A fair protocol for fast resource assignment in wireless PCS networks *

Tsan-Pin Wang, Chien-Chao Tseng and Shu-Yuen Hwang

Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, ROC

Efficient sharing of communication resources is essential to PCS networks since the wireless bandwidth is limited. The Resource Auction Multiple Access (RAMA) protocol was recently proposed for fast resource assignment and handover in wireless PCS networks. The RAMA protocol assigns available communication resources (e.g., TDMA time slots or frequency channels) to subscribers one at a time using a collision resolution protocol based on subscriber ID's. However, the RAMA protocol encounters an *unfairness* problem; furthermore, performance results also indicate that it is inefficient at transmitting fixed-length subscriber ID's. Moreover, the emerging services such as teleconferencing have been presenting new challenges to dynamic-priority resource assignment. In this paper, we propose a modification to the RAMA protocol to improve its performance and resolve the unfairness problem. The proposed protocol also adopts dynamic priority assignment to improve the QOS for subscribers in overload environments.

1. Introduction

A *personal communication services* (PCS) network [5,6] is a digital communication system that enables subscribers to communicate with each other at any time from any location. To support the user mobility, a wireless link should be established before connection. Efficient sharing of communication resources (e.g., TDMA time slots or frequency channels) is essential to PCS networks since the wireless bandwidth is limited and very scarce.

As the demand for new services increases, next-generation wireless PCS networks will need to support integration of various types of data, such as voice, video, and multimedia data in mobile computing environments [7,14]. Consequently, there is a need for fair access and fast resource assignment for call origination and handoff due to the huge numbers of users and the small sizes of cells (e.g., microcells or picocells [8]) in future wireless PCS networks. Moreover, it is hard to provide acceptable quality of service (QOS) for emerging services such as teleconferencing in fixed-priority resource assignment. The emerging services have been presenting new challenges to dynamic-priority resource assignment [14].

Multiple access is one of the most important issues in communication networks, especially in wireless networks. In the literature, several categories of multiple access protocols have been studied [2,9,10,13], including *fixed-assignment*, *random access*, and *demand-assignment*. Fixed-assignment protocols are inefficient because the assigned bandwidth is wasted when the user has nothing to transmit. On the other hand, contention-based random access protocols [13] encounter stability problems in heavy load environments. Under heavy traffic, the throughput of contention protocols decreases rapidly because time slots are wasted in collisions between subscribers accessing at random times.

To eliminate these problems, a demand-assignment protocol, the Resource Auction Multiple Access (RAMA) protocol, was proposed [2–4,10]. The RAMA protocol is a deterministic algorithm that can provide good performance even under heavy loads. Each subscriber ID consists of a 9-digit phone number and a priority digit. In each assignment cycle, the subscriber with the *highest ID value* is the unique winner. Using this collision resolution method, the RAMA protocol creates unfairness problems. For example, among subscribers with the same priority, the one with the largest phone number will always 'win' auctions. Furthermore, even when only one subscriber requests a communication resource, the entire fixed-length subscriber ID is still transmitted one digit at a time in the RAMA protocol. This fixed-length auction cycle significantly degrades the performance of the RAMA protocol.

It is shown that RAMA is one of the most promising multiple access protocol. However, although RAMA provides enough assignments per second for many applications in the cellular environment [4], it still poses us unfairness and performance inefficiency problems. Furthermore, adopting dynamic priority assignment is impossible in RAMA. Therefore, we propose an efficient and fair protocol called random RAMA to adopt dynamic priorities for fast resource assignment in future high-capacity wireless PCS networks.

The remainder of this paper is organized as follows. In section 2, we present a brief introduction to the RAMA protocol. In section 3, we propose the random RAMA protocol. Then in section 4, we present a performance analysis of the random RAMA protocol. Section 5 shows a scheme for dynamic priority assignment. Section 6 discusses some issues for random RAMA. Finally, we conclude this paper in section 7.

* This research was supported by the National Science Council, ROC, under grant NSC 85-2213-E-009-063 and NSC 86-2213-E-009-076.

