

Traffic-Aware Resource Management in Heterogeneous Cellular Networks¹

Cheng-Fu Chou

Department of Computer Science
and Information Engineering
National Taiwan University
Taipei, Taiwan
Email: ccf@csie.ntu.edu.tw

Ching-Ju Lin

Graduate Institute of
Networking and Multimedia
National Taiwan University
Taipei, Taiwan
Email: cjlin@cmlab.csie.ntu.edu.tw

Chung-Chieh Tsai

Department of Computer Science
and Information Engineering
National Taiwan University
Taipei, Taiwan
Email: goose@cmlab.csie.ntu.edu.tw

Abstract—Due to a rapid progress in mobile cellular networks and increasing demand for diverse mobile services, lots of researchers have focused on how to efficiently integrate multiple heterogeneous sub-systems into the next-generation network. Since the channel is the most critical resource in the cellular network and how to allocate and de-allocate the channels has great impact on the system performance, resource management in heterogeneous cellular networks is an important issue. In this paper, we propose a traffic-aware resource management scheme based on the traffic pattern and the penalty function to maximize the system utilization and the revenue for the system provider. There are two main components – the resource management module and the online traffic monitor module – in our approach. The resource management module can adaptively control all resources in the same cell rather than separately manage the resource on each base station. Specifically, the whole resource on 2G and 3G base stations is partitioned to multiple sub-pools, which are reserved for each specific traffic class. The size of each sub-pool is dynamically adjusted according to the traffic load. The online traffic monitor module is used to trace online traffic condition so as to react the network change. The results of performance evaluation show that our approach can achieve significant improvement in terms of resource utilization and system revenue.

I. INTRODUCTION

There has been a rapid progress in mobile communication network for the past decade. The second-generation (2G) cellular systems such as GSM and IS-95, which have evolved from analog to digital technology, have got successful and popular nowadays. 2G cellular systems are designed to provide voice and low-rate circuit-switched data services, while the third-generation (3G) cellular systems are aimed at providing multimedia mobile service and achieving maximum bit rate of 2Mb/s. With the increasing demand for diverse mobile services at anywhere on anytime, lots of researchers have focused on how to efficiently integrate multiple subsystems into a homogeneous system – the next-generation cellular network.

Integrating multiple subsystems into next-generation networks brings about many challenges such as subsystem interworking and end-to-end quality of service for different appli-

cations. Also, the channel is the most critical resource in the cellular network, so how to efficiently allocate the channels will heavily impact the system performance. A poor channel management scheme could lead to higher call blocking or dropping probability for users as well as lower system utilization for the service provider. In this work, we consider the channel management problem in the context of a heterogeneous environment, i.e., different sub-mobile systems and different user mobile devices. That is, due to the heterogeneity of multiple mobile systems, some mobile devices are only able to communicate with certain types of base stations, so, in other words, various mobile devices and diverse service requirements has made the channel management problem more difficult. Therefore, we focus on how to efficiently integrate the 2G and 3G cellular systems. However, our approach can also be applicable to the integration of any multiple subsystems.

There are lots of studies, e.g., [10] [11] [13], discussing channel management problem for 3G system, but they only focus on how to maximize the utilization of scarce resource under the homogeneous 3G environment and meanwhile maintain the quality of service (Qos) for each call. Nonetheless, to the best of our knowledge, no previous works discuss how to integrate the radio resource of a cell that contains both 2G and 3G base stations (BS). In such heterogeneous network environment, the communication capability of base stations and mobile devices will be different, i.e. 2G BSs can only support digital-speech service, single-mode devices are only able to communicate with 2G BS, but 3G BSs can provide both speech as well as multimedia services. The traditional resource management schemes do not take these issues into account and will be inefficient under such heterogeneous environment. The used *separate-control* scheme – which assumes that 2G BS and 3G BS in the same cell apply their own schemes to manage their radio resource separately, i.e. the single-mode user equipments (UEs) can only be served by 2G BS, and the multi-modes UEs can only be served by 3G BS – cannot efficiently utilize all available resources and may cause load imbalance between BSs. For example, a multi-mode device, which requests the audio service, will be blocked if there is no sufficient resource on 3G BS, even though there is still enough resource on 2G BS for this multi-mode device. Consequently,

