# A Semi-Distributed Load Balancing Architecture and Algorithm for Heterogeneous Wireless Networks

Md. Golam Rabiul Alam, Choong Seon Hong*

Kyung Hee University, Korea

robi@networking.khu.ac.kr, cshong@khu.ac.kr

## Abstract

At present, dissimilar wireless access networks including WLAN, UMTS and WiMAX coexist in the mobile computing environment. Resource management, load balancing, reliability and QoS are the most significant research issues on heterogeneous wireless networks. Considering the high speed features of third and fourth generation heterogeneous wireless networks, an efficient semi-distributed load balancing architecture and a soft load balancing algorithm is proposed in this paper. Finally, a simulation result shows the improved load balancing rate and lower call blocking probability.

## 1. Introduction

Load balancing is one of the key technologies in the convergence of heterogeneous wireless networks. It is a way to achieve the resource sharing over heterogeneous wireless networks, and it improves resource utilization, enlarge system capacity as well as provide better services for users [1]. Load balancing depends on network architecture and load balancing algorithm. In this paper, firstly we provide efficient network architecture so that entire system load can be properly balanced. Network architecture can be classified as centralize, distributed and semi centralize and semi distributed [2][3]. There are some problems in the first two mechanisms: the centralized one is relatively low reliability, while the distributed one has a huge overhead [4]. To overcome the limitations of network architecture, we have used a Semi-Distributed Architecture (SDA) [5] for heterogeneous network that efficiently reduce signal overhead, improves reliability [6].

Load balancing algorithm is also required to balance loads efficiently among the different access technologies of heterogeneous wireless environment. Traditional hard load balancing should not be the solution of the high speed features of the future wireless network traffic. So, soft load balancing will be the solution, which can take full advantage of resources when dealing with high speed traffic over the heterogeneous wireless networks (HWN).

In this paper, we proposed a load balancing algorithm based on dividing IP flow for speedy traffic over the HWN. Considering multi-path and path loss in multiple networks, optimal flow dividing ratio is obtained through numerical analysis. Based on the load balancing algorithm, radio access selection and vertical handovers is realized for parallel transmission, which increases resource utilization and reduces call blocking probability

The rest of the paper is organized as follows. Section 2 reviews some of the related work found in the literature. The load balancing architecture is elaborately depicted on section 3. The protocol description and a soft load balancing algorithm are presented in section 4. In section 5 a simulation result is illustrated. Finally we conclude the paper in section 6 with discussion.

## 2. Literature Review

Lots of research had done in Load balancing specially in wired homogeneous network. But Load balancing in wireless network is the present research issue as wireless communications are expanded day by day and multiple wireless networks works on the same grid.

In the paper [7], authors introduced hierarchical semi centralized architecture for heterogeneous wireless network. In this architecture, it creates a huge message passing overhead and partially centralized nature may have the probability of central point of failure.

Reference [8] proposed an adaptive soft handover schema for multimedia cellular communication system. Depending upon the loads of the neighboring cells, the soft handover thresholds and time hysteresis are adapted by using fuzzy principals.

The concept of soft load balancing was introduced by the authors of paper [9] and [10]. In those papers they suggested to divide the IP traffic of a user into sub flows of different access networks. They have shown that the utilization of resources of heterogeneous wireless network is increased significantly. But both of the paper simulated the optimal load balancing ratio either in certain environment or in ideal channel condition. So, the method may not work well in real situation i.e the method have the deficiency in generalization.

Reference [11] designed a load balancing algorithm based on dividing IP flow for high speed traffic (LBDH), where they use the hierarchical semi-centralized architecture as their network structure. Though their algorithm performs well in traffic blocking probability, but their algorithm has the problem of definiteness. When an MS reaches at the boundary line of basic grids, it is not clear in the algorithm that who will (RS or IS) take the responsibility of balancing loads using dividing IP flow. So, there lies the possibility of frequent handover for balancing loads among different access networks. As they use hierarchical semi-centralized network structure still it is vulnerable for total system failure.

## 3. Details of the Semi-distributed Architecture

A distributed architecture for wireless network is appealing due to the single point failure of centralized design. The architecture is partially/semi distributed in nature which is free from the huge overhead and delay problem of fully distributed design. In the architecture, a number of adjacent hexagonal cells create a basic grid. Since the system is heterogeneous, each cell may consist of different access network. For simplicity three types of access networks (UMTS, WLAN and WiMax) are shown in Fig 1.

