A Multicast Search Scheme Based on Bipartite Graph Matching Model

Yuan Yao, Zhang Dalin, Wang Qing and Shi Jinglin

Beijing Key Laboratory of Mobile Computing and Pervasive Device; Institute of Computing Technology, Chinese Academy of Sciences zhangdalin@ict.ac.cn

Abstract

This paper proposes an efficient multicast search scheme based on bipartite graph matching model, aiming at one of the most important problems in the multicast service, how can the wireless network track and locate the mobile uses in the idle state, under the tight bandwidth and delay constraints,. By quantifying the location uncertainty of mobile users with Shannon's entropy, the scheme adopts the LZ78 compression algorithm to design location update scheme and predict the location probability, in order to reduce the cost of location update. The multicast search system need to guarantee the maximum total probability that each user resides at the assigned paging area, for the purpose of an optimal performance on both expected paging delay and paging cost. Therefore, the bigraph-based multicast paging scheme (BMPS) firstly builds the bipartite graph model of multicast search problem, through converting the location probability into weight of the edge. BMPS decides the optimal allocation between the mobile users and location areas, which is mainly achieved by the maximum weight perfect matching in bipartite graph, while modifying the weights dynamically. Simulation results show that BMPS yields a low search delay as well as a low search cost, and reduces the impact of user collision.

Keywords: multicast search; bipartite graph matching; location probability prediction; paging delay; paging cost

1. Introduction

The future mobile communication system [1] will provide accessing network services in anytime for mobile users with different status. Mobility management is an important guarantee to achieve this goal. However, the current mobility management mainly studies on the switching mechanism when the user is in the connection status; doesn't put much attention on how to protect the signaling overhead and the delay requirement of the network access when user is in idle status. With the extensive development of various multicast applications service [2, 3] (such as conference calls, *etc.*), in the wireless network, to realize the requirement of tracking and positioning a group mobile users (mobile station, MS) in idle status is higher and higher. Therefore, the mobility management is an urgent problem to be solved in the mobile wireless network when users are in idle status in multicast service.

For the mobility management problem when the MS is in an idle status, the mobile communication network uses the location management mechanism to divide the entire network coverage area into a plurality of location areas (location area, LA) for tracking and locating mobile users. Location management has two basic operations: location update and paging. Location update is a process that mobile users periodically report the current position information to the network and network updates the relevant location database; paging is a process that when a call arrives, the network searches mobile users according the location information recorded in the location update. For a single user,

ISSN: 2005-4254 IJSIP Copyright © 2015 SERSC paging strategy can always achieve a successful search the mobile users within a specified delay through dynamic programming and can ensure the number of paging location area is the minimum in the entire network coverage. However, for multicast search problem, Reference [4] proved that in the case of limited bandwidth delay, minimizing paging delay and overhead is a hard NP problem.

In recent years, in order to improve the quality of paging, experts have done a series extension study on the traditional paging theme, and have put forward a series of solutions from different point. However, these solutions when applied in the multicast search still have some significant defects.

Location update mechanism: Dynamic location update mechanism can be divided into three types, which are based on the distance, based on the movement and based on the time. By comparing these three update mechanisms in Reference [5], it displayed that the update mechanism based on the distance has the optimal performance; however, its performance depends on the accurate measurement and the effective threshold distance set. Reference [6] presented a statistical analysis method to establish the location update model based on the movement, discussed the optimization of the movement model threshold and effectively improved the accuracy of location probability prediction. Reference [7] proposed a dynamic programming based on the Partially Observable Markov process and designed the joint optimization strategies and the iterative algorithm of paging and registration.

Location probability prediction: Many studies on paging algorithms are dependent on assuming the location probability distribution of mobile subscriber is known thus it can't reflect the user's actual movement model. To solve this problem, references [8-11] proposed a variety of location prediction models, such as: the model based on dynamic Bayesian network, the model of multilayer perceptron prediction, the model of Markov and so on. Reference [12] compared the performance of these types of position prediction algorithms on the accuracy, the stability, the computational complexity and the modeling overhead, respectively. Reference [13] proposed a rule-based paging strategy (RBPS), aimed to use the artificial intelligence in the rule base to determine the probability distribution of the mobile terminal location. However, RBPS requires the paging system understanding the entire network topology and has assumed user movement model and call arrival model.

Paging system: Reference [14] designed a non-blocking pipelined probability paging system (PPP) for a plurality of paging request (PR). This paging mechanism is superior to the sequential probability paging and the blanket paging used by GSM. However, the PPP paging system is established on assuming the location probability distribution is known; when parallel executing a plurality of paging requests, it doesn't fully utilize location probability distribution of the mobile users, thus its paging overhead and delay cannot achieve optimal performance.

