Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(7):651-658

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Dynamic load balance scheme based on access selection and calls transfer in heterogeneous wireless networks

Zhu Si-feng[*1,2], Zhang Xin-gang[3], Shen Lian-feng[1] and Chai Zheng-yi[1]

[1]*National Mobile Communications Research Laboratory, Southeast University, Nanjing, China*
[2]*School of Mathematics & Statistics, Zhoukou Normal University, Zhoukou Henan, China*
[3]*School of Computer Science, Nanyang Normal University, Nanyang Henan, China*

ABSTRACT

In heterogeneous network environment integrated multiple radio access technology, the nonuniform distribution of users' calls in time and space is the important factor affecting the performance of heterogeneous wireless network. In order to improve the performance of heterogeneous wireless network, a novel dynamic load balance scheme based on access selection and calls transfer was proposed. In this paper, access selection was modeled as a constrained optimization problem, the access calls was assigned to all base station evenly in accordance with the optimization result. In addition, in order to weaken the destructiveness of sudden calls from hot area to the system load balance, a calls transfer strategy based on the threshold of base station load rate was given. Experimental result shows that the proposed solution performs better than literature solutions in access blocking rate, calls dropping rate, load balance and system utilization rate, and has the advantage of good application value.

Key words: Heterogeneous wireless network; loads balance; access selection; calls transfer strategy

---

INTRODUCTION

Heterogeneous wireless network is a heterogeneous mixed network [1] of GSM, WLAN, Bluetooth, 3G, B3G, 4G and etc. Because there are differences in calls capability, coverage capability and technical level of various wireless access networks, any of a radio access network can not fully meet the demand of coverage, delay, transmission rate, and cost and so on. In order to provide a continuous, seamless mobile calls to users, joint radio resource management (JRRM) system came into being. Load balancing control is an important task of joint radio resource management, and its purpose is to enable users calls evenly distributed in various wireless access networks, in order to ensure the stability of the entire heterogeneous wireless network system. Load balancing can not only improve the stability of the entire heterogeneous wireless network system, but also bring the system throughput, blocking rate, call dropping rate and other aspects of performance gain. From the point of view of queuing theory, the heterogeneity of spatial and temporal distribution in the user traffic network can be overcome by load balancing control, thereby improving overall service performance across the entire heterogeneous wireless network system.

The development trend of next generation in wireless communication network is the coexistence of multiple wireless access technologies and merging, in a heterogeneous network environment, when a new service access, if we choose an optimal network access, then you can bring multi-radio access gain(MAAG).

In the user-intensive areas, a communication hot spot is usually formed. If we introduce flow equalization in the heterogeneous cell with different bandwidth and capacity, we can reduce the blocking rate of new calls and the dropping rate of switching operations, also the utilization across heterogeneous systems will increase accordingly.

Access selection and calls transfer are an effective way to achieve load balance, in view of this, we propose a

dynamic load balance scheme based on access selection and calls transfer in the heterogeneous wireless network, and we conducted experiments and analysis for the proposed scheme under simulation environment.

In recent years, the load balancing issue in heterogeneous network environment attracted wide attention of scholars at home and abroad. Considering the impact of MAC layer packet retransmission on load balancing in link adaptation mechanism, a multi-access option for packet calls was proposed in the literature [2]; Based on the assessment of average packet calls in the access and network resources consumed by the current load situation, a load-balanced access control scheme was proposed in the literature [3]; Based on fuzzy neural network of radial basis function, the access selection method was proposed in the literature [4], which can be connected to the network by blocking rate equal to reinforce learning objectives for fuzzy neural network parameters, and it had a great dynamic adaptability for the network load degree, also to achieve intelligent access decision; Literature [5,6] proposed a heterogeneous network access control scheme based on game theory; Literature [7] proposed a dynamic load balancing mechanism based on terminal mobile and calls heterogeneous, which focused on the calls potential users, through the joint service access control of dynamic optimization, so as to keep the traffic of every access point and communication resources maintain matched to realize the load balancing of networks; literature [8] proposed an improved dynamic threshold joint load control method, in order to adapt to load balancing requirements in different load conditions, to achieve efficient utilization of resources. While combined reconfigurable features of terminals and diversity of calls requirements, and used a terminal selection algorithm based on AHP in the vertical handover procedure to perform a network-initiated, resulting in a balanced heterogeneous network load and reducing the cost of systems and the impact for users; literature [9] considering the survival time of terminal battery and load balancing, a vertical handover scheme was given by putting two goals into a goal to solve on weights; literature [10] proposed a vertical handover scheme based on SINR and AHP; literature [11] conducted a study on the vehicle heterogeneous network environments; In literature [12], Bidding model was established on the relationship between mobile phone users and access networks, and the vertical handover scheme are given based on cooperative game.