¹This work was partially supported by the National Science Council and the Ministry of Education of ROC under the contract No. NSC93-2213-E-002-094 and NSC 93-2622-E-002-033.

for improving the utilization of radio resource, the *Traffic-aware Resource Management Scheme* is proposed to manage all available resources, both 2G and 3G radio resource, in the same cell. That is, our *Traffic-aware Resource Management Scheme* considers (a) the heterogeneity of mobile systems as well as mobile devices, and (b) dynamic traffic pattern. Thus, in our system there are two main modules: *resource management module* and *online traffic monitor module*; the *resource management module* is used to allocate the resource adaptively for each call, and the *online traffic monitor module* is used to observe the past traffic pattern and predict the possible request in the future.

The paper is organized as follows. Related works that focus on resource management under the homogeneous and heterogeneous wireless cellular environments are given in section II. Our *Traffic-aware Resource Management Scheme* will be described in detail in section III. The performance evaluation of our schemes via simulations is showed in section IV. Finally, conclusions are presented in section V.

II. RELATED WORKS

In the past decade there have been a number of useful strategies [1-14] proposed to efficiently manage the channels in a cellular system. [3] proposes the concept of a shadow cluster – a set of base stations to which a user device is likely to attach in the near future. The scheme partitions time into equal intervals, and estimates the probability of each user device being in any wireless cell within the shadow cluster for future time intervals based on individual user device's dynamics and call holding patterns in the form of probability density functions. During each time interval, the base station exchanges information about the predicted bandwidth demands for future time intervals so as to determine the feasibility of admitting new call requests. In [9] and [14], Soh et al. propose a dynamic bandwidth reservation scheme that utilizes knowledge of road topology, in addition to positioning information. In addition, this scheme is the first one capable of handling irregular cell boundaries. The results show that such approach could potentially achieve more accurate predictions at the cost of increased complexity, but the resulting gain in reservation efficiency may justify this cost. We note that all above works focus on how to make use of the channel in a homogeneous cellular network through more precise and correct prediction for the user device. Our work is to provide an efficient channel management method to deal with heterogeneous cellular systems and mobile devices. Consequently, our approach is orthogonal to the above works and could be combined with approaches mentioned above to improve the system performance.

The only work on resource management scheme that considers the heterogeneity of mobile system as well as user devices is [12]. In this paper, the call request is classified into three different classes, i.e., 2G speech, 3G speech, and 3G packet. To maximize the system utilization, the authors propose to place any 3G speech call into 2G base station first. To be fair to the 2G speech call, the system will migrate a 3G speech call back

to 3G base station when a 2G speech call arrives and there is no available channel in the 2G base station. This approach is simple, but the migration (or vertical handoff) procedure will lead to high delay and control overhead. To avoid the vertical handoff and improve the utilization of the resource in the heterogeneous cellular systems, we propose a *Traffic-aware resource management scheme* that considers the traffic load and traffic type of each traffic class while maximizes the revenue of service provider. The radio resource in a cell is divided into three sub-pools and each sub-pool is reserved for a particular traffic class. According to the amount of reserved resource and the volume of traffic of each traffic class, we can determine if the system will accept a new/handoff call or not.

III. TRAFFIC-AWARE RESOURCE MANAGEMENT SCHEME

The *Traffic-aware resource management scheme* considers the heterogeneity of mobile systems and the specific of different user equipments (UE). Moreover, we take the penalty of blocking or dropping a call into account, because different types of call may have various penalties, e.g. the service provider can get more profits from serving a video call, so for increasing the revenue it would rather block the audio call than video call when there are not sufficient unused resources for both of them. The objective of this paper is to maximize both the utilization of radio resource and the revenue of service provider.

The user equipments are classified into two types – Single-Mode UE (SM-UE) and Multi-Mode UE (MM-UE) – according to their communication capabilities. The main difference between SM-UE and MM-UE is given in the following:

- **Mobile System:** MM-UEs are capable of communicating with both 2G and 3G base stations, while SM-UEs can only communicate with 2G base station.
- **Traffic Class:** SM-UEs can only request audio service, while MM-UEs can ask for either audio or video service.

For easily addressing our scheme, the call requests are grouped into three traffic classes (U_t) as follows.