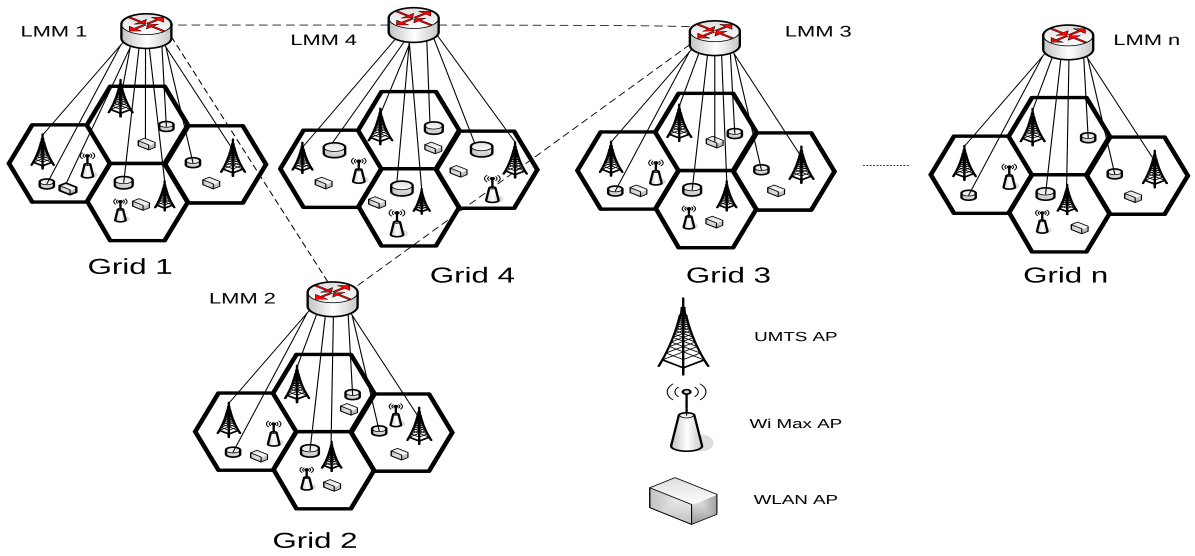


Figure 1. Semi-distributed architecture

In each grid a load and mobile agent management unit (LMMU) is introduced. LMMU consist three units; one unit is responsible for maintaining resource inventory for each type of available access network in each cell, which is called Resource Inventory (RI). Another unit mobile agent (MA) [12] in each cell is accountable for balancing load among various access networks within each cell.

The third unit of LMMU is load and mobile agent manager (LMM), which is responsible for balancing loads among adjacent cells in each grid and among adjacent grids with the help of mobile agents. To ensure reliability, each LMM is associated with its two neighboring LMM so that if any of the LMM is goes down or broken then the backup LMM can take the responsibility for the failed one. In this proposed scheme if one LMM is goes down then only that network management is maintained by backup LMM and the whole system is not hampered.

## 4. Protocol Description to Support Soft Load Balancing

To manage control signal from different access networks, a general link layer is employed in LMMs to support a soft load balancing algorithm shown in Fig. 2. After the common link layer, two individual medium access control (MAC) and physical layers are associated with each access network. RI is located in the MAC layers [9]. MAC protocol for different access network continuously informs its channel status to the upper layer that is in common link layer. Based on the information on attained channel status and load condition, link layer then calculates the flow-dividing ratio. If the mobile subscriber not resides in the boundary line of two or more LMMs, after receiving packets from the MAC layer, the common link layer orders incoming packets on the basis of sequence number and then forwards them to either the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) layer through the IP layer. On the other hand if the mobile subscriber resides in the boundary line of two or more LMMs, then mobile agent communicates with the corresponding LMMs, after receiving packets from the MAC layer, the mobile agent rearranges incoming packets in order of packet number and then forwards them to either the TCP or UDP layer through the IP layer. In the case of heterogeneous network, it is necessary to run the MAC and the physical protocols together, advanced processing speed is required for hardware. So, we introduce a mobile agent to minimize the loads of hardware by rearranging packets of common link layers by incorporating mobile agent.