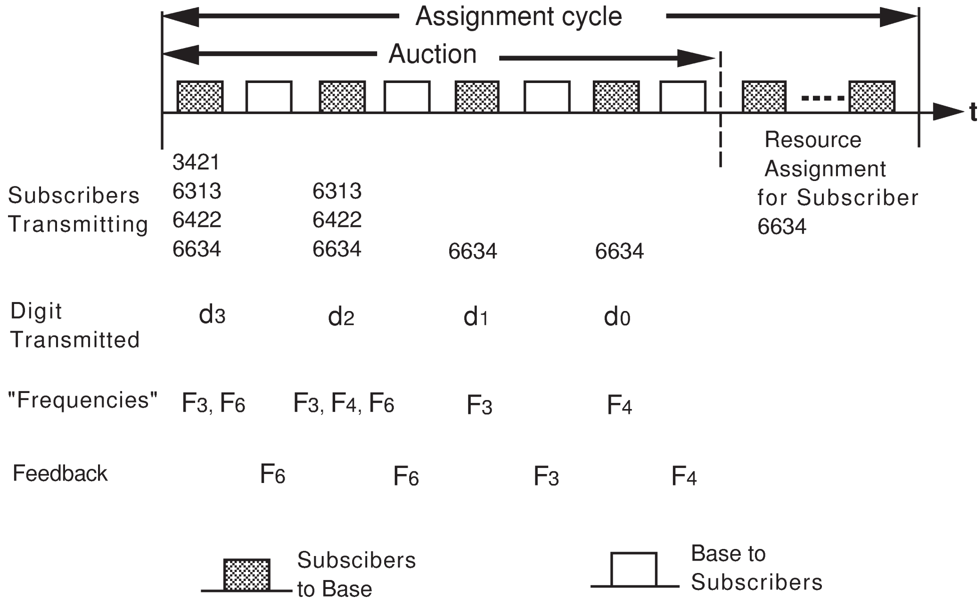


Figure 1. Example of RAMA auction.

2. Resource auction multiple access protocol

The resource auction multiple access (RAMA) protocol that is a demand-assignment access protocol facilitates fast access and resource assignment to spatially distributed subscribers in a deterministic manner irrespective of loading.

In RAMA, each subscriber has a unique ID that consists of a priority digit and a nine-digit phone number. Available communication resources are ‘auctioned’ one at a time using a collision resolution algorithm based on subscriber’s IDs. The subscriber with the *highest ID value* “wins” in an auction cycle. After each auction cycle, the base assigns an available communication resource to the winner. This assignment cycle is repeated for other available resources until either all requests are satisfied, or no more resources are available.

Subscribers requesting communication resources transmit their ID’s one digit at a time. The set of values that can be assumed by a digit is represented by a set of orthogonal signals such as M-ary FSK or binary ASK. Figure 1 shows an example of a RAMA auction. In this example, we assume that each subscriber ID ($d_3d_2d_1d_0$) is represented by a 8-ary FSK (i.e., $0 \leqslant d_i \leqslant 7$), and subscribers with ID’s 3421, 6313, 6422 and 6634 are seeking communication resources. In the first auction time slot, all subscribers transmit their most significant digit d_3, i.e., subscriber 3421 transmits a ‘3’ by transmitting an F_3 FSK signal, and subscribers 6313, 6422 and 6634 transmit ‘6’s by transmitting F_6 FSK signals. The base detects these orthogonal signals, and feeds back the largest digit, ‘6’, to the subscribers by transmitting an F_6 FSK signal. Upon receiving this feedback, all subscribers with most significant digits lower than ‘6’ drop out of the auction, and wait for the next assignment cycle. Those subscribers with $d_3 = 6$ continue by transmitting the next digit (d_2). In this example, subscribers 6313, 6422 and 6634 transmit ‘3’, ‘4’ and ‘6’, respectively. After the base announces that ‘6’ is the large d_2 digit transmitted by this group of remaining subscribers, subscribers 6313 and 6422 drop out of the auction. Subscriber 6634 continues by transmitting the remaining digits (d_1 and d_0) one digit at a time, and the base feeds back the corresponding digits. When the entire fixed-length ID has been transmitted, the base broadcasts a resource assignment for subscriber 6634, the unique winner of the auction. In the next assignment cycle, subscribers that dropped out of the previous cycle (i.e., 3421, 6313 and 6422) participate in a new auction along with requests from other new subscribers.

Note that although the winner 6634, in this example, has been uniquely identified by the base after transmitting the d_2 digit, the entire fixed-length ID still must be transmitted in the RAMA protocol. Moreover, if subscriber 6634 requests an additional resource in the next assignment cycle, subscribers with the same priority but smaller ID values (i.e., 6313 and 6422) will drop out again. In the next section, we propose a fair protocol for fast resource assignment. Improvement is achieved with a modification of the RAMA protocol, and a thorough analysis shows this improvement is significant.

3. Random RAMA protocol

In this section we propose a novel method, random RAMA protocol, for fair and fast resource access in future wireless PCS networks. For simplicity, it is assumed that base stations can detect whether more than one subscriber is transmitting orthogonal signals. The detection of multiple users is beyond the scope of this paper. Related work can be found in [1].