- **Single-Mode Audio (2SM-A):** Owing to the restriction of SM-UEs' communication capability, only 2G base station can serve traffic class. A new 2SM-A call will be blocked if there is no sufficient available resource to serve it on nearby 2G base stations.
- **Multi-Mode Audio (MM-A):** MM-A is the only class that can be served by either 2G or 3G base station in the heterogeneous wireless cellular networks.
- **Multi-Mode Video (3MM-V):** Owing to the restriction of multimedia service, a 3MM-V call request can only be served by 3G base station.

A. System Architecture

Fig. 1 illustrates the system architecture composed of *resource management module* and *online traffic monitor module*.

Resource Management Module (RMM): The radio resources supported by either 2G BS and 3G BS will be divided into multiple sub-pools and each sub-pool is reserved for certain traffic class, such as 2SM-A or 3MM-V. When

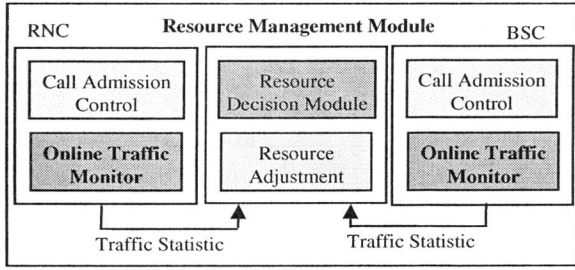


Fig. 1. System Architecture

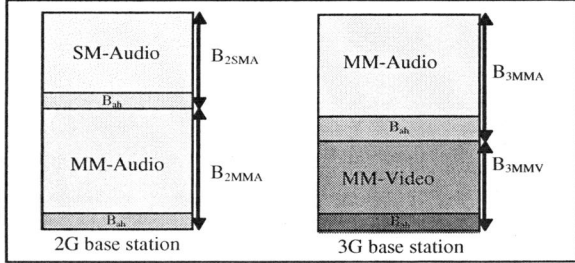


Fig. 2. Partition of Available Resource

a new call arrives, the admission control process in RMM will determine to set up the connection or block this request based on the amount of unused available bandwidth in the corresponding sub-pool.

Online Traffic Monitor Module: To maximize the utilization of radio resources, the size of each reserved sub-pool should be dynamically adjusted based on the traffic pattern and penalty function. Consequently, first of all, traffic pattern should be observed and analyzed by the *online traffic monitor module* on each base station, and subsequently submitted to the RMM. After identifying the traffic statistics, the RMM can adjust the size of reserved sub-pool based on the utilization of all resources, new call arrival pattern, and average service time of each cell.

B. Resource Allocation Scheme

1) *Admission Control Policy:* As Fig. 2, the RMM divides all resources (either on 3G BS or 2G BS) into three parts, and each of them is reserved for corresponding traffic class in the heterogeneous cellular network. The 2SM-A and 3MM-V calls can get the resource from B_{2SMA} and B_{3MMV} , respectively. However, remaining sub-pools, B_{2MMA} and B_{3MMA} , will be reserved for MM-A traffic class. Otherwise, for each sub-pool, there is an area, B_{ah} , reserved for handoff calls.

Next, the admission control policy proposed in our work is showed in Table I. The available bandwidth, $AB(sp)$, for a handoff call is total resources of corresponding sub-pool, while the $AB(sp)$ for a new call is "available bandwidth in sub-pool - $B_{ah}(sp)$ " since some bandwidth should be reserved for handoff calls. After receiving the request, the RMM will detect if unused available bandwidth is sufficient for this request. For example, when the traffic class of request call is 2SM-A, the RMM will check the residual available resource on B_{2SMA} , i.e. $AB(2SM-A) - UB(2SM-A, t)$. If the remainder available resource is larger than request requirement, B_r , the service provider allocates B_r for this request and reset the value of $UB(2SM-A, t)$. The same method can be used to allocate the resources for traffic class 3MM-V. However, when

(a) NOTATIONS IN ADMISSION CONTROL POLICY	
Notation	Meaning
sp (2SM-A/3MM-V)	The id of each sub pool, which is reserved for corresponding type of traffic.
$B(sp)$	The amount of available resource in each sub-pool
$AB(sp)$	The available resource for a new/handoff call in each sub-pool
$UB(sp)$	The resource has been allocated (used) for admitted call in each sub-pool
B_r	The resource requirement of incoming call request
$B_{ah}(sp)$	The resource reserved for handoff call in each sub-pool