Literature [2-7] was to achieve load balancing of heterogeneous network environments through joint access control; literature [8-12] was to achieve load balancing through vertical handover mechanism. Unlike the above ideas, in the paper a new dynamic load balance scheme is proposed, to achieve two strategies of access control and calls transfer.

THE SYSTEM MODEL OF HETEROGENEOUS WIRELESS NETWORK

Assuming there are a lot of wireless access network in area Ω, cooperative mode of wireless access network is a very tight coupling mode, i.e. in the area of Ω each base station of the wireless access network is connected to core network via the same control unit. Usually, the access router (AR) acts as a control unit in the region serving, which is directly connected to the core network, shown in Figure 1.

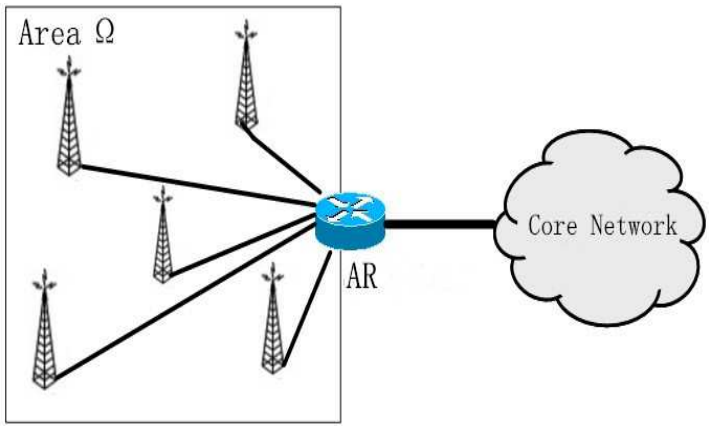

Fig.1 The system model of heterogeneous wireless network

In the network architecture, access routers unify and coordinate data traffics of communication between base stations, to optimize the use of network communication resources through the joint service access control between base stations.

In the system model, when multi-mode mobile terminal (MMT) has a new call or vertical handoff calls arrives, MMT can select the most suitable base station to establish calls links between multiple base stations. From the perspective of optimizing the overall performance of the network, the access router acts as the central control module, according to optimization criteria, to dynamically allocate all users' traffic to each base station within the

coverage area, so that the wireless access traffic of each base station are matched to communication resources, to achieve traffic loads balance of network unit.

People's living and working patterns may lead to the inhomogeneous distribution of traffic on the time domain and space domain, for example, some public places (shopping malls, airports, hotels, etc.) formed a great hot cell when communication traffic grow in certain period of time, resulting that base station of the region is heavily loaded, while other areas are loaded very lightly. In addition, there are also differences in the capacity between each base station, which can lead to load imbalance phenomenon. Therefore, when the load imbalance was found, we need to adopt calls transfer strategy.

Thus, access selection and calls transfer are important means of the loads balancing control, access selection can be seen as a preventive load balancing control strategy, while calls transfer is a remedial load balancing control strategy, and the two are closely related.
Compared with the 2G network, bandwidth of 3G networks has been greatly improved. Voice service takes the less of 3G network's bandwidth, but non-real-time calls (such as network video download) take up a lot of network bandwidth. The load imbalance of wireless access network is mainly because non-real-time calls are unevenly distributed in space and time. Therefore, the study focuses on the access selection and calls transfer of non-real-time calls.