Multicast search algorithm: Reference [15] designed an approximate optimal strategy of a polynomial time for multicast search problem, however, the algorithm only applied to the condition that the largest paging delay and the paging user number are fix. Reference [16] proposed a Sorted Round Robin (SRR) strategy. SRR algorithm considered the fairness among multiple users for multicast paging features and achieved an optimizing balance between the paging delay and the overhead. However, SRR algorithm cannot effectively solve the conflict problem caused by the limited paging bandwidth when multiple users reside in a certain area.

According to the above analysis, the deficiencies of the current paging strategy exist in the following aspects when applied in the multicast search: 1) The designed multicast search theme assumed that the user location or area topology was known, and the established user movement model didn't reflect the various motion of the actual motion and the direction change, so it affected the accuracy of the position probability prediction.

2) The designed multicast search system failed to make full use of the position probability

distribution of mobile users, thus it affected the performance of multicast search delay and overhead. 3) The designed multicast search algorithm didn't effectively solve the assigning area conflict problem for users when performing paging process on the area which had been assigned the limited bandwidth for mobile users.

To solve the above problems, this paper proposes a multicast search theme. First, the construct of the theme does not rely on any assumed motion model and introduces the entropy in information theory to analyze the uncertainty of user location. In order to reduce the location update overhead, the theme uses LZ78 compression algorithm to realize the location update and location probability prediction. Then this paper proposes a multicast search system model, its goal is having the maximum total of residence probability of assigned paging area for all mobile users in each paging cycle, and meeting the bandwidth limitations and user fairness. In order to achieve the system objectives, this paper combines with the multicast search features and proposes the BMPS multicast search algorithm based on bipartite graph matching. The main idea of this algorithm is to ensure the fairness of the user, to obtain the maximum bipartite graph perfect matching, to dynamically adapt to modify the weights according to the current paging results, then achieves optimal assignment scheme between the user and the paging area. The multicast search algorithm based on bipartite graph matching can effectively achieve the overall performance optimization of the paging overhead and delay.

The first section introduces the location prediction mechanism based on information theory. The second section introduces the multicast search system model. The third section details the multicast search algorithm based on bipartite graph matching (BMPS algorithm) and proves the correctness of the algorithm. The fourth section simulates the performance of the algorithm. The fifth section summarizes the full text.

2. Location Prediction Mechanism

2.1. Location Update

Because the movement model affects the analysis of the whole location update theme and the accuracy of user location probability prediction, therefore, in order to make the location update theme can reflect the various motions and the direction change in the actual movement, the constructed model in this paper doesn't make any assumptions about the network topology and user moving.

Definition 1: The movement model of users is a stationary random process $\{v_i\}$ ($v_i \in V$). v is the area set, v_i represents the user is locating in the area v_i for the i times location update.

In the location update, MS changes movement paths into character sequences and sends to the network. Therefore, this paper introduces the entropy concept in the information theory as mentioned in reference [17] to analyze the location uncertainty of users. The location uncertainty of mobile users can be expressed as:

$$H(V) = -\sum_{v_i \in V} p(v_i) \lg p(v_i)$$
 (1)

According to the definition of the stationary source ultimate entropy

$$H_{\infty} = \lim_{N \to \infty} H_{N}(V) = \lim_{N \to \infty} H(v_{i} | v_{1}v_{2}...v_{N-1})$$
 (2)

the average symbol entropy is

$$H_{N}(V) = \frac{1}{N} H(v_{1}v_{2}...v_{N}) = -\frac{1}{N} \sum_{i_{1},i_{2},...,i_{N}} p(v_{i_{1}}v_{i_{2}}...v_{i_{N}}) \log p(v_{i_{1}}v_{i_{2}}...v_{i_{N}})$$
(3)

conditional entropy is

$$H(v_{i} | v_{1}v_{2}...v_{N-1}) = -\sum_{i_{1},i_{2},...i_{N}} p(v_{i_{1}}v_{i_{2}}...v_{i_{N}}) \log p(v_{i_{N}} | v_{i_{1}}v_{i_{2}}...v_{i_{N-1}})$$

$$\tag{4}$$

MS reports the movement path, that is, the character sequence to the network when the location updates; the correlation between v_i and v_j leads the limit entropy \mathbb{H}_{∞} , that is, the MS location uncertainty decrease when the character number N increases. Therefore, when MS reports the movement path to the network, reducing the location update overhead, and does not affect the MS position probability prediction accuracy is the key problem in the location update mechanism design.