(b) ALGORITHM	
1	if (the request is a new call) then
2	$AB(each_sp) = B(each_sp) - B_{ah}(sp)$
3	else
4	$AB(each_sp) = B(each_sp)$
5	endif
6	
7	if ($U_t == 2SM - A$ or $3MM - V$) then
8	if ($AB(U_t) - UB(U_t, t) >= B_r$) then
9	allocate resource and $UB(U_t, t) += B_r$
10	else
11	block the new call request
12	endif
13	else if ($U_t == MM-A$) then
14	if ($AB(3MM-A) - UB(3MM-A, t) >= B_r$) then
15	allocate resource and $UB(3MM-A, t) += B_r$
16	else if ($AB(2MM-A) - UB(2MM-A, t) >= B_r$) then
17	allocate resource and $UB(2MMA, t) += B_r$
18	else
19	block the new call request
20	endif
21	endif

TABLE I
ADMISSION CONTROL POLICY FOR NEW/HANDOFF CALL

1	$B = B_2 + B_3$
2	$\pi = \pi_{SMA} + \pi_{MMA} + \pi_{MMV}$
3	$B_{SMA} = B * (\pi_{SMA} / \pi)$
4	$B_{MMA} = B * (\pi_{MMA} / \pi)$
5	$B_{MMV} = B * (\pi_{MMV} / \pi)$

TABLE II
PROPORTIONAL-ADJUSTMENT ALGORITHM

the traffic class of request call is MM-A, the new call can get the resource from either B_{2MMA} or B_{3MMA} . Therefore, the RMM will check unused available bandwidth of both B_{2MMA} and B_{3MMA} , and the request call can be admitted if residual available bandwidth in at least one of these two sub-pools is larger enough. If neither 2G BS nor 3G BS can be provided B_r for request call, it will be blocked.

2) *Dynamically Proportional Adjustment:* As Table II, the proportion of $B_{SMA}/B_{MMA}/B_{MMV}$, $\pi_{SMA}/\pi_{MMA}/\pi_{MMV}$, must be dynamically adjusted to meet the time-varying traffic condition. The number of resource reserved for each traffic class should be proportional to parameter π . In the following, three schemes, *Proportional-Tr*, *Proportional-Tr_Pn*, and *Proportional-N_Pn*, are specified to calculate the proportional parameter π .

$$\text{Proportional - Tr : } \pi_i = Tr_i \quad (1)$$

$$\text{Proportional - Tr_Pn : } \pi_i = Tr_i * Pn_i \quad (2)$$

$$\text{Proportional - N_P : } \pi_i = (Tr_i / R_i) * Pn_i \quad (3)$$

$$N_i = \lambda_i * T(\text{Little's Theorem})$$

$$Tr_i = N_i * R_i \quad (4)$$

Proportional-Tr, Equation 1, intuitively partitions the resource based on the traffic load in each traffic class. The predicted traffic load of i_{th} traffic class, Tr_i , can be evaluated by the Little's Theorem, Equation 4. Here, the N_i , λ_i , and T are the average number of calls, average call arrival rate and average service time for a call, respectively, and R_i is the resource requirement of each call in i_{th} traffic class. However, the service provider may be able to gain different profit from various types of traffic classes, e.g. the service provider may be able to get more profit by admitting a video call than an audio call. So the system should not only consider traffic load but total revenue to determine the proportion of each traffic class. In *Proportional-Tr-Pn*, and *Proportional-N-Pn*, Equation 2 and 3, the Pn_i means penalty function, the profit that the service provider can earn from admitting a call in i_{th} traffic class (i.e. the profit service provider will lose due to blocking a call in i_{th} traffic class). *Proportional-Tr-Pn* takes penalty function into account, but, however, different types of calls may require various size of bandwidth requirement, R_i . Therefore, the formula divides by R_i in *Proportional-N-Pn*.

