considers the network state information of the nodes of the cognitive network, then, it could adjust the QoS classification.

The structure of the paper is as follows: The second section analyzes the existing QoS mechanisms, the third section presents scalable QoS indicators system, the forth section designs and presents an adaptive QoS mechanism, the fifth section simulates the mechanism proposed, and the last section makes a conclusion.

2 LITERATURE REVIEW

At present, the research on QoS is mainly based on MPLS (Dekeris B and Narbutaite L, 2004), DiffServ (Lee G, 2008) and other models for the mechanisms optimization, as well as mapping algorithm of QoS mechanisms. While relatively less research done on refined and extensible QoS mechanism for diversification of the business. Most of them are at steps of tracking, introduction, digestion and absorption.

In July 2006, ITU-T adopted a recommendation called ITU-T recommendation Y.2111 (Yamada H et al., 2007). The proposal addressed a key area in NGN, provided end to end QoS capability. The recommendations involved resource and admission control functions (RACF) (Kamatani O et al., 2008) which will help operators to ensure the quality of end-to-end multimedia service in NGN, such as VoIP and IPTV. The key of the approach is the ability that operators designated criteria for specific types of traffic capacity, for better organization of the network resources. RACF meets the demand of more intelligent control based on packet network infrastructure. The recommendation defined the relevant requirements and functional architecture, including resource reservation, admission control and access control, Network Address Port Translation (NAPT) and firewall control, and Network Address Translation (NAT traversal).

Currently the more research architecture of end-to-end scalable QoS is a feasible framework presented by a scholar of Australia University of Technology academics (Hoang D B and Phan H T, 2007). This architecture took into account the distinction between the business service capabilities and operational QoS support capabilities for the IP network orientation during the design of network services. This architecture has the capabilities to support the DiffServ extensions and explores a new path congestion control method across multiple data streams of resources between the rational distribution and aggregation. Also, it can make QoS consultations among several domains by using QoS extended functions of BGP. This architecture uses a simple router control panel, and the sending panel does not require other complex PHB scheduling.

For the IP network, it is essential that its network architecture can provide differentiated services to applications, to ensure quality of service, and can be called service-oriented architecture. In this literature (Hoang D B and Phan H T, 2007), a new initiative QoS framework called End-Diff has many features required for this paper. This architecture is more scalable than DiffServ. Through simulation, this End-Diff QoS architecture has achieved good results on scalability, system overhead, routing strategy, fairness, packet delay and jitter conditions.

There was a scalable IP network QoS mathematical model (Fgee E B et al., 2008) proposed, taking into account the requirements of the network forwarding service assurance by VOIP, E-Commerce and other new services. In this model, a new QoS mathematical model was designed based on network calculus theory, which means the maximum rate of a new access transport stream is limited by a correlative arrival curve. The restrictions of transmission rate and time extension are realized by the service-by-hop delay time on the transmission path.

Some scholars proposed an internet transmission strategy (Chan H. et al., 2005) which can support scalable QoS security capability. For the disadvantage that the traditional port-based classification methods cannot obtain a valid result of
the business classification, such a mechanism achieves scalable QoS protection by distinguishing Internet transmission level by the application of business type. The strategy is divided into two steps, first, select the characteristics of the business which mainly based on information provided by the ISP; second, balance the system benefits and system complexity through the features of degradation, and then use REPTree bagging mechanism (Park J et al., 2006) to conduct service classification. From the above analysis we can see that most of the above research is QoS mechanism optimization based on service demands of single business, under the balance between the interests of the various network elements, optimizes the QoS guarantee mechanism, but not for the diverse needs of the business, particularly in the characteristics of cognitive network, to carry out the research of QoS mechanisms scalability and its implementation strategy.

There was a literature (Junghun P and Hsiao-Rong T, 2006) proposed a new approach that makes classification for network business based on scalable QoS application type. The results of using traditional port-based classification method were not satisfactory because the same port number may be shared by multiple applications. The non-equilibrium route, such as the PCF and the Net Flow as well as errors caused by modern measurement tools are reasons for this problem. In order to identify those problems, the paper described the classification process consists of two steps: feature selection and classification. First check those optional features easily learned by the ISP to balance the service performance and complexity.

Currently there is a new mathematical model for QoS (Ducatelle F. et al., 2005) in Australia which is based on the principle of network calculus. In this model, the inbound service flow is regulated by the arrival curve, and upper limit of arrival curve is maximum flow of the inbound service flow. On each node, the lower bound and delay time is determined by service waiting time before the business flow arrives the destination node. Compared with other models, the model has an acceptable end-to-end delay.