## A. Protocol Description

In our wireless heterogeneous network we consider that there are *M* (for simulation purpose *M=3*) kinds of various wireless access networks which are fully overlayed. And for simplicity we assume that the transmit power of every AP in one network is fixed and they share the whole available frequency band equally. Here we consider multi-path and path loss and use Rayleigh fading channel model.

The channel gain, transfer characteristics of the physical medium [13] of MS *i* located at the $j^{th}$ cell of system *m* is:

$$g_{i,m,j} = \alpha^2 . A. d_{i,m,j}^{-\beta} \qquad (1)$$

Where, $\alpha$ represents an exponential distributed random variable with mean of 1; $A.d^{-\beta}$ defines path loss, *A* is a constant determined by antenna height and carrier frequency, $\beta$ is a damped exponential, $d_{i,m,j}$ is the distance between MS *i* and AP *j* in system *m*.

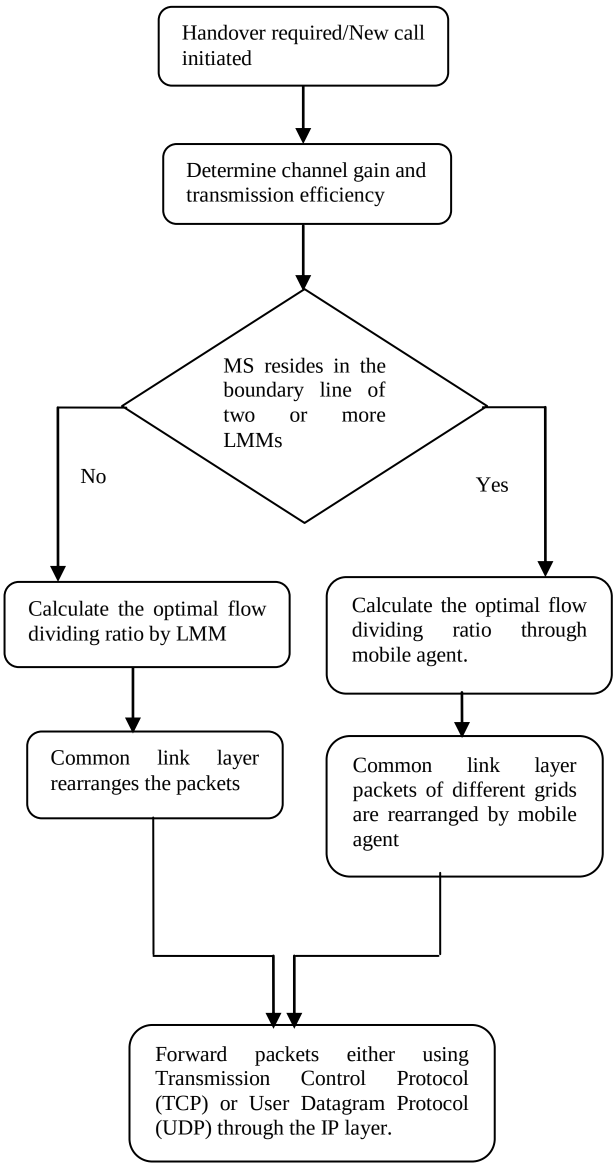


Figure 2. Flow diagram of proposed load balancing algorithm

We define the transmission efficiency $C_{0,i,m,j}$ (bps/Hz) according to Shannon Informatics as the maximum transmission rate got by MS *i* located in the $j^{th}$ cell of system *m* (considering the interference from other APs in the same system),

$$C_{0,i,m,j} = \frac{1}{2}\log_2 . \left( 1 + \frac{1}{\Gamma} . \frac{g_{i,m,j} . P_{0,m}}{\sum_{k=1}^{N_m} g_{i,m,k} . P_{0,m} + N_{0,i,m}} \right) \qquad (2)$$

Where, $N_{0,i,m}$ is the noise power MS *i* received; $\Gamma$ is a constant signal-to-noise ratio gap [13] (SNR gap) related with the required bit error rate and the transmit power of system *m* allocated in unit bandwidth is $P_{0,m}$ .

We take three wireless access systems (such as UMTS, WiMax and WLAN) to calculate the optimal flow dividing ratio.