For accurately predicting the traffic load in dynamically proportional adjustment algorithm, we propose two methods, window-based scheme and interarrival-time-based scheme, to calculate the λ value.

$$\lambda = (N_n + N_h)/\Delta T \quad (5)$$

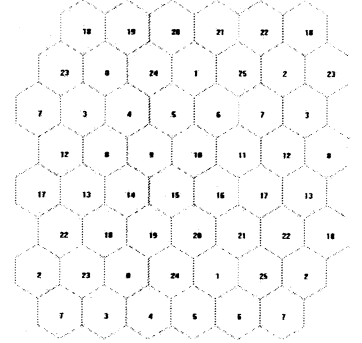
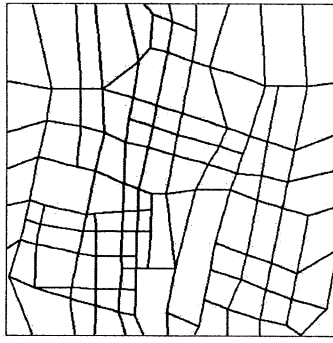
Window-Based Scheme: As Equation 5, the arrival of new call and handoff call, N_n and N_h , will be accumulated in each window size, ΔT , respectively. However, the window size will directly influence the estimated arrival rate. The small window size reduces the time of collecting information but further decrease the accuracy of estimation, while the larger window can increase the accuracy but cannot immediately reflect the real traffic pattern due to delay of getting information. Because of the trade-off in selecting window size, the second method, time-Base scheme, is proposed.

$$\lambda = 1/(t_n - t_{n-1}) \quad (6)$$

$$\lambda_n = \alpha * \lambda_t + (1 - \alpha) * \lambda_{n-1} \quad (7)$$

Interarrival-Time-Based Scheme: As Equation 6, the t_n and t_{n-1} are the arrival time of the n_{th} and $n-1_{th}$ incoming call, respectively. Because the average arrival rate can be estimated in each call request, the system can quickly react to the real traffic pattern. However, a single event may not represent the average traffic pattern, so we use the Exponentially Weight Moving Average (EWMA) scheme to smoothly adjust λ as Equation 7. The EWMA scheme can avoid the estimation results from oscillating. The λ_t is currently estimated value and the λ_{n-1} is last statistic, and the values α and $1 - \alpha$ are the weight of λ_t and λ_{n-1} , respectively. (The α is set as 0.2 in our simulation)

When the estimated arrival rate exceeds a pre-defined threshold, the *online traffic monitor module* will inform the *resource adjustment module* to adjust the proportion of reserved resource for each traffic class.



(a) The Road Layout in simulation (b) Simulation network with wrap-around at network boundary
Fig. 3.

IV. PERFORMANCE EVALUATION

The performance of our resource management scheme is evaluated by the simulation. In the following sub-section, we introduce the simulation environment and baseline schemes used to be compared with our scheme, and finally show the results of simulation.

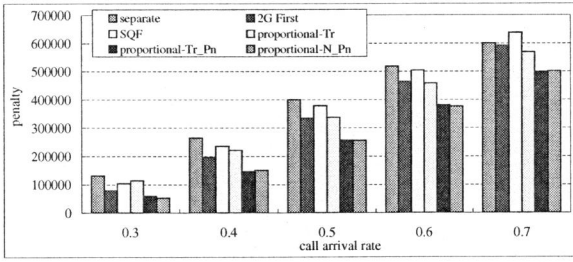
A. Road Layout Database

Each base station needs to keep a database of the roads within its coverage area. As Fig. 3(a), the road layout, which constraints the movement of UE, is adopted to more realistically simulate the cellular network. The base station should keep some information in the database for each road segment as follows:

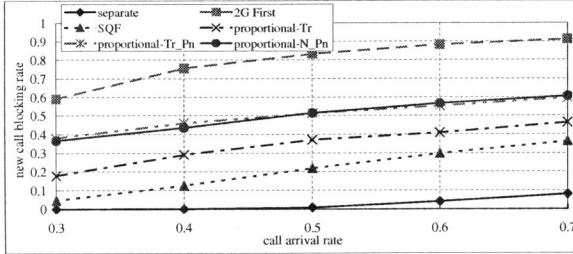
- 1) The identify of each segment
- 2) The coordinate of two end point
- 3) The length of road segment
- 4) Two neighboring segments which is contiguous to this segment
- 5) The probability that the UE travels from this segments to its neighboring segment