Among the domestic research results, there is a paper (Yunbo W, 2006) that analyzed the business features and characteristics of network needs of the Internet Large-scale multimedia applications. And it focused on requirements of the Large-Scale Application on network transmission reliability and scalability. Combined comparative analysis of the existing IP QoS plans and meets the QoS requirements of Large-Scale application, it proposes QoS control thought based on PCF and builds the corresponding support for Internet Large-Scale Multimedia Applications E2E QoS framework for the implementation of the system ON - QoS. Finally it do research about the aggregate flow of the ON-QoS scheduling strategy.

Another domestic study results (Suogang L et al., 2007) proposed a scalable multicast program for unordered QoS. The mechanism uses the order feature of the QoS level and chooses an appropriate multicast tree for multicast group by heuristic algorithm. So that the multicast receivers who are belong to different multicast groups and request the same QoS level can share the multicast tree. The simulation result showed that this program can effectively improve the scalability of multicast state, and in certain experimental settings the ratio of multicast trees and multicast groups may be less than 1 / 8; and can satisfy the different QoS receiver demand at the same time.

In the ChinaCom 2007 meeting, the Chinese scholars proposed a solution (Meng S and Bertrand M, 2008) that can make overlay service management aware of QoS information, so that QoS reminder service (QSON) with overlay network can be provided. In the QSON, the service components are organized into different subnet, and the QoS of service in the same subnet are similar. With the help of QSON, QoS has been unitopia integrated into such a similar information retrieval, so that the service query and choice can be easily obtained. Through a P2P-based services positioning system, service routings become with QoS conscious and scalable capability, and service routings are able to achieve good service performance.

3 SCALABLE QOS INDICATOR SYSTEM

As an important approach to improve resource utilization and overall performance, QoS security system almost penetrates any network with limited capacity in the design process and evaluation criteria. The most typically ones include distributed computing systems of the Internet, packet-based systems of telecommunication networks and computer networks. The QoS parameters of former ones include packet loss rate, delay, jitter, order delivery and error string, while the latter ones also include the availability, throughput and so on. At present, a lot of cutting-edge real-time media
streaming technologies, such as Voice over Internet Protocol (VoIP) and Internet TV technology (IP-TV), are very sensitive to the delay and require a relatively fixed bit-transfer rate. How to optimize and integrate the coding method, the priority distribution of grades (Prioritization) and real-time scheduling mechanism, under certain bandwidth constraints, to meet the stability and real-time network communication requirements is the core issue considered with priority.

Based on the feature that cognitive networks can be dynamically reconfigured, according to the principles of making and distribution QoS parameter index, in this section we designed a scalable QoS mechanism. This mechanism can increase the QoS business indicators dynamically based on business application requirements to provide accurate specific QoS classification for business with different characteristics to enable businesses to obtain more suitable services.

4 SELF-ADAPTIVE MECHANISM

Self-adaptive mechanism is an application-level QoS assurance mechanism, which mainly deal with the strategy of the parameter adjustment according to the user’s demands. Self-adaptive mechanism is the scope of QoS guarantee which can be defined as constraint by the user (application) for the QoS request strategies, through testing the state of the system and network resources and environment to adjust application strategy. It implements the suitable behavior to guarantee the user acceptable QoS request. It mainly reflects the characteristics are as following:

First, It is designed on the base of the specific type of application-centric. Therefore, different application types have different mechanisms of implementation. There is no single realization mechanism which is suitable for all application types.

Second, to solve the user's QoS Strategy level demand for parameter adjustment strategy, this kind of strategy demand proposed is mainly to adapt the network resource change.

Third, It shows primary role of the scope of end-node in the network, end system and above network level. Forth, It is one end-to-end realization mechanism, not only realize on an end system.

Fifth, It is one much more fine-grained network-level QoS to ensure the application of personalized expression.

4.1 Mechanism Constitution

The mechanism is mainly composed of three key components: detection, memory base, index mapping. As shown in Figure 1:

Before the service request arriving at the processor, first it wait in the waiting stack, and then to detect the network environment. Including the service type, the data type, as well as the network type of the goal terminal located. After obtaining the comprehensive record of network environment, then put the record into the memory base for comparing, examine that the historical grade decision-making which the computer adopts under the same network environment.

Based on the multi-dimensional analysis of the memory data base according to the different environmental information, if it has the same historic record, then calls the historic record in the memory base to carry on the index mapping as well as the grade decision-making directly. If the records are not the same, then calls the environmental information to the index collection to compare directly, takes the corresponding classification measure, and stores the top level result and the environmental information into the memory base to prepare for future use.