Conceptually, the random RAMA protocol can be viewed as a RAMA protocol in which each subscriber has

Base
while any subscriber requests are pending
begin {assignment cycle}
broadcasts the *begin-auction* symbol
repeat {auction cycle}
receives all active orthogonal signals F_i
$F_{\max} = \max F_i$
if more than one requesting subscriber
then feedbacks $F_{\max}$ to the subscribers
else
feedback $[ack]$ to subscribers
until feedback is $[ack]$
resource assignment for the 'winner'
end

Figure 2. Random RAMA protocol for a base.

Subscriber
while additional resource is required
begin
wait for the *begin-auction* symbol
$d \leftarrow P$ {the priority digit}
repeat {auction cycle}
transmit an orthogonal signal F_d to the base
receive a feedback F from the base
if $F = [ack]$
then waits for resource assignment
else if $F_d \neq F$
then drop out
generate a random number d
until $F = [ack]$ or $F_d \neq F$
end

Figure 3. Random RAMA protocol for subscribers.

a *virtual* ID. A virtual ID consists of a priority digit P and a variable number of randomly-generated digits with P as the most significant digit (MSD). The priority digit is used to designate the service priority, and the variable length of the *virtual* ID is used to uniquely identify the winner. Note that the length of virtual subscriber's IDs are variably dependent on the numbers of requesting subscribers (e.g., call originations or handoffs) in a cell and the randomness characteristics of transmitting digits. Each requesting subscriber transmits its virtual ID one digit at a time until it drops out or becomes the winner. The winner is the subscriber with the *longest length of virtual ID* in an auction cycle. After each auction, the base assigns an available communication resource to the winner. Like RAMA, this cycle is then repeated for other available resources until either all requests are satisfied or no more resources are available.

Figure 2 depicts how a base uses the random RAMA protocol. Whenever there are available communication resources and subscribers requesting, the base broadcasts a *begin-auction* symbol to inform the subscribers that a new auction cycle has begun. In an auction cycle, the base listens to all active orthogonal signals. If there is only one subscriber requesting a communication resource, the base feeds back an acknowledgement symbol $[ack]$ to the subscriber, and assigns a resource to the subscriber; otherwise, the base feedbacks the maximum active orthogonal signal to the subscribers. This procedure is repeated until the winner is uniquely identified. Then the base broadcasts a *begin-auction* symbol to start the next auction cycle.

Figure 3 depicts the random RAMA protocol for subscribers. After receiving the *begin-auction* symbol, the requesting subscribers transmit their priority digits in orthogonal signals simultaneously. If the base feeds back an acknowledgement, the auction cycle is completed and the subscriber waits for resource assignment from the base. If the base feeds back any symbol other than its own, the subscriber drops out of further participation in this assignment cycle. The remaining subscribers continue in this auction cycle by transmitting a randomly-generated digit, and

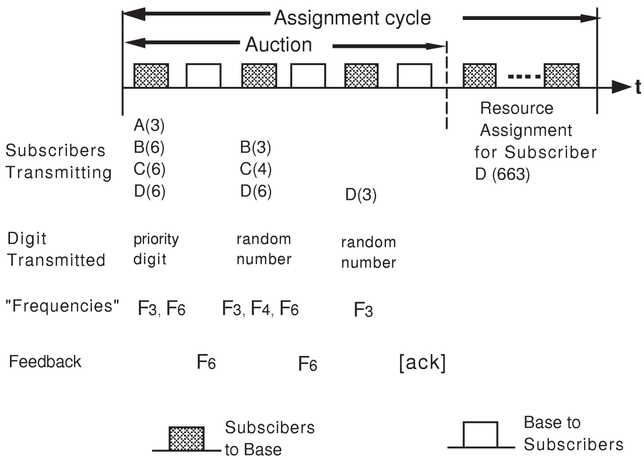


Figure 4. Example of random RAMA auction.

reacting according to the feedback. This transmit-reaction process is repeated until one subscriber becomes the unique winner or drops out of further participation in this assignment cycle. The winner is then assigned an available resource by the base, and those dropout subscribers participate in a new auction during the next assignment cycle.