B. Simulation environment and Traffic model

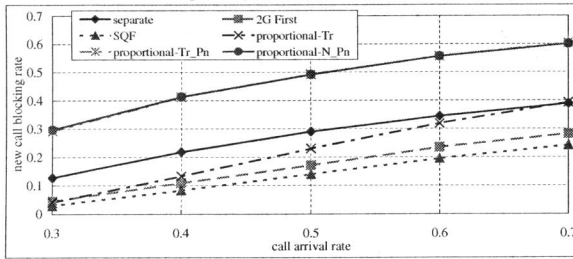
In the simulation, the coverage area of each cell and its base station is 1000m. There are 26 wireless cells in the simulation area, and for each cell there exist one 2G and one 3G base-stations. All simulation results are the averages of 26 cells. Each 2G and 3G base station is assumed to contain a fixed link capacity C , 100BU s and 300BU s respectively. The BU means the resource requirement for each audio call request. In order to eliminate boundary effects, a approach proposed in [14] is used in our simulation. The traffic model used here is also similar to [14]. For each base station, the connection requests are generated based on Poisson distribution with rate λ (connection/sec/base station). The users who request the service are uniformly distributed among all road segments. The probabilities of requests belonged to different traffic classes, SM-A, MM-A, and MM-V, are Pue , $(1 - Pue) * Pau$, and $(1 - Pue) * (1 - Pau)$, respectively. Pue , ue ratio, is the ratio of SM-UEs to total number of user equipments, and Pau , audio ratio, is ratio of audio traffics of the MM-UEs to the total traffics of MM-UEs. The lifetime for each connection is



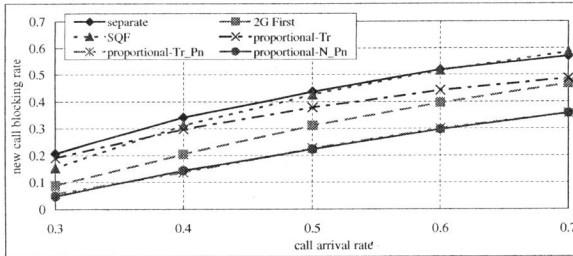
(a) Total penalty



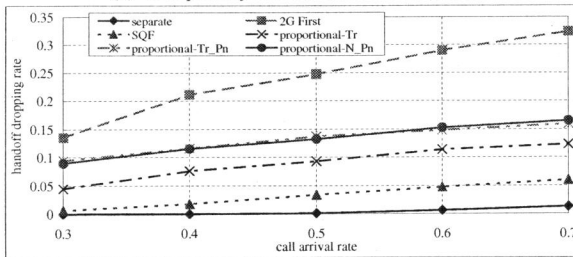
(b) The penalty of new call SM-A traffic



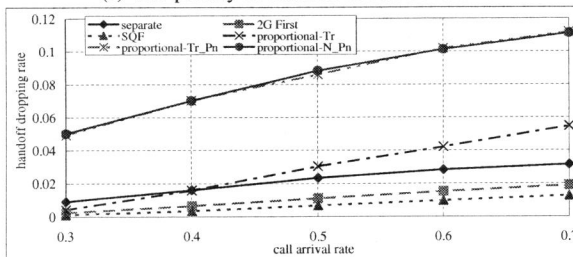
(c) The penalty of new call MM-A traffic



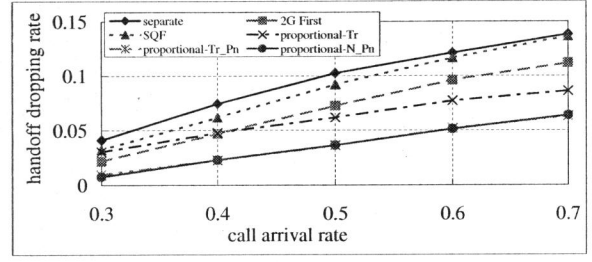
(d) The penalty of new call MM-V traffic



