A Network Architecture for Load Balancing of Heterogeneous Wireless Networks

Wenxiao Shi, Bin Li, Na Li, Chuanjun Xia
Jilin University, Changchun, China
Email: swx@jlu.edu.cn

***Abstract*—The traditional centralized load balancing had a relatively low reliability, and the distributed load balancing had a huge overhead. To solve these problems, this paper mapped heterogeneous wireless networks to distributed grids by introducing Resource Management Unit, and then presented a hierarchical semi-centralized architecture for load balancing of heterogeneous wireless networks drawing on the idea of grid in computer networks. The analytical models for the integrated reliability and signaling overhead of the architecture were established. Theoretical analysis and simulation results indicate that the architecture can reduce the signaling overhead and improve the system reliability effectively.**

***Index Terms*—network architecture, signaling overhead, integrated reliability, load balancing, heterogeneous wireless networks**

I. INTRODUCTION

Radio systems are moving toward forming heterogeneous wireless networks: collaborations of multiple radio access networks, which in some cases operate different radio access technologies [1]. The deployment of heterogeneous wireless networks is spreading throughout the world as users want to be connected anytime, anywhere, and anyhow [2]. 3GPP specifies some recommendations for methods, architectures, and design of heterogeneous wireless network interconnection (in particular, the one formed by UMTS and WLAN networks), such as those presented in [3]. The key topics in heterogeneous wireless networks are referred to as spectrum sensing, coexistence, resource management, reliability and QoS support avoiding interference etc. [4].

As a recent research focus, load balancing [5, 6] which belongs to Radio Resource Management (RRM) is one of the key technologies in the convergence of heterogeneous wireless networks. Load balancing is a significant method to achieve the resource sharing over heterogeneous wireless networks, and it can improve resource utilization, enlarge system capacity, as well as provide better services for users.

Generally, load balancing is divided into two parts [7]: network architecture and load balancing algorithm. The former is the foundation of load balancing, and a good network architecture can improve the efficiency of load balancing. In the perspective of control mode, load balancing mechanisms can be classified as centralized, distributed and semi-centralized and semi-distributed [8, 9]. There are some problems in the first two mechanisms: the centralized one has a relatively low reliability, while the distributed one has a huge overhead [10].

A grid [11, 12] is a service for sharing computer power and data storage capacity over the Internet. Grid infrastructure is a large virtual organization that integrates a large mount of distributed heterogeneous resources and high performance computing capabilities into a super service, which can provide huge computing services, storage capability and so on. Grid allows the use of geographically distributed computing systems that belong to multiple organizations as a single system. Resource management and scheduling is the important components of the Grid. It efficiently maps jobs submitted by the user to available resources in grid environment [13].

There are some similarities between grid and heterogeneous wireless networks, such as the dynamic variation of resources, heterogeneous structure, and the key technology of integrating and deploying the distributed resources. Enlightened by the similarities, we propose a hierarchical semi-centralized architecture (HSCA) based on basic grids for load balancing of heterogeneous wireless networks. Since the reliability and overhead are important performances to a network, we analyze the two performances of HSCA we proposed in this paper.

The remainder of this paper is organized as follows. In Section II we briefly discuss the related researches into overhead and reliability of network architectures. Section III introduces the HSCA we proposed, and gives the signaling flows of the HSCA. In Section IV, the modeling of integrated reliability and signaling overhead for HSCA is presented followed in Section V by our simulation and the results. Finally in Section VI we conclude this paper with discussion of our work.

II. RELATED WORK

The authors of [14] developed a mathematical framework that can be used to compactly represent and analyze heterogeneous networks that combine multiple entity and link types. They generalized Bonacich centrality, which measures connectivity between nodes

This work was supported by the National Natural Science Foundation of China (No. 60972028).

doi:10.4304/jnw.6.4.623-630

by the number of paths between them, to heterogeneous networks and used this measure to study network structure. The authors of [9] proposed a semi-centralized and semi-distributed architecture (SCSDA), in which a BS just exchanges load information with several neighboring BSs. Although the architecture can reduce the overhead of control signaling, the authors neither expressed the overhead in mathematical formula, nor proved it by simulation.

Reference [15] designed a hybrid wireless network architecture, [16] proposed a multiple mobile routers based network architecture to support seamless mobility across heterogeneous networks, and they both tested the overhead by NS2 simulator. However, the model of the overhead was not derived in neither [15] nor [16]. Route overhead was analyzed in theory in [17], by calculating the number of control messages generated in a BS/AP service area due to maintaining route. Nevertheless, the simulation for overhead was not given. The communication overhead of the scheme presented in [18] was calculated, and an algorithm for minimizing the communication overhead was given, which was proved to be effective through simulation. The authors of [19] considered a general heterogeneous network architecture with two basic entities in the system: mobile nodes (MNs) and access points (APs). They formulated the overhead of AP discovery which is divided into hello messages and RREQ messages, and gave the simulation results. Reference [20] proposed a hierarchical and distributed (HD) architecture with three hierarchical levels of mobility management being distinguished: end terminal remains connected to the same radio access network but it changes its point of attachments, end terminal changes its radio access network but it remains associated to the same operator and end terminal changes its operator network. And it also studied signaling cost generated by QoS negotiation during handover process in both theory and simulation.