Figure 4 shows an example of a random RAMA auction. In this example, we assume that each subscriber 'ID' (either the priority digit P or a random number d) is represented using radix 7 notation (i.e., $0 \leqslant P, d \leqslant 6$, the acknowledgement symbol is denoted by 7), and subscribers A, B, C, and D with respective priority digits P = 3, 6, 6, and 6 are seeking for communication resources. In the first auction time slot, all subscribers transmit their priority digits, i.e., subscriber A transmits a '3' by transmitting an F_3 FSK signal, and subscribers B, C, and D transmit '6's by transmitting F_6 FSK signals. The base detects these orthogonal signals, and feeds back the largest digit '6' to the subscribers by transmitting an F_6 FSK signal. After receiving this feedback, all subscribers with priority digits P lower than '6' drop out of the auction, and wait for the next assignment cycle. Subscribers with priority digit $P = 6$ continue by transmitting a randomly-generated number d. In this example, subscribers B, C and D randomly generate and transmit 3, 4, and 6, respectively. After the base announces that '6' was the largest digit transmitted by this group of remaining subscribers, subscribers B and C drop out of the auction, and subscriber D continues by randomly generating and transmitting a '3'. Finally, the base feeds back the acknowledgement symbol $[ack]$ by transmitting, for example, an F_7 FSK signal, and broadcasts a resource assignment to subscriber D, the unique winner of the auction. In the next assignment cycle, subscribers that dropped out of the previous cycle (i.e., subscribers A, B, and C) participate in a new auction along with any new subscribers requesting service.

Unlike RAMA, an auction cycle lasts until a winner is uniquely identified in the random RAMA protocol. A variable-length ID in a random RAMA auction yields lower delay, particularly under light loads. When there is only one requesting subscriber, each auction cycle requires only two slots: one slot is the priority digit from the subscriber to the base, and the other is the acknowledgement from the base to the subscriber. In contrast, the entire fixed-length auction cycle must be completed in the RAMA protocol, wasting valuable slot time. Furthermore, the random RAMA protocol is a fair protocol because subscriber's IDs are randomly generated here.

4. Performance analysis

In this section, we present preliminary performance results for the random RAMA protocol. In our model, performance metrics include mean service time and mean waiting time (access delay). In the real world, the amount of available resource units might affect the waiting time of a customer. However, the wireless bandwidth in future wireless PCS networks is much larger than that in current wireless networks. We assume that the amount of resource units is enough for assignment so that the waiting time in the resource assignment periods is negligible, i.e., the period of resource assignment is short for handoff or initial access. Moreover, there is no difference in the cost of resource assignment and message transmission between RAMA and random RAMA.

For fair comparison with RAMA, we focus on the auction cost and model the waiting time as the time in auction instead of the waiting time until a customer gets service. In the steady state, the number of requesting subscribers is derived in equation (1) and the service time of a requesting subscriber is equal to the ID length. The time for transmitting an M-ary symbol is the time unit in this analysis. Thus mean service time for random RAMA can be considered as the mean ID length. In our study we considered only mobile subscribers, and excluded fixed-network subscribers. With the mobile subscribers, two types of requested resources, call setups and handoffs, were investigated. To simplify our model, we did not distinguish between call setups and handoffs. Furthermore, error-free transmission was assumed.

Suppose that there are N mobiles in a cell on average. Let P_{oc} denote the probability of mobiles originating calls and P_{handoff} denote the probability of mobile handoffs. In the steady state, the number of active mobiles n is the sum of the number of mobiles originating calls and that of handoffs as derived below. The number of mobiles originating calls is equal to $N \cdot P_{\text{oc}}$. A mobile only requests its handoff to a specific cell and the cell will inform the counterpart of the handover via the wireline network. For example, a mobile requests its handoff to the new base station with the mobile-controlled handover scheme [12]. Let each cell have *nc* neighboring cells and the direction of handover to each neighboring cell is uniform for a mobile. The number of handover is equal to $N \cdot nc \cdot (1/nc) \cdot P_{\text{handoff}} = N \cdot P_{\text{handoff}}$. Thus,

$$n = N \cdot (P_{\text{oc}} + P_{\text{handoff}}). \tag{1}$$

In RAMA, each subscriber has a unique ID that consists of 10 decimal digits (a priority digit and a nine-digit phone number). Let M denote the number of orthogonal signals, e.g., M-FSK. Thus the length of the ID (in M-ary symbols) is

$$L = 10 \cdot \log_M 10. \tag{2}$$

The mean waiting time (access delay) $\overline{w_i}$ (in M-ary symbols) for one mobile when there are i active mobiles is equal to the total waiting time of i mobiles divided by i

$$\overline{w_i} = \frac{10 \cdot \log_M 10}{i} \sum_{j=2}^{i} (j - 1). \tag{3}$$

Note that the first mobile gets service in time equal to the ID length, the second gets service in two ID lengths, etc., and that this is the source of the total waiting time.