DYNAMIC LOAD BALANCE SCHEME BASED ON ACCESS SELECTION AND CALLS TRANSFER
The assumption that in the region Ω, there are a plurality of wireless access networks, m base stations (or wireless access point) are deployed: BSs= {BS1, BS2,…, BSm}, base station BSi reserved maximum bandwidth Bi for non-real-time service.

This article assumes that: a terminal at a time can be up to a non-real-time service, the bandwidth of the non-real-time service to terminal MTj is bj. Assuming the region Ω, with n multi-mode terminals MTs={MT1, MT2,…, MTn}, when the multi-mode terminal MTj has calls arrived, we can choose the most suitable base station BSi to establish calls links in the m base stations.

From the perspective of optimizing the overall performance of the heterogeneous network system, according to optimization criteria, the access router dynamically assignes all users' traffic to the base station BSi within the coverage area, so that the wireless access traffic of each base station BSi are matched to communication resources, to achieve traffic loads balance of the region Ω.

(1) Access Selection
Because the wireless access network use different access technologies, and the load and the form of the same type of calls generated by access network may be different. Therefore, it is necessary to targetedly select the most appropriate parameters to describe and measure the load factor of heterogeneous wireless access network,for comparison. In this paper, take the network UMTS (Universal Mobile Telecommunications System) and WiMAX (Worldwide Interoperability for Microwave Access) for example, the load rate metrics were given.

Considering the solving method of downlink access selection problem to UMTS is very similar to uplink access selection, so this article only take the downlink case as an example to model , on the basis we can get the corresponding uplink model slightly modified. Assuming the downlink power of UMTS base station is always constant, and is uniformly distributed over the entire available bandwidth. The downlink total power of the UMTS base station is by means of P_i, then the transmission power of the base station i in the UMTS network per unit bandwidth is:

$$p_i^0 = \frac{P_i}{B_i} \tag{1}$$

According to (1), if the bandwidth of non-real time service needs to terminal MT_j is b_j, then the required total transmit power is $b_j \cdot p_i^0$. This has for the wireless access technology UMTS of feature of "soft capacity", the paper is to measure the load level based on the load power factor of the cell [8]. Taking the asymmetry (usually the downlink traffic are more than the uplink traffic) of the non-real-time calls into account, and the downlink capacity is generally less than the uplink capacity, the paper examines the downlink load factor, as in formula (2).

$$\eta_{DL}^{UMTS} = P_{\text{current}} / P_{\max} \tag{2}$$

Where, P_{current} and $P_{\max}$ represent the total transmission power and the maximum total transmit power required for the current base station downlink.

WiMAX is an emerging broadband wireless access technology, which uses high technologies (OFDM/OFDMA, MIMO and etc) represent the future direction of development of communication technology, and can provide high-speed Internet connection for data transmission distances of up to 50km. The network WiMAX has advantages of QoS guarantee, high transfer rate, and diverse calls. WiMAX (IEEE802.16e) as a new network can meet the needs of high-speed data transmission in the communication environment. Interconnection of WiMAX (IEEE802.16e) and existing 3G wireless network is a complement to existing 3G wireless network, it can take advantage of the core network of existing wireless network, reducing the investment of the WiMAX's core network; also it can fully tap existing network users' information, sharing the billing, authentication and encryption mechanisms of existing network, to fully rely on the resources of existing network and service platform to launch new calls. WiMAX (IEEE802.16e) and 3G wireless network can provide data calls under certain mobile conditions, forming the situation of competing and complementary in a certain range. WiMAX (IEEE802.16e) is a wireless access technology of high bandwidth, the paper uses bandwidth-based cell load factor to measure the load factor level, as in formula (3).

$$\eta^{WiMAX} = B_{\text{current}} / B_{\max} \tag{3}$$

Where, B_{current} and $B_{\max}$ represent respectively the total current required bandwidth and the maximum bandwidth of the base station.