The research on reliability of telecommunication network starts at the study on switched telecommunication network by Lee [21]. Lee defined call blocking as the link failure, and measured reliability taking connectivity [22] as standard. Reference [23] mentioned the concept of integrated reliability, which took call loss as the evaluation indicator of network reliability, and proved that the integrated reliability can reflect the practical situation much better than taking connectivity as standard. The authors of [24] analyzed the reliability aspects of some access network topologies to insure a certain level of quality of service at the lowest cost for the end users, which are a little alike to our works.

III. Hierarchical Semi-centralized Architecture

A. *Description of the architecture*

The hierarchical semi-centralized architecture based on basic grids is depicted in Fig. 1, which takes three different types of access networks (UMTS, WLAN and WiMax) for example. A basic grid is made up of several adjacent cells. IS (Information Server), RA (Resource Allotter) and RS (Resource Statistics) are collectively referred to as Resource Management Unit (RMU), which are responsible for managing the resources of basic grids. Installed in the access point (AP), a RS is used to calculate resources of its jurisdiction cell. A RA collects load information from RSs, and balances the load in virtue of the load and resources of the basic grid. Normally an IS allocates resources for the borders of basic grids and stores information of cell identification, location and load states. However, it can take over the broken RA immediately. In order to improve the system reliability, a main IS and a standby IS are set up, additionally, a RA and IS are connected by two junction lines. Once the main IS (one junction line between RA and IS) stops running, the standby IS (the other junction line) will take over it.

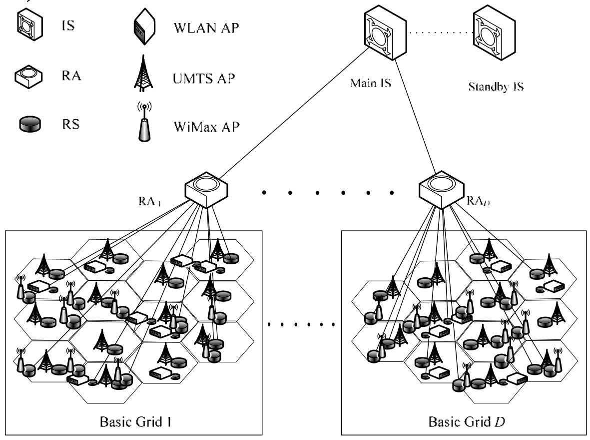


Figure 1. Hierarchical semi-centralized architecture.

A RA and the RSs located in the basic grid administered by the RA compose the first layer centralized architecture, while an IS (including a main IS and a standby IS) and the RAs constitute the second layer centralized architecture. The whole first layer is made up of a series of basic grids and the corresponding RAs. Thus, we call it the hierarchical semi-centralized architecture.

The signaling overhead of transferring load information between RSs and RAs (we call it the RSs to RAs signaling overhead for short, and the similar terms are used below) can be reduced by limiting the number of RSs in a basic grid; the RAs to RS signaling overhead is also small because of the limited number of RAs. The IS can take over the RA immediately if the latter breaks down, which ensures the communication of the basic grid; the standby IS and the two junction lines between each RA and IS make further improvement on system reliability. Therefore, HSCA has the advantages of low signaling overhead and high reliability, which are the advantages of centralized architecture and distributed one respectively.

There are two ways to fix the RA and IS. One way is to install them in the existing equipments: RS in RNC (Radio Network Controller) and RA in GGSN (Gateway GPRS Support Node), the other way is to set them separately. The advantage of the former is that it can reduce the housing construction and maintenance cost, while the disadvantage is that it is restrained by the existing network topology. The latter can design the network topology flexibley while increases the cost. Weighing the strong points and weaknesses of the two methods, a proper method to fix the RA and IS can be chosen.

B. Signaling flow

Normally a RS calculates load and resources of its jurisdiction cell according to the wireless parameters received from the AP, and then transfers the load, resources and location information to the RA. On the basis of load balancing algorithm (is not in the scope of this paper), the RA transfers load balancing information to RS. Then the RS changes load balancing information into load balancing instructions and transfers the load balancing instructions to the AP. The signaling flow chart is shown in Fig. 2.

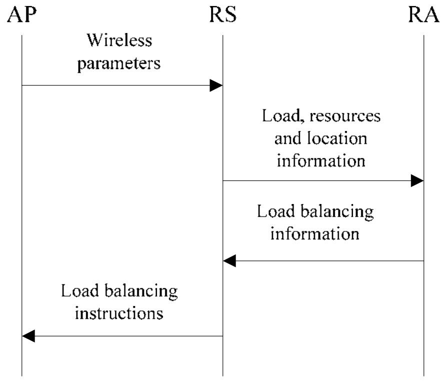


Figure 2. Signaling flow chart 1.

When a subscriber is on the border of basic grids, the IS will allocate resources for the subscriber, and the signaling flow is shown in Fig. 3.