(e) The penalty of handoff call SM-A traffic

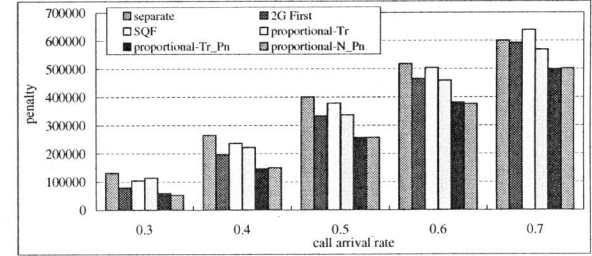


(f) The penalty of handoff call MM-A traffic

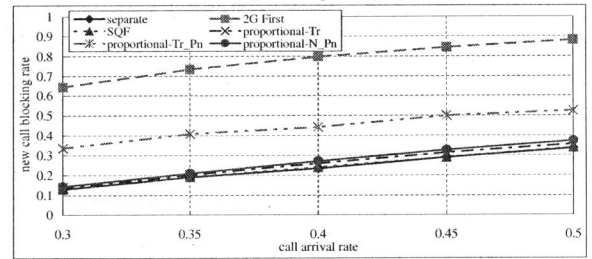


(g) The penalty of handoff call MM-V traffic

Fig. 4. Call arrival rate vs. call blocking penalty (0.18 0.55 0.27, 1 1 20)



(a) Total penalty



(b) The penalty of SM-A traffic

Fig. 5. Call arrival rate vs. call blocking penalty (0.48 0.48 0.02, 1 1 6)

exponentially distributed with mean 180sec, and the resource requirement for calls in three traffic classes, SM-A, MM-A and MM-V, are 1 BUs, 1 BUs, 4 BUs respectively.

C. Other compared schemes

The following three schemes are chose as baselines to be compared with our scheme.

- 1) **Separate control:** The 2G BS and 3G BS manage their own resources. Also, the 2G BS can only serve single-mode users, and the 3G BS can only server multi-mode users.
- 2) **2G First:** The MM-A request will be served by 2G BS only as long as there is still sufficient available resource on 2G BS. Otherwise, the request will be served by 3G BS.
- 3) **Shortest Queue First (SQF):** The MM-A request will be served by the base station which remains maximum unused available resource.

D. Performance Metrics

$$CBP = \Sigma PN_n * (PN_i * NB_i) + PN_h * (PN_i * HB_i) \quad (8)$$

The performance metric used to evaluate the performance is *Call Blocking Penalty (CBP)*. The penalty is defined as the profit that the service provider loses due to blocking a new call or dropping a handoff call. The *CBP* is defined as Equation 8. PN_n, PN_h and PN_i denote the penalty of new call request,

the penalty of handoff call request, and the penalty of traffic class i , respectively. In our simulation, PN_n and PN_h are set as 1 and 2, and NB_i and HB_i are the number of blocked new calls and handoff calls in traffic class i .

E. Simulation Result

Multiple traffic patterns are simulated to evaluate the performance of our scheme. In first traffic pattern, the new call arrival probability of class SM-A, MM-A, and MM-V will be 0.18, 0.55, and 0.27, and the penalty of class SM-A, MM-A, and MM-V will be 1, 1, and 20, respectively. Fig. 4(a) shows that our *proportional-based* scheme (*proportional-Tr*, *proportional-Tr_Pn*, and *proportional-N_Pn*) can cause fewer call blocking penalties. In such traffic pattern, the *separate control* performs most inefficiently. Fig. 4(d) shows that *separate control* scheme causes higher P_{NB} in MM-V traffic. Note that the MM-A and MM-V traffics have to contend the resource on 3G BS, so the resource that can be used by MM-V decreases heavily. Also, the penalty of MM-V traffic is much higher than SM-A and MM-A traffic, so the higher penalty caused by blocking more MM-V traffics will greatly reduce the total revenue of service provider. On the other hand, there may be lots of unused resources on 2G BS even though the 3G BS is overloaded, so the utilization of resource degrades.

The *SQF* scheme performs better than *separate control* because some MM-A traffics can be shared by 2G BS. However, the improvement is limited since the MM-V traffic still needs to contend the 3G resource with MM-A traffic even though there are unused 2G resources.

The *2G-First* has best performance among three baseline schemes. Because MM-A calls could be served by the 2G BS as many as possible, the maximum amount of 3G bandwidth is available for the MM-V traffic. However, the system doesn't reserve resources for any traffic class, so all kinds of requests are served by first-come-first-serve scheduling order, i.e. the MM-V traffic may not be able to get resource if the available resources are all allocated to audio traffics.