After MS *i* is accepted, the difference of the relative load among the wireless access systems

$$\Delta_{i,1,2} = \frac{L_{1,j_1,i} + W_{req,i,1}}{W_{T,1}} - \frac{L_{2,j_2,i} + W_{req,i,2}}{W_{T,2}} \qquad (3)$$

$$\Delta_{i,2,3} = \frac{L_{2,j_2,i} + W_{req,i,2}}{W_{T,2}} - \frac{L_{3,j_3,i} + W_{req,i,3}}{W_{T,3}} \qquad (4)$$

$$\Delta_{i,3,1} = \frac{L_{3,j_3,i} + W_{req,i,3}}{W_{T,3}} - \frac{L_{1,j_1,i} + W_{req,i,1}}{W_{T,1}} \qquad (5)$$

Where, $W_{req,i,m}$ is bandwidth requirement of MS *i* in the system *m* and $L_{m,j,i}$ is the load of the $j^{th}$ cell in system *m*.

To gain the best load balancing effect, the optimal flow dividing ratio, $\omega^*_{i,m}$ of user *i* in system *m* is derived by

$$\left[\omega^*_{i,1}, \omega^*_{i,2}, \omega^*_{i,3}\right] = \arg\min \left|\Delta_{i,x,y}\right| \qquad (6)$$

Where $x=\{1,2,3\}$ and $y=\{1,2,3\}$ but $x \neq y$.

Let $|\Delta| = 0$, on the basis of (3) to (6), the optimal flow dividing ratio can be obtained by the following three equations

$$\omega_{i,1} = \frac{R_i.C_{0,i,1,j_1}.W_{T,1} + C_{0,i,1,j_1}.C_{0,i,2,j2}.C_{0,i,3,j_3}.(L_{2.j_2.i}.W_{T,2} + L_{3.j_3.i}.W_{T,3} - L_{1.j_1.i}.W_{T,1})}{R_i.(C_{0,i,1,j_1}.W_{T,1} + C_{0,i,2,j2}.W_{T,2} + C_{0,i,3,j_3}.W_{T,3})} \qquad (7)$$

$$\omega_{i,2} = \frac{R_i.C_{0,i,2,j_2}.W_{T,2} + C_{0,i,1,j_1}.C_{0,i,2,j2}.C_{0,i,3,j_3}.(L_{1.j_1.i}.W_{T,1} + L_{3.j_3.i}.W_{T,3} - L_{2.j_2.i}.W_{T,2})}{R_i.(C_{0,i,1,j_1}.W_{T,1} + C_{0,i,2,j2}.W_{T,2} + C_{0,i,3,j_3}.W_{T,3})} \qquad (8)$$

$$\omega_{i,3} = \frac{R_i.C_{0,i,3,j_3}.W_{T,3} + C_{0,i,1,j_1}.C_{0,i,2,j2}.C_{0,i,3,j_3}.(L_{2.j_2.i}.W_{T,2} + L_{1.j_1.i}.W_{T,1} - L_{3.j_3.i}.W_{T,3})}{R_i.(C_{0,i,1,j_1}.W_{T,1} + C_{0,i,2,j2}.W_{T,2} + C_{0,i,3,j_3}.W_{T,3})} \qquad (9)$$

Where, $R_i$ be the required rate of MS *i*.

### B. Load Balancing Algorithm

Step 1: When MS $i$ requires to access to the system or needs to handover intra-system or inter-system during to the factors such as the movement of MS and the decline of service quality, RI lists the available APs.

Step 2: RI measures the channel gains $g_{i,m,j}$ between MS $i$ and each available AP according to (1), calculates the transmission efficiency $C_{0,i,m,j}$ of user $i$ in system $m$ on the basis of (2), and delivers the results to LMM.

Step 3: If MS $i$ reside within the basic grid then LMM calculates the optimal flow-dividing ratio for MS $i$ according to (7), (8) and (9).

Step 4: If MS $i$ reached at the border of basic grid then mobile agent communicates with neighboring LMMs and calculates the optimal flow-dividing ratio for the MS $i$ according to (7),(8) and (9).

Step 5: LMM makes MS $i$ access to or handover to the relevant network(s) according to the optimal flow-dividing ratio.

## 5. Simulation Study

The call blocking probability and load balancing uniformity is simulated for study and comparison purpose. MATLAB simulator is used to simulate the proposed SDA based load balancing algorithm and exiting HSCA [7] based LBDH [11], the latest flourishing architecture of heterogeneous wireless network over semi-centralized and semi-distributed architecture SCSDA [3].