Reference [17] has been proved: in order to achieve the exact location prediction, the average location update overhead cannot be less than its limit entropy. Therefore, in order to reduce the location update overhead, we should design an update strategy to gradually approach the ultimate entropy.

Because the source entropy is the limit of the average code length of the LZ78 code, so the location update can use compression algorithm based on LZ78. By converting the movement path of MS into a character sequence $v_1v_2...v_N$ and the online adaptive learning to analyze the movement model, the LZ78 compression algorithm realizes the location tracking of user location. The segmentation rule used by LZ78 compression coding is taking less attached source symbols to ensure all the segments are not the same. The location update algorithm based on LZ78 compression coding in the MS and the network system are shown in Figure 1 and Figure 2, respectively.

```
initialize dictionary := null; phrasew :=null loop

wait for next symbol v

if (w.v in dictionary) w :=w.v

else encode <index(w),v>

add w.v to dictionary

w := null

endif

forever
```

```
initialize dictionary := null
loop
wait for next codeword <i,s>
decode phrase :=dictionary [i].s
add phrase to dictionary
increment frequency for every
prefix of every suffix of phrase
forever
```

Figure 1. Encoder at the Mobile

Figure 2. Decoder at the System

2.2. Location Probability Prediction

According to the location update theme [18, 19], the network paging controller (PC) uses the structure of LZ compression model trie to record the MS location information. In the process of implement location update, assuming MS movement path is aaababbbbbaabccddcbaaaa, Figure 3 shows the LZ compression model trie of the network terminal, which is structured by the symbol-wise according to the algorithm as shown in Figure 1 and Figure 2.

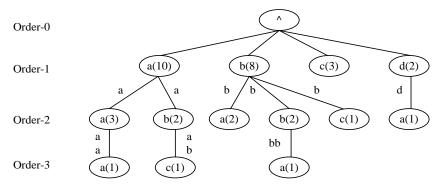


Figure 3. Trie for LZ Symbol-wise Model

MS location probability distribution in each area can use part prediction matching algorithm (PPM) for predicting. Supposing ε indicates the last location update context (that is the movement path), N(w) represents the frequency of the context w in the LZ compression model trie, L(w) represents the length of the context w, $S^k(w)$ represents the k-th suffix of the context w, P(w) represents the prefix of w, the probability of context ψ is calculated as:

$$\Pr(\psi) = \frac{N(\psi \mid P(S^k(\varepsilon)))}{\sum_{i=1}^{k} N(w \mid P(S^k(\varepsilon)))} + \frac{N(\Lambda \mid P(S^k(\varepsilon)))}{\sum_{i=1}^{k} N(w \mid P(S^k(\varepsilon)))} \times \Pr[S^1(P(S^k(\varepsilon)))] \quad (1 \le k \le L(w))$$
(5)

The algorithm process of position prediction which uses the PPM algorithm to calculate the probability of each context is shown as follows:

Step1: Initialization: i:=0, $Pr[\psi]$:=0, h=highest order; the jump probability: $Pr_h^{(e)}$:=1;

Step2: Searching ψ in the order h-i layer of the LZ compression model trie;

Step3: If ψ is found, then the probability $Pr_{h\text{-}i}[\psi]$ of ψ in the order h-i layer will be calculated; otherwise $Pr_{h\text{-}i}[\psi]=0$;

Step4: Calculating the probability Pr_{h-i} jumped to the order h-i layer;

Step5: Calculating the mixed probability: $Pr[\psi] := Pr[\psi] + \prod_{i=h}^{h-i} Pr_i^{(e)} * Pr_{h-i}[\psi];$

Step6: i:=i+1, when $i \le h$, turn to step2, otherwise, the algorithm ends.

The probability of a single character can be calculated by using its weight in each context, the formula is:

$$\rho(v_i) = \sum_{w} \frac{N(v_i)}{L(w)} \times \Pr(\psi) \qquad \left(v_i \in V\right)$$
 (6)

According to the PPM algorithm and the formula (6), the network can use the location update message to calculate the MS probability distribution of each area, so in the next paging, the network can decide the optimization distribution scheme between MS and the area according to the location probability.

3. Multicast Paging System

3.1. System Model Description

Problem Definition: In multicast paging, assuming mobile user set is $U = \{U_1, U_2, U_3, ..., U_M\}$, area set is $U = \{U_1, U_2, U_3, ..., U_M\}$, area set is $U = \{U_1, U_2, U_3, ..., U_M\}$, where the available bandwidth in each paging cycle is B. The user's location probability distribution

is $P = \{p_{ij}\} (1 \le i \le M, 1 \le j \le N)$, p_{ij} represents the stay probability of user U_i in area C_j and satisfies $\sum_{C_j \in C} P_{ij} = 1 (1 \le i \le M)$. In N areas, through no more than D paging cycles, paging

system can page M mobile users and make the expectation paging overhead and delay achieve the overall performance optimization.