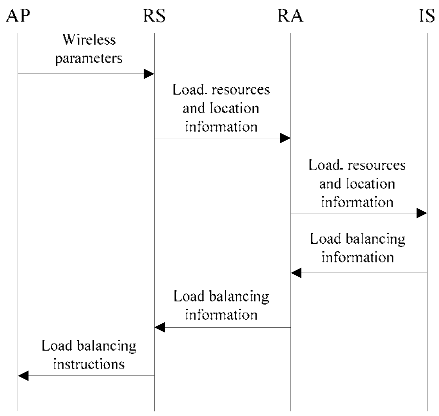


Figure 3. Signaling flow chart 2.

When a RA breaks down, the IS can take over it immediately, and the signaling flow is shown in Fig. 4.

IV. MODELING OF RELIABILITY AND SIGNALING OVERHEAD

In this section, we derive an analytical model for reliability of HSCA and two analytical models for signaling overhead of HSCA. The following notations are

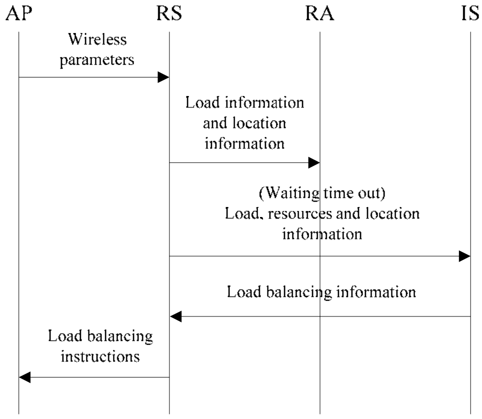


Figure 4. Signaling flow chart 3.

used in our analysis.

i: System type. i=1 denotes UMTS system; i=2 denotes WLAN system; i=3 denotes WiMax system.

b_{ij}: The traffic intensity between RS_i and RA_j.

c_j: The traffic intensity between RA_j and IS.

R_{IS}: The reliability of IS.

R_{RA}: The reliability of RA.

R_c: The reliability of junction line between IS and RA.

a_i: The signaling overhead of transferring load information once between one RS located in system type i and RA.

d: The signaling overhead of transferring load information once between one RA and IS.

e: The signaling overhead of transferring load information once between one main IS and standby IS.

A_i: The number of APs in system type i.

D: The number of RAs.

A_{ij}: The number of APs for system type i in the basic grid j.

λ_i: The traffic arrival rate of system type i.

μ_i: The service rate of system type i.

m_i: The cell capacity of system type i.

k_{i1}: The light threshold of system type i.

k_{i2}: The heavy threshold of system type i.

T: The period of transferring load information among RMU.

To facilitate the analysis, we assume that there are only one main IS and one standby IS.

A. Modeling of integrated reliability

Since the integrated reliability can reflect the practical situation very well [23], we use it to analyze the reliability of HSCA.

The HSCA we presented is actually a tree structure, as shown in Fig. 5.

When the number of RS is large, the breakdown of a RS or the junction line between a RS and the RA has little influence on the total traffic of the system. Thus, the breakdown of a RS or the junction line between a RS and the RA can be neglected.

Taking the main IS and the standby IS as a whole, the probability that the IS is in normal use is

$$p_{IS} = 1 - \left(1 - R_{IS}\right)^2 . \quad (1)$$

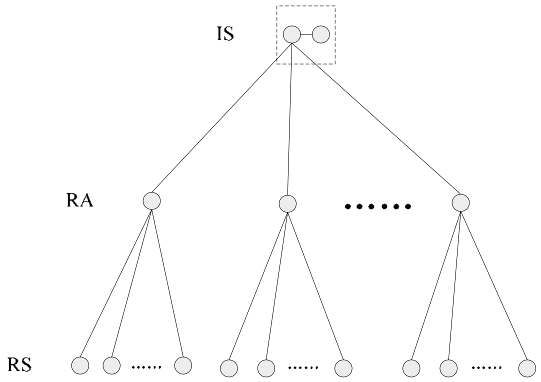


Figure 5. The tree structure of HSCA.

Taking the two junction lines between RA_j (j=1,2,⋯,D) and the IS as a whole, the probability that the junction line is out of fault is

$$p_c = 1-(1-R_c)^2 . \tag{2}$$

The probability that the whole system is in normal use is

$$p_0 = p_{IS} \cdot p_c^D \cdot R_{RA}^D \tag{3}$$

then, the traffic of the whole system does not lose, namely

$$L_0 = 0 . \tag{4}$$

The probability that K_1 (K_1=1,2,⋯,D) junction lines between IS and RA are out of action is

$$p_1 = p_{IS} \cdot \frac{D!}{K_1!\,(D-K_1)!} \cdot (1-p_c)^{K_1} \cdot p_c^{D-K_1} \cdot R_{RA}^D \tag{5}$$

then, the traffic on the K_1 junction lines loses, while all the traffic between RAs and RSs does not lose. Thus, the loss of the whole system traffic is

$$L_1 = \sum_{j=1}^{K} c_j . \tag{6}$$

The probability that K_2 (K_2=1,2,⋯,D) RAs are out of action is

$$p_2 = p_{IS} \cdot p_c^D \cdot \frac{D!}{K_2!\,(D-K_2)!} \cdot (1-R_{RA})^{K_2} \cdot R_{RA}^{D-K_2} \tag{7}$$