### A. Simulation Scenario

For simulation purpose and simplicity we consider that all three types of APs (UMTS, WiMax and WLAN) are present in each basic grid. The APs of center cell of the three access technologies are overlayed and the subscribers are randomly distributed in the overlapped area, where as the subscriber locations are updated in every 100ms. The simulation parameters are summarized in Table 1.

Table 1. Simulation parameters

| Parameters | Cellular System 1 | Cellular System 2 | Cellular System 3 |
|---|---|---|---|
| Radius | 0.38km | 0.75km | 1.5km |
| Carrier frequency | 2.0GHz | 2.0GHz | 3.5GHz |
| Bandwidth | 5MHz | 10MHz | 20MHz |
| TX power of AP | 24dBm | 43 dBm | 46 dBm |
| Noise Power | -118 dBm/Hz | -170 dbm/Hz | -174 dbm/Hz |
| Required BER | 10-3 | | |
| Subscriber speed | 120km/h | | |
| Traffic rate | Follows uniform distribution from 500 kb/s to 1000 kb/s | | |
| Traffic generation time interval | Follows poison distribution with mean of 1s | | |

### B. Simulation Result

A simulation result regarding call blocking probability is shown in Fig. 3. Here, number of new arrivals represents the number of activated subscriber in the cell of cellular system 1. The call blocking probability is increasing with the number of subscribers' increment. But the call blocking probability of the proposed algorithm is always lower than the blocking probability of existing benchmark i.e hierarchical semi-centralized architecture (HSCA) based algorithm (LBDH). Because, in existing architecture, when subscriber reached at the border line of the basic grid, load balancing algorithm based on HSCA tries to make a handover and otherwise drop the packet. On the other hand the proposed load balancing algorithm based on SDA try to balance loads using mobile agents and by determining flow dividing ratio among the available nodes of adjacent grids.

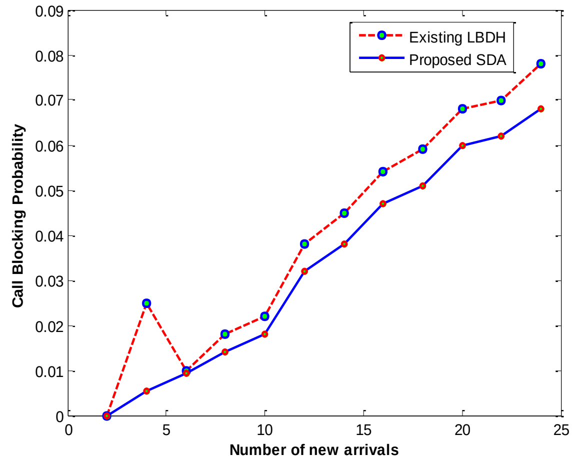


Figure 3. Call blocking probability

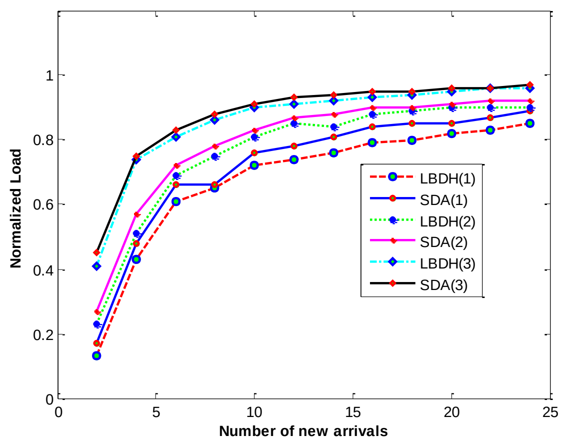


Figure 4. Balanced loads of different access technologies

The load status of different access technologies is shown in Fig. 4. Here, the normalized loads of the different cellular systems are becomes more balanced with increment number of users (number of new arrivals). The figure shows that SDA based load balancing algorithm distributes loads more uniformly than the existing LBDH. Because the proposed algorithm always try to balance loads among the available access technologies, whereas existing approach concentrate on load sharing.