In random RAMA, let $P(i,j)$ denote the probability that j mobiles transmit the relative maximal signal within i active mobiles, and $\overline{x_i}$ denote the mean service time in M-ary symbols, i.e., the mean ID length, for one mobile when there are i active mobiles. $\overline{x_i}$ is recursively defined as current transmissions plus the time servicing j mobiles that transmit the relevant maximal digit, as shown below:

$$\overline{x_i} = \begin{cases} \sum_{j=1}^{i} P(i,j) \cdot (1 + \overline{x_j}), & 1 < i \leqslant n, \\ 1, & i = 1. \end{cases} \tag{4}$$

Note that for any positive integer i, $\sum_{j=1}^{i} P(i,j) = 1$. Also, $P(i,1) = 0$ for $i > 1$. Then, eliminating the $\overline{x_i}$ term on the right-hand side of the above equation, we obtain the following equation:

$$\overline{x_i} = \begin{cases} \dfrac{1 + \sum_{j=2}^{i-1} \overline{x_j} P(i,j)}{1 - P(i,i)}, & 1 < i \leqslant n, \\ 1, & i = 1. \end{cases} \quad (5)$$

Now we derive the probability $P(i, j)$. Let $\mathcal{D}$ represent the relevant maximal digit of transmitted signals ($0 \leqslant \mathcal{D} \leqslant M - 1$). The probability $P(i, j)$ is equal to the summation of the conditional probabilities of all possible values of $\mathcal{D}$. Imagine the number of ways to place i balls of the same color in M numbered boxes. This is the total number of ways i mobiles can transmit M-FSK signals, and is equal to $\binom{M+i-1}{i}$. The total number of ways that j mobiles can transmit the relevant maximal digit $\mathcal{D}$ within i mobiles is the same as the number of ways $i-j$ balls of the same color can be placed in D numbered boxes. Thus the following equations are derived:

$$P(i,j) = \sum_{k=0}^{M-1} P(i,j \mid \mathcal{D} = k) = \sum_{k=0}^{M-1} \frac{\binom{k+i-j-1}{i-j}}{\binom{M+i-1}{i}}$$

$$= \begin{cases} \sum_{k=1}^{M-1} \dfrac{\binom{M-k+i-j-1}{i-j}}{\binom{M+i-1}{i}}, & i \neq j, \\ \dfrac{M}{\binom{M+i-1}{i}}, & i = j. \end{cases} \quad (6)$$

Similar to RAMA, the mean waiting time $\overline{w_i}$ when there are i active mobiles using random RAMA is computed as follows:

$$\overline{w_i} = \frac{1}{i} \sum_{j=2}^{i} (j-1) \overline{x_i}. \quad (7)$$

Numerical results are shown in figures 5 and 6. Figure 5 depicts a comparison of mean ID lengths between RAMA and random RAMA. It shows that the mean ID length of the random RAMA protocol is much shorter than that of the RAMA protocol. The curves also indicate that mean ID length of random RAMA increases slowly (a log-like function) according to the number of active mobiles. The number of active mobiles can be estimated by using equation (1). In the comparison, we considered the number of active mobiles ranging from 0 to 500. Comparison of mean waiting times (access delays) between RAMA and random RAMA is shown in figure 6. It shows that mean delays are also greatly improved by the random RAMA protocol. Additionally, the mean waiting time is almost proportional to the number of active mobiles, i.e., *linear growth*.

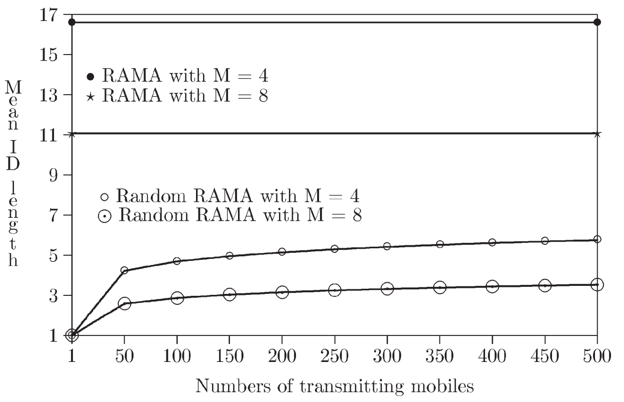


Figure 5. Comparison of mean ID length between RAMA and random RAMA.

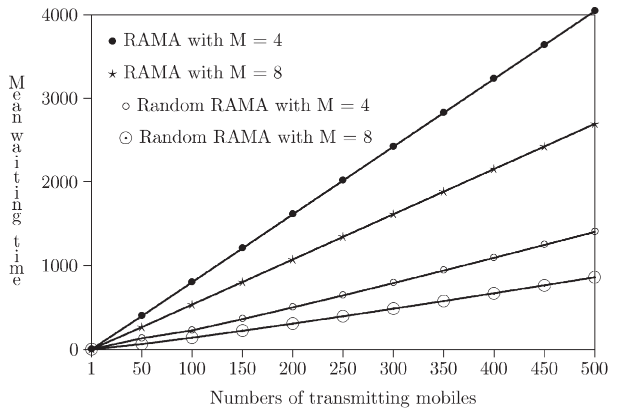


Figure 6. Comparison of mean waiting time between RAMA and random RAMA.