then, the IS can take over the K_2 breakdown RAs. Thus, the traffic of the whole system does not lose, namely

$$L_2 = 0 . \tag{8}$$

The probability that the IS and K_3 (K_3=1,2,⋯,D) RAs are out of action at the same time is

$$p_3 = (1-p_{IS}) \cdot p_c^D \cdot \frac{D!}{K_3!\,(D-K_3)!} \cdot (1-R_{RA})^{K_3} \cdot R_{RA}^{D-K_3} \tag{9}$$

then, the traffic between all the RAs and the IS loses whether the junction lines between the RAs and the IS are out of action or not, and the traffic between the K_3 broken RAs and the IS loses. Thus, the loss of the whole system is

$$L_3 = \sum_{j=1}^{D} c_j + \sum_{j=1}^{K}\sum_{i=1}^{3} b_{ij} . \tag{10}$$

The total traffic intensity of the system is

$$B = \sum_{j=1}^{D} c_j + \sum_{j=1}^{D}\sum_{i=1}^{3} b_{ij} . \tag{11}$$

Therefore, the integrated reliability of the whole system is

$$R = 1-\frac{1}{B}\cdot\left[p_0 \cdot L_0 + p_1 \cdot L_1 + p_2 \cdot L_2 + p_3 \cdot L_3\right]. \tag{12}$$

B. Modeling of signaling overhead

In the perspective of executing time, load balancing mechanisms can be classified as periodic and non-periodic [8]. Accordingly, the mode of transferring load information can be divided into periodic and non-periodic. Following we study the periodic and non-periodic signaling overhead respectively.

1) The periodic signaling overhead

The RSs to RAs signaling overhead in unit time is:

$$O_{11} = \frac{1}{T}\cdot\sum_{i=1}^{3}(a_i \cdot A_i) . \tag{13}$$

The RAs to IS signaling overhead in unit time is:

$$O_{12} = \frac{1}{T}\cdot d \cdot D . \tag{14}$$

The main IS to standby IS signaling overhead in unit time is:

$$O_{13} = \frac{1}{T}\cdot e . \tag{15}$$

The signaling overhead of transferring load information among RMU in unit time is:

$$O_1 = O_{11} + O_{12} + O_{13} = \frac{1}{T}\cdot\left[\sum_{i=1}^{3}(a_i \cdot A_i) + d \cdot D + e\right]. \tag{16}$$

For problem tractability, we assume that a_1=a_2=a_3, then (16) can be reduced to:

$$O_1 = \frac{1}{T}\cdot\left(a_1 \cdot \sum_{i=1}^{3} A_i + d \cdot D + e\right). \tag{17}$$

2) The non-periodic signaling overhead

Enlightened by the ideas from [9] which employed two thresholds to classify cells into three classes, we introduce a heavy threshold and a light threshold to classify cells into three classes in terms of their load states: overloaded, under-loaded and balanced cells. A cell is in the under-loaded state when its load ≤ k_{i1}, in the overloaded state when its load ≥ k_{i2}; in the balanced state when k_{i1} < its load < k_{i2}.

RMU transfer load information when the following conditions are satisfied: RS transfers load information to RA when the load state of its jurisdiction cell changes; RA transfers load information to IS when the load state of one or more cells changes in the basic grid administered by the RA; the main IS transfers load information to standby IS when one or more RAs transfer load information to the main IS.

For problem tractability, we assume that the arrival and service process are Poisson process and each cell is an independent M/M/m(m) queue.

According to queuing theory [25], the state probabilities of system type i are:

$$P_{i0} = \left[\sum_{j=0}^{m_i} \frac{(\lambda_i / \mu_i)^j}{j!} \right]^{-1} \tag{18}$$

$$P_{ik} = \frac{(\lambda_i / \mu_i)^k}{k!} \cdot P_{i0} \,. \tag{19}$$

In time T, the probability that a cell in system type i changes from under-loaded state to balanced state is:

$$\Pr_i(\mathrm{U} \rightarrow \mathrm{B}) = \mathrm{P}_{ik_{i1}-1} \lambda_i T \left[1 - (k_{i1} - 1)\mu_i T\right] . \tag{20}$$

In time T, the probability that a cell in system type i changes from balanced state to overloaded state is:

$$\Pr_i(\mathrm{B} \rightarrow \mathrm{O}) = \mathrm{P}_{ik_{i2}-1} \lambda_i T \left[1 - (k_{i2} - 1)\mu_i T\right] . \tag{21}$$

In time T, the probability that a cell in system type i changes from overloaded state to balanced state is:

$$\Pr_i(\mathrm{O} \rightarrow \mathrm{B}) = \mathrm{P}_{ik_{i2}+1} \left(k_{i2} + 1\right) \mu_i T \left(1 - \lambda_i T\right) . \tag{22}$$

In time T, the probability that a cell in system type i changes from balanced state to under-loaded state is:

$$\Pr_i(\mathrm{B} \rightarrow \mathrm{U}) = \mathrm{P}_{ik_{i1}+1} \left(k_{i1} + 1\right) \mu_i T \left(1 - \lambda_i T\right) . \tag{23}$$