Fig. 4(a) illustrates our *proportional-based* schemes can perform better because the system can flexibly manage the resources between heterogeneous mobile systems. Besides, the *proportional-Tr_Pn* and *proportional-N_Pn* can maximize the revenue for service provider due to considering the differential profit of each traffic class.

In second traffic pattern, the new call arrival probability of class SM-A, MM-A, and MM-V will be 0.48, 0.48, and 0.02, and the penalty of class SM-A, MM-A, and MM-V will be 1, 1 and 6, respectively. As Fig. 5(a), the *2G-First* is the most inefficient scheme among three baseline schemes, because the SM-A traffics should content the 2G resource with MM-A traffics. The SM-A request would be blocked when 2G BS is overloaded even though there still exist many unused resources on 3G BS. But, actually, the MM-A requests, now served by 2G BS, can be moved to 3G BS so that the 2G BS can reserve more resources for serving SM-A traffics. We can also see from Fig. 5(b) that *2G-First* causes higher penalty in SM-A traffic than other schemes. On the other

hand, the *proportional-Tr_Pn* performs worst among three of our proposed schemes because it neglects to consider the number of resource required for different traffic class so that the system over-reserve the resource for MM-V traffics. From above simulation results, we conclude that our *proportional-N_Pn* scheme can perform better in average and maximize the revenues for service provider.

V. CONCLUSION

In this paper, we investigated the issues of resource management in the heterogeneous wireless network. We propose two modules, *resource management module* and *traffic monitor module*. In *resource management module*, all resources under heterogeneous environment are partitioned into sub-pools, and each sub-pool is reserved for each traffic class. The *traffic monitor module* can analyze and provide traffic statistics for management module, which can dynamically adjust the proportion of each sub-pool based on monitored traffic pattern. The performance evaluation of our scheme shows that our *proportional-based* scheme, which considers the penalty of blocking call, can achieve better revenue for service provider.

REFERENCES

- [1] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoffs procedures," Aug. 86.
- [2] N. M and S. M., "Distributed call admission control in mobile/wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 4, pp. 711-716, May 1996.
- [3] D. A. Levine, I. F. Akyildiz, and M. Naghshines, "A resource estimation and call admission algorithm for wireless multimedia networks using the shadow cluster concept," *IEEE/ACM Trans. Networking*, vol. 5, no. 1, pp. 1-12, Feb 1997.
- [4] Y. R. Haung, Y. B. Lin, and J. M. Ho, "Performance analysis for voice/data integration on a finite-buffer mobile system," Feb., 2000.
- [5] M. GD and G. WD, "Strategies to maximize carried traffic in dual-mode cellular systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 357-366, Mar. 2000.
- [6] N. M and A. AS., "Qos provisioning in micro-cellular networks supporting multimedia traffic," in *Proceedings of IEEE INFOCOM*, April 1995, pp. 1075-1084.
- [7] B. Epstein and M. Schwartz, "Reservation strategies for multi-media traffic in a wireless environment," in *Proc. IEEE Veh. Technol. Conf.*, 1995.
- [8] Z. D and Z. M., "Performance and efficiency evaluation of resource allocation schemes for hscsd in gsm," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1B, 1999, pp. 1084-1088.
- [9] W.-S. Soh and H. S. Kim, "Dynamic guard bandwidth scheme for wireless broadband networks," in *Proceedings of IEEE INFOCOM*, Apr. 2001, pp. 572-581.
- [10] A. Hac and A. Armstrong, "Resource allocation scheme for qos provisioning in microcellular networks carrying multimedia traffic," *Int. J. Network Mgmt*, vol. 11, pp. 277-307, 2001.
- [11] C. Lindemann, M. Lohmann, and A. Thummler, "Adaptive performance management for umts networks," in *Computer Networks*, 2002, pp. 477-496.
- [12] S. L. Salecker and C. S. Hood, "Integrated networks that overflow speech and data between component networks," *Int. J. Network Mgmt*, vol. 12, pp. 235-257, 2002.
- [13] C. Lindemann, M. Lohmann, and A. Thummler, "A unified approach for improving qos and provider revenue in 3g mobile networks," *Mobile Networks and Application*, vol. 8, pp. 209-221, June 2003.
- [14] W.-S. Soh and H. S. Kim, "Dynamic bandwidth reservation in cellular networks using road topology based mobility predictions," in *Proceedings of IEEE INFOCOM*, Mar. 2004.