5. Dynamic priority assignment

In the random RAMA protocol, communication resources are fairly assigned to subscribers according to their service priorities. When communication resources (radio bandwidth) are not sufficient to satisfy the requirements of all active subscribers (applications), the ones with lower priorities may not retain services. This forced termination is inconvenient and sometimes unacceptable to subscribers since it is more harmful to the quality of service (QOS) than initial access blocking. Therefore, it is desirable when allocating channel capacities to assign priorities dynamically to various applications or subscribers according to their relative urgency.

Assigning priority according to relative urgency, on the one hand, can alleviate the problem of forced termination for low-priority subscribers; on the other hand, it might degrade the QOS for high-priority subscribers under overload conditions. Fortunately, a graceful degradation of QOS is allowable for voice and multimedia data that can tolerate some loss of information. It is appropriate to transmit these types of data on sub-rating channels [11] or to directly discard some frames. In the following, we present a dynamic priority assignment scheme to improve the quality of service for subscribers under overload conditions.

Table 1
Typical values of the reliability requirements and tolerance.

	Voice	Video	Image	Text
Requirement (r)	0.98	0.90	1.0	1.0
Tolerance ($1-r$)	0.02	0.10	0.0	0.0

Our scheme is based on the fact that each PCS subscriber must define his or her QOS requirement which is an agreement between the PCS subscriber and service provider, in order to obtain services. The QOS information includes service priority, *reliability tolerance, deadline constraints,* and other related parameters. The reliability tolerance indicates the maximum percentage of multimedia data that can be dropped when the wireless bandwidth is insufficient. Typical reliability requirement values and various types of data tolerance are listed in table 1 [14]. The deadline constraint specifies the real-time characteristics of voice and multimedia services. It depends on the types of transmitted data and resources in mobile units such as the sizes of buffers. Note that the larger the buffers, the looser the deadline constraint will be. To satisfy these QOS requirements, each subscriber must measure his or her current reliability tolerance value and *deadline urgency*. The current reliability tolerance value is the percentage of multimedia data that have been dropped and the deadline urgency is a countdown value that represents the remaining time to meet the deadline constraint.

In this dynamic scheme, the assigned priority of an applicant (subscriber) is based on service priority, and can be dynamically adjusted according to other QOS metrics such as reliability tolerance and deadline urgency. Let s denote a subscriber's service priority, and w_r and w_d represent the relative weights of reliability tolerance and deadline urgency, respectively, to service priority. The priority p of an applicant (subscriber) transmitted by the proposed protocol is dynamically assigned as follows:

$$p = s + \lfloor w_r \cdot \Delta r \cdot \delta(\Delta r) + w_d \cdot \delta(d) \rfloor, \qquad (8)$$

where Δr is the difference between current drop percentage and reliability tolerance, d is the value of deadline urgency, $\lfloor X \rfloor$ represents the largest integer less than or equal to X, and δ is a step function:

$$\delta(x) = \begin{cases} \dfrac{1}{d}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases} \qquad (9)$$

Note that the value of deadline urgency d is larger than zero. When $d = 0$, the request will be given up since it fails to meet the deadline. As stated in equation (8), when either the reliability tolerance is violated or the transmission is near deadline, the priority of the subscriber is dynamically increased to prevent forced termination under heavy loading. The level of increment is proportional to the level of reliability tolerance violation and inversely proportional to the value of deadline urgency. Note that although the QOS parameters considered in our scheme are reliability requirements and deadline urgency, it is easy to adopt other QOS parameters.

Time	t_1	t_2	t_3	t_4
Requesting subscribers	A(5) B(4) C(3)	B(4) C(3)	C(3) D(4)	C fails to meet the deadline
Winner	A(5)	B(4)	D(4)	

(a) Static priority assignment

Time	t_1	t_2	t_3	t_4
Requesting subscribers	A(**5**,5,3) B(**4**,4,3) C(**3**,3,3)	B(**5**,4,2) C(**4**,3,2)	C(**5**,3,1) D(**4**,4,3)	D(**5**,4,2)
Winner	A(**5**,5,3)	B(**5**,4,2)	C(**5**,3,1)	

(b) Dynamic priority assignment

Figure 7. Effect of deadline urgency.

The example shown in figure 7 illustrates the proposed scheme. Because the effect of reliability tolerance is similar to that of deadline urgency, only the effect of deadline urgency is concerned for demonstration. In this example, we assume that the deadline constraint for each request is 3, that is, subscribers should be able to gain the required communication resources in 3 auction cycles after requesting. Subscribers A, B, and C request communication resources at time t_1 and subscriber D requests at time t_3. The service priorities of subscribers A, B, C and D are 5, 4, 3 and 4, respectively.