4.2 Detection

The testing process is divided into three categories, categories of detecting content shown in Figure 2:

(1) Service type detection: to detect the terminal of the service applicants for those who are using application softwares. It identifies the service request is issued by which applications, and thus obtains the information of the service type.
request sent out by the e-mail system, the audio document synchronized transmission request sent by the skype software and so on.

(2) Data type detection: it distinguishes between the service type, while detecting the data type of the files in the waiting stack. The file block size, block number, and the format of the waiting files, suffixes are mainly information for carrying on the type division.

(3) Terminal type detection: it uses cognitive radio technology to detect the signal types of the terminal service, such as measuring radio frequency band, spectrum and so on. So that it obtains the type of terminal is the wireless network or wired network, thus gets the corresponding transport strategy.

4.3 QoS Index Map

QoS is a comprehensive indicator to measure the satisfaction of a service using. Different multimedia applications require different service requirements, so these services must be parameterized. At the same time, different objects have different descriptions. For example, the QoS requests proposed by clients are only some brief descriptions, such as poor, in general, better, the best and so on. It is essential to map the user QoS to the application QoS in order to control the QoS consultation and the permission through the RSVP protocol. Application QoS re-mapped to the system QoS including network QoS for consultations. QoS description is that the user demands of the service quality, because users may not be computer professionals. Therefore, the audio QoS description is a brief, unspecific description of several service levels. But it is insufficient to describe the complex QoS structure which depends upon these brief level descriptions obviously. Therefore, the mapping from user QoS to application QoS, then to the system QoS, last to the network QoS must be completed before the QoS consultation.

4.4 Memory Base

The storage form of the classification historic record is the database relation table. It should call the memory data for comparing during the process of each new task pushed on, so that the classification step could be simplified. Simultaneously the historic records which are similar to the network environment make the logical partition.

First the simple two-dimensional relational table memorizes each historic record of network service, whose recording information includes: the network environment of the terminal services such as service type, data type, other side terminal service type, during the time service occurred, the record name (auto-coding). And it stores with the record name and the environment name as the primary keys. Service time, service type, and other fields are as foreign keys. They connect with the two-dimensional relational table of other types named in order to facilitate the logical partition.

<table>
<thead>
<tr>
<th>Field</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Environment</td>
<td>Text (primary key)</td>
</tr>
<tr>
<td>Time of reservoir inflow</td>
<td>Time (foreign key)</td>
</tr>
<tr>
<td>Type of service</td>
<td>Text (primary key)</td>
</tr>
<tr>
<td>Type of data</td>
<td>Text (primary key)</td>
</tr>
<tr>
<td>(Each indicator corresponds to the data)</td>
<td>......</td>
</tr>
<tr>
<td>Grading result</td>
<td>Numeric</td>
</tr>
</tbody>
</table>

In order to simplify the management of a database large table, we divide the mechanism into two types according to the modules: service type, data type, to carry on logical partition management. Along with the accumulation of system operation time, the memory base has a gradual increase in total memory. If everyone thinks that searching key words will cut the working efficiency, the partition separate the...
table into several different table spaces. It uses “divide and rule” method to support an unlimited expansion big table, which makes the large table controllable at the physical level. Dividing the large tables into smaller partitions, include collection, excavation and analysis of the multi-dimensional data such as the service type, data type, run-time, so the table can be improved in the performance of maintenance, backup, restore, transaction and query.

The specific reason of using logical partitions as following: Enhancing usability which means if one partition of the table can not be used due to system failure, the remaining good partition of the table can still be used; reducing the closing time which means if the system failure affects only a part of the table partition, then only this part of the district need to be repaired, it could take less time to repair than the large table; easier maintaining which means if you need to rebuild the table, the independent management of each partition is much easier than manage a single large table; Balancing I/O which means we can balance the I/O by assigning different partitions of the table into different disk; Improving performance which means query, add, modify and other operations of the large table can be broken down into different partitions of the table to execute much faster. The partition is transparent to the users, users will not feel the partition existent.

4.5 Features of Self-adaptive Mechanism

4.5.1 Intelligent Learning Mechanism

The key point of cognitive networks making improvements of the performance is that it has specific learning mechanism. The behavior model of their network elements is: According to different network conditions, the results of adjusting and saving will be classified for storage. When meeting similar problems again, taking into account of historic record, CN makes appraisals to each kind of question solution's fit and unfit quality and selects a relatively optimal strategy. Therefore, based on the network behavioral model, CN may carry on the revision and the expansion to the existing QoS classification mechanism, and satisfy the various service type and the personalized user demands. Also CN can predict network behavior by network status sensors to make reasonable allocation of traffic flow; can use the smart probe packets to establish redundant transmission path to support real-time streaming media business better and better; can use the resource reservation of network neatly to set aside resources to achieve optimal use of distribution according to need and goals.