Let the service access request case of base station i receiving multi-mode terminal j to $x_{ij} \in \{0,1\}(1 \le i \le m, 1 \le j \le n)$, under the action of joint access selection strategy, if access router selects the service of multi-mode terminal j to access to the base station i, i.e. assigns b_j bandwidth to multi-mode terminal j from the base station i, then the value of x_{ij} is 1; otherwise, 0. Since each call can only access to a base station at the same time, so formula (4) establishes.

$$\sum_{i=1}^{m} x_{ij} \le 1, \forall j(1 \le j \le n) \tag{4}$$

In access selection mechanism of the paper, the optimization of network performance takes the load rate of each base station as the target, to minimize the gap of load factor between BSi(i=0,1,2,…,m), explore how to achieve load balancing between different BSi at the maximum extent. Therefore, we define an access selection optimization model, which is as follows:

$$\min \ (\frac{\max \ \eta_i}{\min \ \eta_i}), \ 1 \le i \le m \tag{5}$$

$$S.T. \begin{cases} \sum_{i=1}^{m} x_{ij} \le 1, \ \forall j(1 \le j \le n) \\ \sum_{j=1}^{n} x_{ij} \cdot b_j \le B_i, \ \forall i(1 \le i \le m) \end{cases} \tag{6}$$

Wherein, the formula (5) represents the gap of load factor to the minimum base station cell; formula (6) shows a non-real-time service can access up to a base station, and each station can not be overloaded.

Access selection processes designed in this paper are as follows:
Step 1: Timer is reset, start counting.
Step 2: When new calls reaches in multi-mode terminal MTj, multi-mode terminal MTj sends the access request contained calls demand for bandwidth to the cell's base station BSi.

Step 3: The base station BSi sends the received access request information to the access router.
Step 4: The access router gains optimal access scheme by formula (4) and (5).
Step 5: According to obtained optimal access scheme, the access router notifies a base station BSi is responsible to access the calls by broadcasting.
Step 6: After the base station BSi receives broadcast information, assign the appropriate bandwidth to access calls of the multi-mode terminal MTj.
Step 7: If the timer expires, all base stations send the current loads, assigned bandwidth information to the access router, go to Step 1; otherwise, go to Step 2.

(2) Calls Transfer
For the capacity of different base stations to access networks may be different, the throughput index can not be used to measure the load of base station. By using the method of setting two load-rate thresholds for each base station: a high load-rate threshold, if exceeds the threshold, it is considered the base station is heavily loaded; a low load-rate threshold, if below the threshold, lightly. If the base load rate is between the two thresholds, the base station is considered to be in equilibrium.

Let the current load rate of base station i is ηi, which high load-rate threshold is δi, low load-rate threshold is σi. Load state of base station i is defined as follows:

$$f(\eta_i)=\begin{cases}-1,\text{if } 0\le\eta_i\le\sigma_i;\\ 0,\text{ if } \sigma_i<\eta_i<\delta_i;\\ 1,\text{ if } \eta_i\ge\delta_i;\end{cases} \tag{7}$$

Wherein, -1 indicates a light load; 0 indicates equilibrium; 1 indicates overloading.

Access selections only try to ensure that the calls loads are evenly distributed within every base station cells of heterogeneous network systems. When there is a sudden event (a communication hotspot is suddenly formed in certain time and region), this may result in a rapid rise in the calls of certain base station's cell, so that the serving base station is in overloaded state, and the calls of other adjacent base stations are few, in light-load condition. In the case, its need to execute the calls transfer strategy, transfer part non-real-time calls of high-load base station's cell to light-load base station's cell, to ensure the loads balance across entire heterogeneous network system.

For non-real-time calls, which are transferred from one cell to another through the vertical handover, the impact of smaller delay may be not large for QoS. But if it is unfortunate to be asked to perform passive vertical handovers many times, the QoS will be seriously affected. Therefore, in order to ensure the QoS of calls transfer, which performs passive vertical handover is seen as a mark, when a base station's cell selects calls to transfer, non-real-time calls with the sign of transfer are not selected.