In a static priority assignment scheme (cf. figure 7(a)), subscribers A, B and D become the winners at times t_1, t_2 and t_3, respectively. Even though subscriber D is not an urgent subscriber at time t_3, he or she gets the communication resources immediately regardless of the urgency of subscriber C. Thus subscriber C fails to meet the deadline constraint. This scenario shows an example of the QOS degradation which can be alleviated by a dynamic priority assignment scheme.

In a dynamic priority assignment scheme (cf. figure 7(b)), the deadline urgency of a subscriber is an integral part of his/her deadline constraint. Thus subscriber A with service priority s and deadline urgency d can be represented by a triple $A(p, s, d)$, where p is the assigned priority in equation (8). After each auction cycle, the subscriber either becomes the winner or counts down its deadline urgency until it fails to meet the deadline constraints (i.e., it counts down to 0). In this example, we assume w_r is equal to 0 and w_d is equal to 2. Thus the winners are the same as in the static assignment scheme at time t_1 and t_2, as shown in figure 7. At time t_3, however, subscriber C counts down its deadline urgency to one and thus has an assigned priority of $p = 3 + \lfloor 2 \cdot \delta(1)/1 \rfloor = 5$, while subscriber D has an assigned priority $p = 4 + \lfloor 2 \cdot \delta(3)/3 \rfloor = 4$. Consequently, subscriber C becomes the winner at time t_3, thus meeting his/her deadline constraint. Also subscriber D will get his/her communication resources at time t_4 without violating his/her deadline constraint. This example shows the advantage of dynamic priority assignment for deadline urgency. In a similar manner, we can improve the QOS using reliability tolerance and other QOS parameters in dynamic priority assignments under heavy load conditions.

6. Discussion

The issues of fairness, fault-tolerance, and dynamic priority are all critical to resource assignment in wireless PCS. In this section we discuss these design issues and some practical issues for the random RAMA protocol.

- Fairness.
  We define that an auction-based protocol is *fair* if and only if with the protocol each portable with the same priority has the same ID distribution. As a result, RAMA is unfair since each portable has a fixed, but different ID. Conceptually, with Random RAMA, each portable has a random-generated ID. We assume that the random number generator is fair so that each portable with the same priority has the same ID distribution. Therefore, we conclude that Random RAMA is a fair protocol.
- Fault-tolerance and error-recovery.
  Because wireless networks are prone to error, fault-tolerance and error-recovery are critical when designing a wireless multiple access protocol. Fault-tolerance and error-recovery can be easily be achieved in the random RAMA protocol as follows. In random RAMA, if channel errors lead to exclusion of all subscribers (mobiles) from transmission of their next random IDs, the mobiles will drop out. If channel errors cause multiple mobiles to each think they have received an acknowledgement symbol $[ack]$, then the mobiles will wait for resource assignment from the base. In both cases, the base receives no active orthogonal signals from the mobiles in the subsequent transmission cycle, and thus detects these errors. Consequently, the base broadcasts a *begin-auction* symbol to initiate another auction cycle. Besides, the random RAMA protocol is fault tolerant to other kinds of errors because the subscriber ID's are randomly generated.
- Dynamic priority assignment.
  Static priority assignment is inherent in the RAMA protocol since each subscriber has a fixed ID that designates its priority. In contrast, random RAMA adopts dynamic priority assignment. The basic idea of dynamic priority assignment for random RAMA is to use a priority that is dynamically assigned according to the QOS information to replace the random-generated digit. The advantage of dynamic priority assignment is to improve the quality of service for subscriber, especially for the subscriber that transmits multimedia data under overload conditions.

We then conclude this section with some practical issues of the random RAMA protocol. First, a random number (digit) can be generated by the system time clock. There is no need to embed a real random number generator (RNG) in mobile units. For the sake of resource and power consumption, we use the system clock to replace the RNG even though random and pseudo-random numbers can be generated easily. This advantage is based on the assumption that the system clocks of mobiles are not synchronized in the last n LSD (Least Significant Digit) digits. On the contrary, RNGs are needed if the system clocks are well synchronized. Secondly, the memory requirement is small. Only a one-digit memory is required to store either the priority digit or a random digit. Thirdly, the acknowledgement symbol $[ack]$ can be represented by a reserved M-FSK signal, for example, the F_{M-1} signal. In contrast, it is not necessary to reserve an M-FSK signal for the begin-auction symbol since the begin-auction symbol is not transmitted simultaneously with the feedbacks.