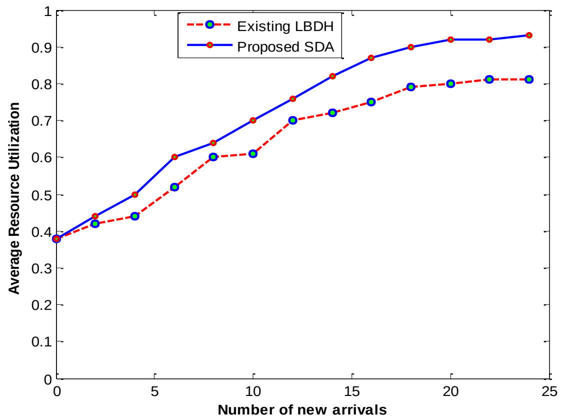


Figure 5. Overall heterogeneous system utilization

Fig. 5 presents the average resource utilization of heterogeneous system of existing LBDH and proposed SDA based load balancing algorithm. The overall resource utilization of heterogeneous system using the proposed algorithm is higher than that of the LBDH algorithm. The major cause of the higher resource utilization is its balance distribution of loads among different wireless network access technologies, so it can be served more number of arrival calls.

## 6. Conclusion

The proposed load balancing algorithm deployed on the SDA architecture reduces call blocking probability significantly. The simulation result also articulates its uniform distribution of loads among different access technologies. Obviously the proposed load balancing architecture and algorithm introduce an efficient load balancing method in heterogeneous wireless networks. Resource utilization, uplink-downlink power consumption, handoff requirements, interference and QoS parameters are still subject to study in the proposed method.

## Acknowledgement

This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2011-0020518). *Dr. CS Hong is corresponding Author.

### References


[1] B. Li, W. X. Shi and N. Li, "A Hierarchical semi-centralized architecture for load balancing of heterogeneous wireless networks," *Second International Conference on Networks Security Wireless Communications and Trusted Computing* (NSWCTC), vol.2, Apr. 2010, pp.28–31.

[2] Y. R. Lan and T. Yu, "A dynamic central scheduler load balancing mechanism," IEEE *Fourteenth Annual International Phoenix Conference on Computers and Communications*, Mar. 1995, pp.734–740.

[3] G. Q. Ning and G. X. Zhu, "Load balancing based on traffic selection in heterogeneous overlapping cellular networks," *Journal of Chinese Computer Systems*, vol.27, no.11, Nov. 2006, pp.2036–2041.

[4] G. Q. Ning, G. X. Zhu, L. X. Peng and X. F. Lu, "Load balancing based on traffic selection in heterogeneous overlapping cellular networks," *The First IEEE and IFIP International Conference in Central Asia on Internet*, Sept. 2005.

[5] M. G. R. Alam, C. Biswas and N. Nower "A Reliable and Semi-distributed Load Balancing Architecture of Heterogeneous Wireless Networks", *Proceedings of 14th International Conference on Computer and Information Technology (ICCIT 2011)*, 22-24 December, 2011, Dhaka, Bangladesh.

[6] M. G. R. Alam, C. Biswas, N. Nower and M. S. A. Khan "A Reliable and Semi-distributed Load Balancing Architecture of Heterogeneous Wireless Networks", *International Journal of Computer Networks & Communications (IJCNC)*, Vol. 4, No.1, January 2012, pp. 97-111.

[7] W. Shi, B. Li, N. Li and C. Xia, "A Network architecture for load balancing of heterogeneous wireless network", *Journal of Networks*, Volume 6, No 4, Apr. 2011, pp. 623-630.

[8] K. Yang, B. Qiu and L. S. Dooley, "An Adaptive Soft Handover Schema using Fuzzxy Load-balancing for WCDMA Systems," *IASTED International Conference on Networks and Communication Systems, Thailand*, March 2006.

[9] H. Son, S. Lee, S. C. Kim, Y. S. Shin, "Soft load balancing over heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2632 – 2638, July 2008.

[10] L. Kuheli, J. J. Hyun, D. P. Agrawal , "Exploring load balancing in heterogeneous networks by rate distribution," *IEEE International Performance, Computing and Communications Conference*, 2008. IPCCC 2008, pp. 427 – 432, December 2008.

[11] B. Li, W. X. Shi and Y. Zhao, "A Load Balancing Algorithm Based on Dividing IP Flow for High-speed Traffic over Heterogeneous Wireless Networks," *3rd International Congress on Image and Signal Processing (CISP)*, 2010, pp.4294–4298.

[12] W. X. and Z. Q. Qiang, "The analysis of the application about MA in E-Commerce," *International Conference on E-Business and E-Governance*, 2010, pp.380-383.

[13] J. Huiling, "Access selection and call admission control in heterogeneous wireless networks," Ph.D Thesis, Zhejiang University, April 2000.