Different base stations of access network, it is greatly different in capacity. After non-real-time calls are transferred from high-load base station's cell s to light-load base station's cell t, the bandwidth consumed by the non-real-time calls may result in base station cell t becomes from a light-load state to a overloaded state, and it's need to transfer non-real-time calls from base station cell t to light-load base station's cell y again. This may lead to ping-pong switching phenomenon. To avoid the ping-pong switching phenomenon, in the paper, we make a pre-judgment operation when calls transfer, to ensure until the transferred non-real-time calls don't lead to the base station cell becomes from light-load state to overloaded state, the calls transfer.

This design of calls transfer process is as follows:
Step 1: Scheduling cycle timer is reset, start timing.
Step 2: The access router collects information of each base station cell, computing the load state of each base stations cell .
Step 3: The access router ascendant arranges the base station in the light-load state for queue L1 according to the load rate; reducibly arranges the base station in the overloaded state for queue L2 according to the load rate.
Step 4: If the queue of overloaded base station L1 is empty, go to Step 10; otherwise, go to Step 5.
Step 5: Take the head BS11 from overloaded base station queue, take the head BS21 from light-load base station queue.
Step 6: Select calls MT11 of the most consuming bandwidth from the base station BS11.
Step 7: If the calls MT11 is transferred from the base station BS11 to BS21, BS21 remains in a light-load state, calls transfer operation is executed, and move BS21 to the queue L2's tail; if the calls MT11 is transferred from the base

station BS11 to BS21, BS21 will become in an equilibrium state, calls transfer operation is executed, and remove BS21 from the queue L2.
Step 8: Delete the head station BS11 from the queue L1.
Step 9: If the scheduling cycle timer times out, then go to Step 1; otherwise go to Step 3.
Step 10: End.

EXPERIMENTS AND ANALYSIS

Assuming: in the region Ω that the cellular networks WiMAX and UTMS overlapping coverage, there are 3 UTMS base stations and 2 WiMAX base stations, 100 multi-mode terminals. Maximum bandwidth reserved for non-real-time calls are 2 Mbps and 20 Mbps in UTMS and WiMAX base stations, the average bandwidth of each non-real-time calls needs is 128 kbps. Overloaded threshold and light-load threshold of the base load rateare respectively 0.8 and 0.3.

Experimental computer systems' configuration: HP Z800 Workstation, dual-core CPU (Intel Xeon Quad-Core W5580 3.2GHz), 12GB RAM, 450GB 15000 rpm HDD, windows 7. In order to verify the performance of dynamic loads balance scheme proposed in this paper, in the experimental environment of MatLab7, conducted comparative experiments with the literature [3], [12]. At the beginning of the experiment, multi-mode terminals uniformly distributed in the region Ω. In order to simulate the terminal moving case of the actual environment, we use a random mobility model[13].

The main indicators to evaluate the performance of loads balance algorithm are: new calls access to blocking rate, dropping rate of calls switching, load balance and system utilization rate. In the paper, the schemes of literature [3] [12] and the article are respectively of running 1000 seconds until the results are to converge in the same simulation environment, the following is comparative analysis of simulation results.

Figure 2 describes the performance differences of three schemes in non- real-time calls access blocking rate. As can be seen from the Figure 2: With the increase of arrival rate in non-real-time calls, calls access blocking rate of three schemes emerge upward trend; calls access blocking rate of the article is rising the slowest, while the literature [3] rising the fastest. This shows that the scheme of this paper outperforms the literature [3] and [12] on the index of non-real-time calls access blocking probability. The reason is that the scheme gets the best access to distribution scheme by using of optimization algorithm when access selection, the loads distribute balanced on all base stations of heterogeneous systems, so that the system has a good calls access capability.