Whether there is a feasible solution for the multicast paging problem, reference [4] has proved: assuming when D=d, multicast paging system always can page all the M mobile users in the maximum delay; and when D=d-1, there doesn't exist a paging strategy which can page all the users in the maximum delay, then B satisfies the following relation:

$$B = \left\lceil \frac{M}{D} \right\rceil \tag{7}$$

The above theorem shows that when $D \ge \left\lceil \frac{M}{B} \right\rceil$, there must exist a paging strategy $\{s^1, s^2 ... s^3 ... s^{p^*}\}$ making multicast paging system page all the users in the prescribed

Multicast paging system model is shown in Figure 4. To reduce the paging overhead and delay, multicast paging system need consider the mobile user probability distribution, available bandwidth of each area, paging delay and other factors, then implements the multicast paging strategy and dynamically distributes the N paging areas to the m $(m \le M)$ users who haven't successfully been searched yet in each paging cycle.

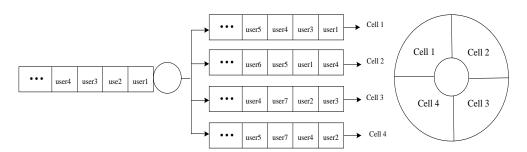


Figure 4. Multicast Paging System

Definition 2: Multicast paging strategy can be expressed as $\{s^1, s^2 ... s^3 ... s^{p^k}\}$, s^d indicates the paging decision matrix in the d-th paging cycle; only when multicast paging system pages the user U_i in area C_j during the d-th paging cycle, $s_{ij}^d = 1$; otherwise $s_{ij}^d = 0$ $(1 \le i \le M, 1 \le j \le N)$.

The definition of multicast paging delay D^* is successfully paging the last user in the D^* paging cycle.

Definition 3: Multicast paging overhead is the total number of multicast paging system paging M users in N areas during the whole paging process that is the D paging cycles.

$$Cost = \sum_{1 \le d \le D^*} \sum_{1 \le I \le M} S_{ij}^{\ d}$$
 (8)

To achieve the approximate optimization of D^* and Cost, the target of multicast paging system is:

maximum delay.

Target 1: Obtaining a multicast paging strategy $\{s^1, s^2 ... s^3 ... s^{p^k}\}$, which can make $\sum_{1 \le i \le M} s_{ij}^{-d} p_{ij}$ maximize in each paging cycle and satisfy the following constraints:

Constraint 1: $\forall d (1 \leq d \leq D^*), C_j \in C, \sum_{U_j \in U} S_{ij}^d \leq B$;

Constraint 2: if $s_{ii}^{d} = 1$, then $\forall d (d \neq d, 1 \leq d \leq D^{*})$, otherwise $s_{ii}^{d} = 0$.

Constraint 1 shows that in each paging cycle, the paging user number which will be distributed by each paging area does not exceed the available paging channel number B. Constraint 2 shows in different paging cycle, the user cannot execute the paging procedure in the same paging area.

In addition, for the features of the multicast paging, multicast paging delay p depends on the paging cycle numbers of successfully paging the last user. So, multicast paging algorithm should consider the user fairness, that is, as far as possible to ensure the paging area number for distributing to each user is equal.

Definition 4: Paging user fairness index:

$$F = \frac{\left(\sum_{i=1}^{M} B_{i}\right)^{2}}{M * \left(\sum_{i=1}^{M} B_{i}^{2}\right)}$$
 (9)

 $B_i = \sum_{d=1}^{D^*} \sum_{j=1}^{N} S_{ij}^d$ represents the assigned paging numbers by the i-th user during the whole paging process.

It is not difficult to prove that, for F = (0, 1), the bigger F is, the better the system fairness is; Specially when $B_1 = B_2 = ...B_M$, F = 1. In this paper, we set F = (0.9,1), and the algorithm satisfies the fairness condition.

3.2. Multicast Paging Strategy Analysis

Considering the multicast paging problem of M = 12, N = 4, B = 2, D = 6, that is, under the condition that there are two available paging channels in each paging area and the maximum paging delay is 6 paging cycles, multicast paging system need page 12 mobile users in 4 paging areas. Within each paging cycle, multicast paging strategy need to choose an allocation scheme between the user and the paging channel, ensure the scheme satisfy the constraints 1 and 2 and make $\sum_{1 \le i \le M} s_{ij}^{\ \ \ \ \ \ } P_{ij}$ maximized.