7. Conclusions

To meet the need for fair and fast resource assignment in future wireless applications and services with integrated traffic (e.g., in mobile computing environments), we have proposed an extension of the RAMA protocol. The random RAMA protocol offers fast and fair access to available communication resources using a randomly-generated virtual ID. Fairness is achieved by the randomness of virtual IDs. The variable length of the virtual ID in a random RAMA auction also yields lower delay, particularly under light loading. The features of small resource requirements, inherent fault tolerance, and high performance make the random RAMA protocol attractive for low-cost mobile units. Moreover, the proposed protocol also uses dynamic priority assignment to improve the quality of service for subscribers under overload conditions. Although our dynamic priority assignment scheme is based on service priority, reliability tolerance, and deadline urgency, adopting other QOS parameters would be quite straightforward.

Acknowledgement

The authors would like to thank Dr. J.R. Cruz for his assistance in preparing this paper.

References

[1] D.-H. Alexandra, J. Holtzman and Z. Zvonar, Multiuser detection for CDMA systems, IEEE Personal Communications (April 1995) 46–58.

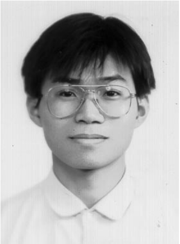
[2] N. Amitay, Resource Auction Multiple Access (RAMA): Efficient method for fast resource assignment in decentralized wireless PCS, Electronics Letters 28(8) (1992) 799–801.

[3] N. Amitay, Distributed switching and control with fast resource assignment/handoff for personal communications systems, IEEE Journal on Selected Areas in Communications 11(6) (1993) 842–849.

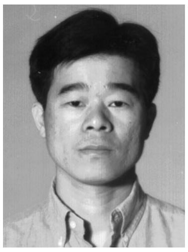
[4] N. Amitay, Resource Auction Multiple Access (RAMA) in the cellular environment, IEEE Transactions on Vehicular Technology 43(4) (1994) 1101–1110.

[5] D.C. Cox, Personal communications – A viewpoint, IEEE Communications Magazine 128(11) (1990).

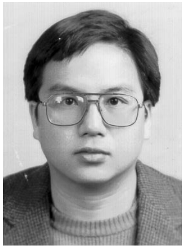
[6] D.C. Cox, Wireless personal communications: What is it? IEEE Personal Communications Magazine 2(2) (1995) 20–35.
[7] G.H. Forman and J. Zahorjan, The challenges of mobile computing, IEEE Computer (April 1994) 38–47.
[8] R. Ghai and S. Singh, An architecture and communication protocol for picocellular networks, IEEE Personal Communication (Third Quarter 1994) 36–46.
[9] D.J. Goodman et al., Packet reservation multiple access for local wireless communications, IEEE Transactions on Communications 37(8) (1989) 885–890.
[10] M.J. Karol and C.H. I, A protocol for fast resource assignment in wireless PCS, IEEE Transactions on Vehicular Technology 43(3) (1994) 727–732.
[11] Y.B. Lin, S. Mohan and A. Noerpel, PCS channel assignment strategies for hand-off and initial access, IEEE Personal Communications (Third Quarter 1994) 2–11.
[12] A. Noerpel and Y.B. Lin, Handover management for a PCS network, submitted for publication (1995).
[13] A.S. Tanenbaum, *Computer Networks*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[14] M. Woo, N. Prabhu and A. Ghafoor, Dynamic resource allocation for multimedia services in mobile communications environments, IEEE Journal on Selected Areas in Communications 13(5) (1995) 913–922.



Tsan-Pin Wang received the B.S. degree in applied mathematics and the M.S. degree in computer science and information engineering, both from National Chiao Tung University, Taiwan, ROC, in 1990 and 1992, respectively. From 1992 to 1993, he was a system engineer in the R&D Division of Taiwan NEC Ltd. He is currently working toward his Ph.D. degree at National Chiao Tung University. His research interests include PCS networks and mobile computing.
E-mail: tpwang@csie.nctu.edu.tw



Chien-Chao Tseng is currently a professor in the Department of Computer Science and Information Engineering at National Chiao Tung University, Hsinchu, Taiwan. He received the B.S. degree in industrial engineering from National Tsing-Hua University, Hsinchu, Taiwan, in 1981, and M.S. and Ph.D. in computer science from the Southern Methodist University, Dallas, Texas, USA, in 1986 and 1989, respectively. His research interests include mobile computing and parallel and distributed processing.
E-mail: cctseng@csie.nctu.edu.tw



Shu-Yuen Hwang was a professor in the Department of Computer Science and Information Engineering, National Chiao Tung University. He received the B.S. and M.S. degrees in electrical engineering from National Taiwan University in 1981 and 1983, respectively, and the Ph.D. degree in computer science from the University of Washington in 1989. His current research interests include artificial intelligence, computer simulation and mobile computing.