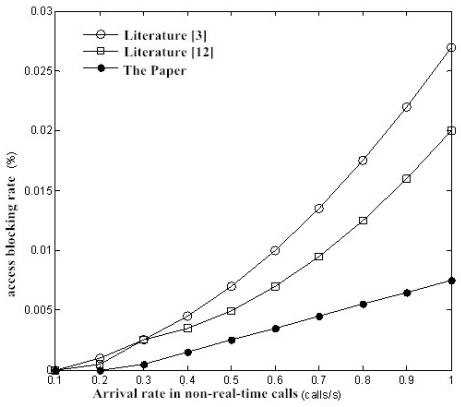

Fig. 2: The performance differences of three schemes in non- real-time calls blocking probability

Figure 3 depicts the performance differences of three schemes in calls dropping probability. As can be seen from figure 3, with the increase of arrival rate in non-real-time calls, calls dropping probability of three schemes emerge upward trend; switch dropping rate of this article is rising the slowest, while the literature [3]'s rising the fastest. This shows that the scheme of this paper outperforms the literature [3] and [12] on the dropping probability index of non-real-time calls. The reason is that the scheme sets up the load rate threshold based on switching strategy of calls transfer, avoiding the ping-pong switching phenomenon, thus effectively reducing the handover dropping rate.

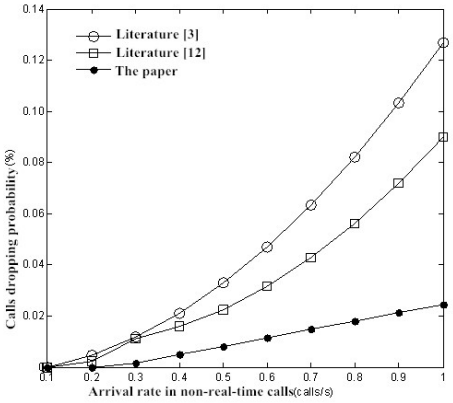

Fig. 3: The performance differences of three schemes in calls dropping probability

The ratio of standard deviation and mean is called the Coefficient of Variation (CoV)[9], which is a common indicator of measuring loads balance, the smaller the value, the better loads balance. The CoVs of the three schemes in test scenarios are shown in Figure 4. As can be seen from figure 4, the loads CoV of the scheme is the smallest that is loads balance of the article is the best; the scheme of literature [12] worst; the scheme of literature [3] is slightly better than the literature [12]. The reason is: This scheme, on the access selection strategy, obtains the optimal access solution by use of optimization algorithm, making calls distribution located in each base station cell relatively uniform. In addition, the paper uses calls transfer strategy, to effectively balance sudden business of the hotspots, so that the loads are more evenly distributed throughout the base station cells of heterogeneous system. The scheme of literature [3] adopts the access selection strategy based on loads balance, so it can be slightly better than the literature [12] on the index of loads balance.

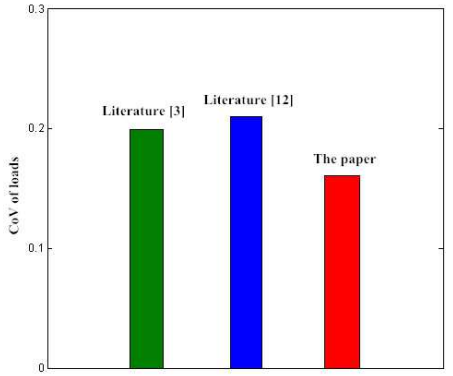

Fig. 4: The performance differences of three schemes in load balance

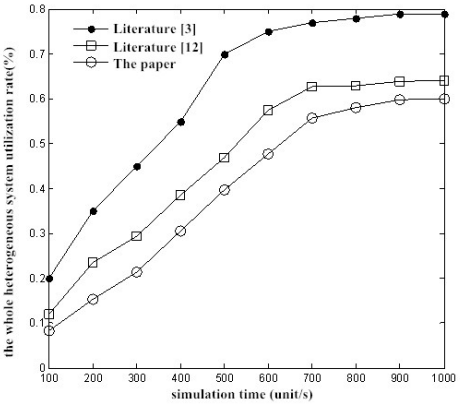

Fig. 5: The performance differences of three schemes in system utilization rate

Figure 5 depicts the performance differences of three schemes in system utilization rate. As can be seen from Figure

5: With the passage of simulation time, system utilization rates of three schemes emerge upward trend; the scheme is rising the fastest in this article, the scheme of literature [12] is followed , and the literature [3] are the slowest . The reason is: the scheme of this paper uses joint optimization strategies of access selection and calls transfer, so that the whole heterogeneous systems keep the lower of new calls access blocking rate and calls switched dropping rate, thereby effectively improving system utilization.

