Advanced Science and Technology Letters
Vol.111 (NGCIT 2015), pp.191-195
http://dx.doi.org/10.14257/astl.2015.111.37

Load Balancing Algorithm Based on QoS Awareness Applied in Wireless Networks

CHEN Xiangqian, MA Shaohui

Department of Computer Science and Technology, Henan Mechanic and Electrical Engineering College, Henan Xinxiang, 453000, China

Zhao Kai

Teaching Affairs Division, Pingdingshan University, Henan Pingdingshan, 467000, China
xqChenxx@126.com

Abstract. In order to reduce business latency and packet loss rate, to improve the throughput of integration of heterogeneous wireless network, to achieve load diversification and to improve end quality of service, while there are many problems on the dropout rate of the traditional load balancing algorithm in processing throughput, delay and business, therefore, gateway load balancing algorithm is proposed in this paper. All terminals in the access network can reflect the effectiveness of the average network load level of the network, this algorithm will gain weight of the load in the network of small business terminal to switch to the network load which is light. First, it defines the heterogeneous networks and network terminal payoff function utility function, which are used to characterize the experience and network QOS terminal load situation, and then presents the specific processes of the gateway load balancing algorithm. Finally, switching decisional load balancing algorithm, proposed by Yan X, etc. is compared with simulation experiments and it shows that: the proposed gateway load balancing algorithm has strong robustness to achieve network load balancing and to achieve a balanced use of network resource.

Keywords: efficiency, loading differences, real-time business, switching decision

1 Introduction

Load balancing is an important way, which makes full use of heterogeneous wireless network resources. Through the load balancing among the heterogeneous networks, the high probability of network load can be reduced, the overall utilization of network resources is improved and the blocking probability is reduced to provide users a better QOS, so the load balancing between the networks is also an important aspect of considering the access algorithm selected [1-3]. The user can only access a kind of network at the same time with the traditional hard load balancing service, which can not satisfy differentiation of the user service, resulting in partially utilized heterogeneous network resources and higher traffic blocking ratio.

ISSN: 2287-1233 ASTL

2 Proposed Scheme

The core idea of algorithm is: the QOS income of terminal end can reflects the current level of network quality of service received, the greater the benefits is, the better the resulting quality of service is, and vice versa; the average QOS benefit of all terminals in a accessing network (hereinafter defined as a network utility) can reflects the load level of the network, the larger the average gain of the terminal is, the load on the network is lighter, and vice versa. In order to achieve diversification and to increase the service quality of load terminal, the algorithm will switch small businesses with gain heavy load in the network to a lighter terminal network load.

In order to carry out the terminal to select the RAN and its access effectively and dynamically, the multimode terminal will experience unified management from the network side, which will be completed by the network side management entity (Network side manager, NSM). The terminal will interact with NSM through the terminal-side management entity (Terminal-side manager, TSM), dynamically achieving refactoring of switching/accessing. The interaction between NSM and TSM is completed through management control channel (Management and control channel, MCC). NSM is deployed in the core network, and is shared by a plurality of RAN. RAN will convey each context information to NSM, then NSM transfer each context information of RAN to the terminal for the decisions through the MCC downlink transmission. TSM of each terminal sends context information of the terminal through uplink transmission of the MCC to NSM. Based on context information of RAN and terminal, NSM uses the appropriate network selection algorithm to develop strategies and policies issued under the various terminals. Terminal then chooses according to their needs and network reconfiguration decisions and configures to access the appropriate RAN. The paper will assume the terminal in the network / inter-cell handover fast enough, and thus, the load balancing process, due to switching delay caused by the upper risk of business disruption can be ignored.

Different types of wireless network services have different QOS requirements. Based on the characteristics of the various services, wireless services can be classified into three types of basic services, as shown in Table 1.

Table1. Wireless network traffic types and their QOS requirements

	Rate of change	The minimum rate	Maximum rate	Time delay requirements	A typical business
Real-time business	constant	$r_{\min} = r_{\max}$	$r_{\min} = r_{\max}$	$d_{\max}$	VoIP
	variable	$r_{\min}$	$r_{\max}$	$d_{\max}$	MPEG

The real-time business	variable	r_{min}	r_{max}	φ	HTTP

According to real-time requirements, the wireless business can be classified into real-time business (Real time, RT) and non-real-time service (Non-real time, NRT). According to whether the rate can change or not, real-time service has been divided into two types of constant rate and variable rate. Minimum service time at a constant rate min and the maximum rate requirements max are equal, that is the rate are unchanged. The variable rate real-time service is possessed of the minimum and maximum rate requirements. When the waiting time of a real-time service exceeds the maximum packet delay tolerance d_{max} , the packet will be discarded. Non-real-time service needs not to delay, and the minimum bandwidth can become zero.