In time T, the probability that the load state of a cell changes in system type i is:

$$\Pr_i = \Pr_i(\mathrm{U} \rightarrow \mathrm{B}) + \Pr_i(\mathrm{B} \rightarrow \mathrm{O}) + \Pr_i(\mathrm{O} \rightarrow \mathrm{B}) + \Pr_i(\mathrm{B} \rightarrow \mathrm{U}) \,. \tag{24}$$

In time T, the probability that RA j transfers load information to IS is:

$$\Pr_j^{'} = 1 - \prod_{i=1}^{3} \left(1 - \Pr_i\right)^{A_{ij}} . \tag{25}$$

In time T, the probability that main IS transfers load information to standby one is:

$$\Pr^{''} = 1 - \prod_{j=1}^{D} \left(1 - \Pr_j^{'}\right) . \tag{26}$$

The signaling overhead of transferring load information among RMU in unit time is:

$$O_2 = \frac{1}{T} \cdot \left[a_1 \cdot \sum_{i=1}^{3} \left(A_i \cdot \Pr_i\right) + d \cdot \sum_{j=1}^{D} \Pr_j^{'} + e \cdot \Pr^{''} \right] . \tag{27}$$

V. Simulation Study

To evaluate the performance of HSCA, we employ a simulation model using the simulator MATLAB, and compare the signaling overhead of HSCA with that of SCSDA.

A. Simulation Scenario

The scenario is a medium urban area, where both UMTS system and WiMax system cover the whole area while WLAN system covers the hot spots only. In order to reduce the complex of simulation, we assume that there are all the three types of APs in each basic grid, and the number of APs for the same system is equal in every basic grid.

The values of parameters used in simulation are as follows. T=0.1s, A_1=600, A_2=900, A_3=600; R_{IS}=0.99, R_{RA}=0.98, R_c=0.97; b_i=1erl; K_1=1, K_2=1, K_3=1; a_1=1, d=1, e=1; m_1=60, m_2=20, m_3=80; η_{iTH1}=0.7, η_{iTH2}=0.9; μ_i=1/180s. Where i=1, 2, 3.

B. Simulation Results

Fig. 6 indicates the integrated reliability of HSCA with different number of RAs. Because of the assumption in *Simulation Scenario*, the number of RAs is discrete, and its value set is {1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 25, 30, 50, 60, 75, 100, 150, 300}. We can see that the integrated reliability of HSCA rises with the number of RAs increasing, and the integrated reliability is always very high. When there is only one RA, the HSCA degenerate to be a centralized architecture with a lower reliability. The larger the number of RAs is, the more distributed the architecture is, and the higher the reliability is.

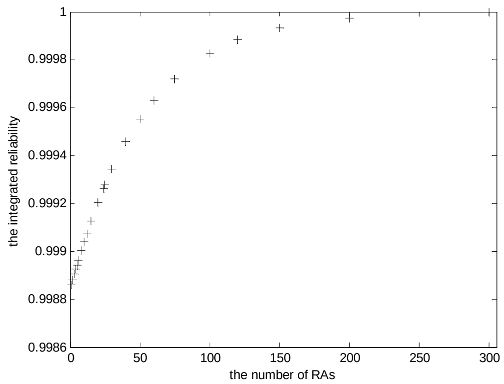


Figure 6. The integrated reliability with different number of RAs.

Fig. 7 and 8 illustrate the signaling overhead of HSCA and SCSDA with different number of RAs. The signaling overhead of SCSDA, which has nothing to do with the number of RAs, is associated only with the number of APs. Therefore, the signaling overhead of SCSDA is constant when the number of RAs varies. It can be observed that the signaling overhead of HSCA is always much smaller than that of SCSDA despite whether transferring load information is periodical or non-periodical, which indicates that HSCA has great advantages in reducing system signaling overhead.

Equation (17) tells us that the periodic signaling overhead includes three parts: the RSs to RAs signaling overhead, the RAs to IS signaling overhead and the main IS to standby IS signaling overhead. Since the number of RSs, main IS and standby IS is constant, the signaling overhead of the first part and the third part is invariable, while the second part is variable. We can see in (5) that the signaling overhead of the second part has a linear relationship with the number of RAs. Therefore, the periodic signaling overhead of HSCA increases linearly with the rising of the number of RAs, which can be seen in Fig. 7.

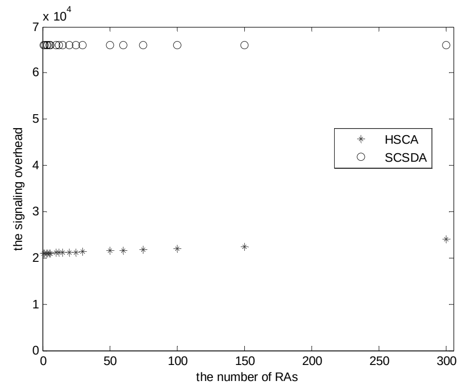


Figure 7. The periodic signaling overhead with different number of RAs

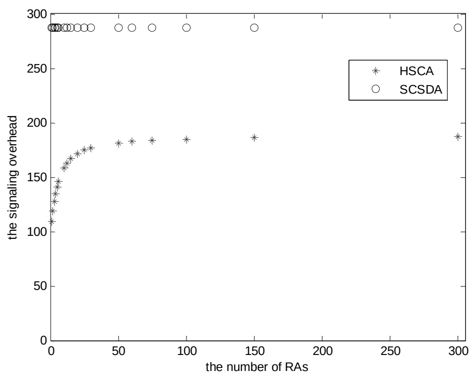


Figure 8. The non-periodic signaling overhead with different number of RAs (λ_1=0.3, λ_2=0.2, λ_3=0.5).