CONCLUSION

In heterogeneous wireless network scenarios, different access networks are very different in terms of technical superiority and operational capacity, etc., the JRRM is an effective mechanism to improve resource utilization for heterogeneous wireless networks. Loads balance control is one of the main tasks of the JRRM, and it has important significance to improve the entire performance across heterogeneous networks.

The paper, on the process of access selection, derives the optimal allocation scheme based on the optimization algorithm, to assign access tasks for the base station cell by using the scheme, so that the loads are more evenly distributed among every base cell. In order to weaken adverse effects of the system loads balance caused by hot cells' sudden business, the paper also uses a calls transfer strategy based on the threshold of load rate. Compared with the schemes of literature [3] and [12], the proposed scheme based on access selection and calls transfer performs better in access blocking rate, calls dropping rate, loads balance and system utilization rate, and has the advantage of good application value.

Acknowledgments
The authors wish to thank the National Natural Science Foundation of China for contract U1204618, the foundation of education department of Henan province of China for contract 13HASTIT041, 12A520055, 2012GGJS-183 and the research foundation of Zhoukou Normal University, under which the present work was possible.

REFERENCES

[1] Paskaleva K. *International Journal of Innovation and Regional Development,* v.1, n.4, pp.405-422, 25 January, **2009**.
[2] Wu G., Mizuno M, Havinga P. *IEEE Communications Magazine*, v.40, n.2, pp. 126-134,**2012**.
[2]NING Guo-qin, ZHU Guang-xi,PENG Lie-xin,et al. Load balancing based on traffic selection in heterogeneous overlapping cellular networks[C]//*The First IEEE and IFIP International Conference in Central Asia on Internet*.pp.26-29,**2012**.
[3] Sun Zhuo, Zheng Xu-fei, Wang Wen-bo. *University Journal of Electronic Technology,* v.39, n.4, pp. 532-536,**2012**.
[4] Shi Wen-xiao , Fan Shao-shuai, Wang Nan et al. *Journal of Communication,* v.31, n.9, pp. 151-156,**2013**.
[5] Li Ming-xin, Chen Shan-zhi, Xie Dong-liang et al. *Journal of Software,* v.21, n.8, pp. 2037-2049,**2013**.
[6] Chen Qian-bin, Zhou Wei-guang, Chai Rong et al. *Journal of Computers,* v.33, n.9, pp.1633-1642,**2013**.
[7] Liu Qi, Yuan Jian, Shan Xiu-ming et al. *Journal of Computers,* v.33, n.9, pp. 1569-1579,**2013**.
[8] Feng Zhi-yong, Zhang Ping, Zhang Yong-jing et al. *Electronics and Information Journal,* v.31, n.4, pp. 893-896,**2012**.
[9]SuKyoung Lee, Kotikalapudi Sriram, Kyungsoo Kim, et al. *IEEE Transactions on Vehicular Technology,* v.58, n.2, pp.865-881,**2012**.
[10]Liu Sheng-mei,Meng Qing-min,Pan, Su, et al..*Journal of Electronics and Information Technology,* v.33, n.1, pp.235-239,**2011**.
[11]Kaveh Shafiee,Alireza Attar,Victor C M. *IEEE Journal on Selected Areas in Communications,* v.29, n.3, pp.534-544,**2013**.
[12]Liu Xing-wei,Fang Xu-ming,Chen Xu, et al. *Journal of Network and Computer Applications,* v.34, n.4, pp.1263-1271,**2013**.
[13] Bettstetter C, Resta G, Santi P. *IEEE Trans on Mobile Computing,* v.2, n.3, pp. 257-269,**2012**.