In order to characterize the obtained terminal from the current network quality of service, based on the QOS requirements of different business, si_{gmond} function to construct the terminal payoff function is used. So suppose J is the set of kinds of RAN, I represents the set of terminals in the network, defining an accessing terminal $i \in I$ of RAN $j \in J$ QOS benefit function is as follows:

$$U_{ij} = \begin{cases} (1 - \dfrac{1}{1 + \exp(-\partial_m \times \dfrac{d_{ij} - d_{ij}^{e}}{d_{ij}^{max} - d_{ij}^{m}})}).BT \\ (1 - \dfrac{1}{1 + \exp(-\partial_m \times \dfrac{d_{ij} - d_{ij}^{e}}{d_{ij}^{min} - d_{ij}^{r}})}).BRT \end{cases} \quad (1)$$

For real-time services, at the premise to meet the minimum bandwidth, using an average delay d_{ij} to measure user gains, the smaller the time delay is , the higher the gain is; for non-real-time services to the user, using average speed r of the user gains to measure, the larger the rate is, the higher the income is;

$$\frac{d_{ij} - d_{ij}^{e}}{d_{ij}^{max} - d_{ij}^{r}}$$

If delay for real-time services the normalized, wherein d_{ij}^{e} denotes the expectancy of average delay of real-time services:

$$\frac{\beta_{ij} - \beta_{ij}^{e}}{\beta_{ij}^{max} - \beta_{ij}^{r}}$$

As the normalization about Non-real-time services' rate, it helps to ensure a minimum rate of non-real-time services; β_{ij} and β_{ij} is constant parameters, which determines the steepness of the curve of the function, the larger the value is, the steeper the curve changes, the higher the sensitivity to the end quality of service is. Formula (1) as defined in revenue function reflects QOS-awareness of terminal, the function maps a plurality of QOS parameters with reasonable perception or experience for the user to QOS level, gives a measurement of the QOS of different users by using uniform quantization levels standards.

To characterize the load level of the accessing network, the wireless access network defines the utility of all the terminals connected to the network average of QOS benefit. Suppose at a time, a terminal can only access a RAN, then the gain RAN $J \in I$ can be expressed as

$$R_{ij} = \begin{cases} \dfrac{\sum_{j \in i} \alpha_{ij} n_{ij}}{\sum_{J \in I} \alpha_{ij}}, & \sum_{I \in J} \partial_{ij} \neq 0 \\ 1, & \sum_{i \in j} \partial_{ij} = 0 \end{cases} \quad (2)$$

Wherein, $\partial_{ij} = \begin{cases} 1 \text{ terminal } i \text{ into } RAN_j \\ 0 \quad other \end{cases}$

Obviously, the heavier network load will result in lower average QOS benefit of terminal; otherwise the terminal average QOS gains will be higher. Therefore, the average QOS benefit of the terminal, namely the network load of the network utility can reflects the situation. Network utility is higher, indicating that the network load is lighter, otherwise it indicates the network load is heavier.

3 Conclusion

For the integration of heterogeneous wireless network communication scenario, the network load balancing algorithm was proposed based on the QOS, which is fallen into switchable load balancing algorithm. Algorithm is based on QOS gains and network utility of a generalized feature for wireless business users, and it is able to characterize the quality of service terminal network experience and network load conditions, and it is universally applicable for a variety of heterogeneous network. Among them, the universality of network utility to heterogeneous network makes heterogeneous network load comparability comparable, able to achieve switching between heterogeneous network load balancing. Simulation results show that the proposed gateway load balancing algorithm can improve integration of heterogeneous network throughput, reduce business latency and packet loss rate, with strong robustness to achieve network load balancing and to achieve a balanced use of network resources.

References

1. Kasman Suhairi, Ford Lumban Gaol, The Measurement of Optimization Performance of Managed Service Division with ITIL Framework using Statistical Process Control. Journal of Networks, Vol 8, No 3 (2013), 518-529
2. Guang Yan, Zhu Yue-Fei, Gu Chun-Xiang, Fei Jin-long, He Xin-Zheng, A Framework for

Automated Security Proof and its Application to OAEP. Journal of Networks, Vol 8, No 3 (2013), 552-558

3. Yishuang Geng, Kaveh Pahlavan, On the Accuracy of RF and Image Processing Based Hybrid Localization for Wireless Capsule Endoscopy, IEEE Wireless Communications and Networking Conference (WCNC), Mar. 2015
4. Jie He, Yishuang Geng and Kaveh Pahlavan, Toward Accurate Human Tracking: Modelling Time-of-Arrival for Wireless Wearable Sensors in Multipath Environment, IEEE Sensor Journal, 14(11), 3996-4006, Nov. 2014