4.5.2 QoS Needs for the Diverse Business

Before the business start, CN carries on the end-to-end QoS classification consultation, based on the actual situation of the network, to make reasonable adjustments of the QoS business level. When the service is on operation, we could obtain the variation situation of end-to-end path status by using the capacity of information interaction of CN to carry out QoS re-negotiation. And also it makes adaptive adjustment of QoS service classification to ensure the classification of QoS service is reasonable and reliable.

5 SIMULATION

This article uses the OPNET (Kubera E et al., 2004) to simulate, compares and analyzes performances of network delay, scalability, loading between the existing mechanisms and self-adaptive mechanism proposed in this paper.

5.1 Delay

The picture below are describing two sets of simulation results comparing, thick line did not call the memory base delay, thin line represent for the use of adaptive mechanism, which means the delay of calling the memory base. The simulation results are shown in Figure 3.

![Figure 3: The simulation result of delay.](image-url)

Queuing delay is determined by the service rate and packet size under the use of self-adaptive mechanism, that is called delay compared before and after calling the memory base as shown in Figure 3. Figure 3 indicated the average time of two nodes data packets waiting in the service side. By the chart,
delay time is about 2.40 sec before calling in the memory base; and the delay time is about 0.25 ms after calling the memory base. For servers, the shorter delay time, the higher efficiency is in response to user requests, which will help to improve the quality of the service. Therefore, the self-adaptive mechanism plays a significant role in improving service quality.

5.2 Network Scalability

It named the original network under the use of the adaptive mechanism Scene 1, and then expand the number of network clients to twice of the original, and named it the Scene 2. It shows network delay curves after running simulations as Figure 4 below.

![Figure 4: The simulation result of network scalability.](image)

In the figure, the thin curve represents for the delay curve after using adaptive mechanism, the thick curve represent for the delay curve after expanding the number of network terminals into twice of the thin one which also uses self-adaptive mechanism. As shown in Figure 4, the two curves almost overlapped. Although the number of the network users increased doubled to the original one, the network delay has not increased significantly. The good network scalability is a reflect of ensuring high quality of service during meeting a large quantity of service in the network, in other words, it is a necessary condition of satisfying the diverse service to assure the service delay basically invariable in dealing with large quantity of diversified service. Therefore, the use of adaptive mechanism can supply grading and coordinating services quickly to ensure smooth network, and it is a good guarantee of network expansion.

5.3 Network Load Capacity

This section simulated two network environments, in which an adaptive mechanism is not invoked, and the other is called adaptive mechanism. The figure 5 shows the network load capacity curve after running simulation.

![Figure 5: The simulation results of network load capacity.](image)

As shown in the figure, the thin curve represents for the original network load curve, the thick curve represents for adapting the mechanism, the thick line is higher than the thin line level equally, thick line index is 2000 - 5000 bit/sec compared with thin line index at each time point, which means capacity of load improved. A good load capacity of the network means network can complete the large number of services and data quantity at the same time point, that is, shorten the delay time, while ensuring the efficient data transmission. Therefore, the use of adaptive mechanism can improve load capacity.

6 CONCLUSIONS

This paper proposes a set of scalable QoS indicators and designs and makes simulation of self-adaptive mechanism based on the current situation in the CN. The mechanism will detect the actual status of the network and make end to end QoS classification consultations and adjustments before business start. We can obtain the changing situation of end-to-end path status on time by using the information interaction capability of CN. Then we make QoS re-negotiation to adjust QoS classification of business adaptively when the user business is on operating. At last, the results are classified to import into memory base due to the characteristic of service. It will greatly reduce the negotiation time and improve the efficiency of the network service when meeting the same situation we could directly call the classification results of the memory base.

This mechanism focuses on solving the problem of personalization and diversification users for protection of QoS in the cognitive network. It fully enhances the utilization of network resources and
meets the complexity, heterogeneity, and reliability requirements of the new generation information network, and ultimately it achieves performance optimization of end-to-end.

In this paper, the design of adaptive mechanism still need improved, such as increasing memory base integrated, the fuzzy inquiry function, storage capacity and so on. Besides, though the simulation of this self-adaptive mechanism is only a simulation for three indicators, we could make simulation and comparison more comprehensive to much more indicators.

REFERENCES

Zhang, L., Yang, W., Wang, W., Guo, L., Yang, Y., 2007. A Method to Enhance Service Survivability Based on Autonomous Configuration. DOI10. 1109 / IMSCCS. 44.