Fig. 8 shows that the non-periodic signaling overhead of HSCA rises rapidly with the number of RAs increasing from 1 to 20, while it rises slowly with the number of RAs increasing from 20 to 300. When there are a few RAs (less than 20), the number of APs in a basic grid is large, and the probability that one or more cells change load states in a basic grid is large, which makes the probability of transferring load information from RA to IS large, as a result, the signaling overhead of HSCA rises rapidly with the number of RAs increasing from 1 to 20; vice versa.

Comparing Fig. 7 with Fig. 8, it can be seen that the non-periodic signaling overhead of the two architectures is much smaller than the periodic one. Therefore, it can save lots of resources to employ the mode of transferring load information non-periodically.

The relationship between the non-periodic signaling overhead of the two architectures and the traffic arrival rate of the three systems are shown in Fig. 9, 10, and 11.

The following conclusions can be drawn from Fig. 9: in UMTS cells, the signaling overhead of the two architectures remains the same with λ_1 increasing from 0 to 0.15, rises gradually with λ_1 increasing from 0.15 to 0.32, drops gradually with λ_1 increasing from 0.32 to 0.65,

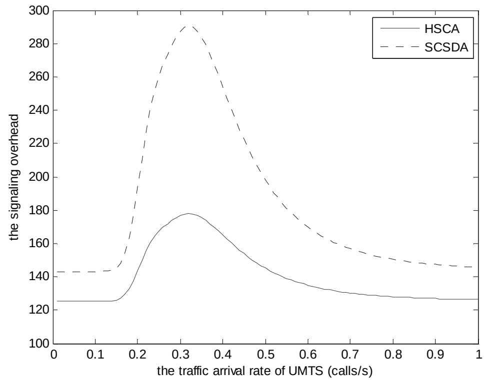


Figure 9. The signaling overhead of the two architectures with different traffic arrival rate of UMTS (λ_2=0.2，λ_3=0.5，D=30）

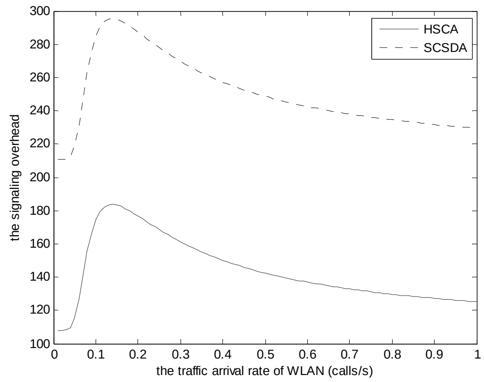


Figure 10. The signaling overhead of the two architectures with different traffic arrival rate of WLAN (λ_1=0.3，λ_3=0.5，D=30）

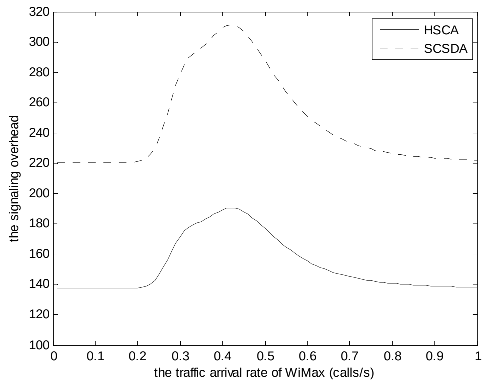


Figure 11. The signaling overhead of the two architectures with different traffic arrival rate of WiMax (λ_1=0.3，λ_2=0.2，D=30）.

and drops slowly with λ_1 increasing from 0.65 to 1. The above phenomenon is due to the following aspects. When λ_1 increases from 0 to 0.15, UMTS cells always have small amount of subscribers and are always in under-loaded state, which causes RSs hardly to transfer load information to RA. When λ_1 increases from 0.15 to 0.32, the probability that the number of subscribers varies in the vicinity of k_{11} and k_{12} increases gradually, and system states change frequently. As a result, the probability that

RSs transfer load information to RA increases gradually. When λ_1 increases from 0.32 to 0.65, the probability that the number of subscribers exceeds k_{12} increases gradually, and the system states hardly change, which makes the probability that RSs located in UMTS APs transfer load information to RA reduce gradually. When λ_1 increases from 0.65 to 1, UMTS cells are almost always in the overloaded state as a result of large number of subscribers, which leads to the probability that RSs transfer load information to RA smaller and smaller.

The tendencies of the curves in Fig. 10 and 11 are similar with that in Fig. 9, and the reasons are also similar. It is unnecessary to give more details.

VI. CONCLUSIONS

In this paper, we proposed a hierarchical semi-centralized architecture based on basic grids for load balancing of heterogeneous wireless networks. An IS is not only a superior server of RAs but also a backup for RAs, which improves the system reliability; the standby IS and the two junction lines between each RA and IS make a further improvement on system reliability. It has been shown in our simulation that HSCA can reduce signaling overhead to a great degree while maintaining a very high reliability. The proposed network architecture has properly solved the problem of low reliability in centralized load balancing and high overhead in distributed load balancing.

APPENDIX A SIGNALING OVERHEAD OF SCSDA

According to the SCSDA presented in [3] and the scenario mentioned in this paper, the LBAs transfer load information as follows: a WiMax LBA transfers load information to its six neighboring WiMax LBAs, the UMTS LBAs and the WLAN LBAs whose coverage area overlaps part coverage area of the WiMax; a UMTS LBA transfers load information to the WiMax LBA whose part coverage area overlaps that of the UMTS, its six neighboring UMTS LBAs, and the WLAN LBAs whose coverage area overlaps part coverage area of the UMTS; a WLAN LBA transfers load information to the WiMax LBA and the UMTS LBA whose part coverage area overlaps that of the WLAN. WLAN system only covers the hot spots, and usually their coverage area doesn't overlap. Therefore, The WLAN LBAs do not transfer load information to each other.

On the basis of the analysis above, we can derive the periodic signaling overhead and the non-periodic signaling overhead of SCSDA.

The periodic signaling overhead of transferring load information among LBAs in unit time is

$$\begin{aligned} O_1' &= \frac{1}{2}\cdot\frac{1}{T}\left[\left(6a_1'A_3 + a_2'A_1 + a_3'A_2\right) + \left(a_2'A_1 + 6a_4'A_1 + a_5'A_2\right) + \left(a_3'A_2 + a_5'A_2\right)\right] \\ &= \frac{1}{T}\left[\left(a_2' + 3a_4'\right)A_1 + \left(a_3' + a_5'\right)A_2 + 3a_1'A_3\right] \end{aligned} \quad \text{(A1)}$$

where, a_1' denotes the signaling overhead of transferring load information once between two WiMax LBAs, a_2' denotes the signaling overhead of transferring load information once between one WiMax LBA and one UMTS LBA, a_3' denotes the signaling overhead of transferring load information once between one WiMax LBA and one WLAN LBA, a_4' denotes the signaling overhead of transferring load information once between two UMTS LBAs, a_5' denotes the signaling overhead of transferring load information once between one UMTS LBA and one WLAN LBA.

For problem tractability, we assume that $a_1' = a_2' = a_3' = a_4' = a_5'$, then (A1) can be reduced to:

$$O_1' = \frac{1}{T}\cdot a_1'\cdot\left(4A_1 + 2A_2 + 3A_3\right). \quad \text{(A2)}$$

The non-periodic signaling overhead of transferring load information among LBAs in unit time is

$$O_2' = \frac{1}{T}\cdot a_1'\cdot\left(4A_1\cdot \mathrm{Pr}_1 + 2A_2\cdot \mathrm{Pr}_2 + 3A_3\cdot \mathrm{Pr}_3\right). \quad \text{(A3)}$$

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 60972028).

REFERENCES

[1] S. Buljore, H. Harada, S. Filin, P. Houze, K. Tsagkaris, O. Holland, et al., "Architecture and enablers for optimized radio resource usage in heterogeneous wireless access networks," The IEEE 1900.4 Working Group Communications Magazine, vol.47, no.1, pp.122–129, Jan. 2009.

[2] K. Piamrat, A. Ksentini, J. M. Bonnin, C. Viho, "Radio resource management in emerging heterogeneous wireless networks," Computer Communications, in press.

[3] G. Djukanovic, M. Sunjevaric, N. Gospic, H. H. Chen, "Dynamic guard margin CAC algorithm with ensured QoS and low CDP in heterogeneous wireless networks," Computer Communications, in press.

[4] J. Sun, H. B. Zhu, "Optimization and Scheduling of Spectrum Sensing Periods in Heterogeneous Wireless Networks," 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2009, pp.1–5, 2009.

[5] P. Andreas, D. Fariborz, J. Enrico, M. T. Andreas, "Force-based load balancing in co-located UMTS/GSM networks," IEEE Vehicular Technology Conference, vol.6, no.6, pp.4402–4406, Sept. 2004.

[6] 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, Jul. 2008.

[7] B. Li, W. X. Shi, N. Li, "A Hierarchical Semi-centralized Architecture for Load Balancing of Heterogeneous Wireless Networks," 2010 Second International Conference on Networks Security Wireless Communications and Trusted Computing (NSWCTC), vol.2, pp.28–31, Apr. 2010.

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

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

[10] G. Q. Ning, G. X. Zhu, L. X. Peng, 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.

[11] I. Bird, B. Jones, K. F. Kee, "The Organization and Management of Grid Infrastructures," Computer, vol.42, no.1, pp.36–46, Jan. 2009.

[12] F. Yin, C. J. Jiang, R. Deng, J. J. Yuan, "Grid resource management policies for load-balancing and energy-saving by vacation queuing theory," Computers and Electrical Engineering, vol.35, no.6, pp.966–979, Nov. 2009.

[13] C. L. Li, L. Y. Li, "Optimization decomposition approach for layered QoS scheduling in grid computing," Journal of Systems Architecture, vol.53, no.11, pp.816–832, Nov. 2007.

[14] G. Rumi, L. Kristina, "Structure of heterogeneous networks," 12th IEEE International Conference on Computational Science and Engineering, CSE 2009, vol.4, pp.98–105, Aug. 2009.

[15] B. Venkata Ramana, Devesh Agrawal, C. Siva Ram Murthy, "Design and performance evaluation of meghadoot – A hybrid wireless network architecture," 2006 IEEE International Conference on Networks, vol.2, pp.605–610, Sept. 2006.

[16] P. Arun, V. Rajesh, T. Rajeev, N. Kshirasagar, "Multiple mobile routers based seamless handover scheme for next generation heterogeneous networks," First International Conference on Networks and Communications, NETCOM 2009, pp.72–77, Dec. 2009.

[17] M. X. Li, D. L. Xie, B. Hu, Y. Shi, S. Z. Chen, "A multi-hop routing mechanism based on fuzzy estimation for heterogeneous wireless networks," IEEE Vehicular Technology Conference, pp.1–5, Sept. 2009.

[18] K. Mokhtarian, M. Hefeeda, "Authentication of Scalable Video Streams with Low Communication Overhead," IEEE Transactions on Multimedia, in press.

[19] K. Sharif, L. J. Cao, Y. Wang, T. Dahlberg, "A Hybrid Anycast Routing Protocol for Load Balancing in Heterogeneous Access Networks," International Conference on Computer Communications and Networks, ICCCN 2008, pp.99–104, Aug. 2008.

[20] A. Zouari, L. Suciu, J. M. Bonnin, K. Guillouard, "A Proactive and Distributed QoS Negotiation Approach for Heterogeneous Environments: An Evaluation of QoS Signalling Overhead," The 2nd International Conference on Next Generation Mobile Applications, Services, and Technologies, NGMAST 2008, pp.41–46, Sept. 2008.

[21] A. Behr, L. Camarinopoulos, G. Pampoukis， " Domination of k-out-of-n systems," IEEE Transactions on Reliability, vol.44, no.4, pp.705－708, Dec. 1995.

[22] H. Frank, I. Frisch, "Analysis and Design of Survivable Networks," IEEE Transactions on Communication Technology, vol.18, no.5, pp.501–519, Oct. 1970.

[23] W. X. Shi, L. C. Zhang, K. G. Hu, Communication networks theory (in Chinese). Changchun, China: Jilin University Press, 2001.

[24] F. Houeto, S. Pierre, R. Beaubrun, Y. Lemieux, "Reliability and cost evaluation of third-generation wireless access network topologies: a case study," IEEE Transactions on Reliability, vol.51, no.2, pp.229–239, Jun. 2002.

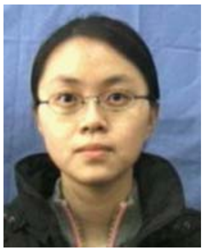
[25] W. X. Shi, L. C. Zhang, K. G. Hu, Y. Dong, Communication network: principle and applications(in Chinese). Beijing, China: Publishing House of Electronics Industry, 2008.

Wenxiao Shi received the Master degree of Engineering from Harbin Institute of Technology in 1991 and Ph.D in Communication and Information System from Jilin University in 2006. He is now a Professor of Department of Communication Engineering in Jilin University. His research interests include broadband communication theories and wireless communication network.

He is a co-author of the textbook principle and applications of communication network (Publishing House of Electronics Industry, 2008), Fast dynamic channel allocation algorithm based on priority channel reservation (Journal on Communications, 2009), and Hot-cell clusters deriving algorithm for slow dynamic channel allocation (Journal on Communications, 2009). He has a long-term cooperation with China Mobile Group Guangdong Company, engaging in the theory and technology research of programming and optimizing, performance evaluation and quality analysis in mobile communication network. He engaged in the research of wireless resource management technology, and was responsible for the project of The Science-Technology Plan of Jilin Province: Fast Dynamic Channel Allocation technology in TD-SCDMA; current is responsible for the project of The National Natural Science Foundation of China: researching on distributed Grid-based heterogeneous wireless network architecture and load balancing methods.

Prof. Shi is a member of Institute of Communications of Jilin Province and China Association of Automation of Jilin Province.

Bin Li received the Bachelor Degree in Communication Engineering from Jilin University in 2008. He is now a postgraduate student of Jilin University. His research interests include Core Network technology of GPRS, Radio Resource Management technology in TD-SCDMA and Load Balancing technology for Wireless Heterogeneous Networks. He is a co-author for two papers and a patent now.



Na Li received the Bachelor Degree in Communication Engineering from Jilin University in 2009. She is now a graduate student of Jilin University. His current research interests include the Load Balancing technology for Wireless Heterogeneous Networks.

Chuanjun Xia received the Bachelor Degree in Communication Engineering from Jilin University in 2008. He is now a postgraduate student of Jilin University. His current research interests include Load Balancing Technology for Wireless Heterogeneous Networks.