PA1 1 3 PA1 4 6 PA1 5 7 PA1 9 11 PA1 10 PA2 4 1 PA2 5 6 PA2 7 8 PA2 12 9 10 PA2 PA3 3 2 PA3 7 4 8 5 PA3 10 PA₃ 6 PA3 2 7 PA4 4 PA4 5 PA4 6 8 PA4 10 11 PA4 Paging cycle 1 Paging cycle 2 Paging cycle 3 Paging cycle 4 Paging cycle 5 Searched user: 1,2,3 Searched user: 4 Searched user: 5,7,8 Searched user: 6,9,11,12 Searched user: 10

(a) Multicast Paging in each Paging Cycle

pmf	PA1	PA2	PA3	PA4											
User1 User2	0.2836	0.5333	0.1468 0.6584	0.0363	PA1	1	3	4	6	5	7	9	11	10	
User3 User4	0.7698 0.3223	0.0866 0.3500	0.1353 0.2965	0.0083 0.0312	PA2	4	1	5	6	7	8	12	9	10	
User5 User6	0.0180 0.1235	0.2300 0.6145	0.2260 0.2453	0.5260 0.0167	PA3	3	2	7	4	8	5	6	10		
User7 User8	0.1057 0.1305	0.3038 0.2858	0.3224 0.5645	0.2681 0.0192	PA4	2	4	7	5	6	8	10	11		
User9 User10		0.4409 0.2128	0.2064 0.3673	0.0952			1		2	1	3		4		5
User11 User12	0.3530 0.1091	0.2934 0.4717	0.2100 0.2137	0.1436 0.2055											

(b) MSs' Location Probability Distributions

(c) allocation between MS and LA

Figure 5. Example Operations in the Multicast Paging Scheme

Figure 5 (a) proposes a schematic diagram; the schematic diagram indicates the paging process of multicast paging strategy which is based on the user's location probability distributions as shown in Figure 5 (b). In the first paging cycle, the multicast paging system uses the two paging channels of paging area PA1 to page user v_1 and user v_2 and user v_3 and user v_4 in area PA2, pages user v_3 and user v_4 in area PA4, respectively. From the probability distribution as shown in Fig.5(b) we can know, as far as possible, the allocation scheme chooses the area with larger location probability page for the user, for example, the theme distributes the most likely areas PA2 and PA1 to page for user v_4 . However, due to the bandwidth limitation of the area, user's conflict will occur inevitably. Therefore, in order to solve the conflict, some users can only select suboptimal probability area for paging. Such as paging strategy assigns paging area PA4 for user v_4 , but user v_4 has the smallest probability in area PA4.

After a round of paging, paging system has been successfully paged user v_1, v_2, v_3 . In addition, according to the first round of paging results, the stay possibility of v_4 in the un-researched area increases. In the next cycle, system will consider how to assign the 8 channels of the 4 areas to user U_4 and U_{12} who hasn't been paged yet. After four rounds of paging, only user v_{10} hasn't been paged yet, and v_{10} doesn't been paged in paging area PA1 and PA2; then in the fifth paging cycle, system assign paging area PA1 and PA2 to user v_{10} for paging. So far, all the twelve users have been successfully paged, multicast paging process ends. The multicast paging overhead that is the paging number of paging in the 4 areas is 34, paging delay is 5 paging cycles.

Based on the analysis of multicast paging strategy, multicast paging algorithm design can be considered from the following four aspects. First, in each paging cycle, according to the user location probability, assigning N areas to M users for paging and making the total probability of user stay in the assigned area be maximum; Second, in each paging cycle, the paging user number cannot exceed the available bandwidth limit of each paging area; Third, avoiding bandwidth waste, that is in the whole paging process, avoiding repeatedly assigning one same area to the user for paging; Fourth, user fairness, that is, in the whole paging process, as far as possible assigning the same paging area number to each user.

4. Multicast Paging Scheme Based on Bipartite Graph Matching

4.1. Definition of Bipartite Graph Model

The essence of multicast paging problem is in the condition of the bandwidth delay limited, how to assign M users into the N*B paging channels of N paging areas, making the total location probability $\sum_{1 \le i \le M} s_{ij}^{\ d} p_{ij}$ of paging decision be maximum in each paging cycle and satisfying both constraint 1: $\forall d \ (1 \le d \le D^*), C_j \in C, \sum_{U_i \in U} s_{ij}^{\ d} \le B$ and constraint 2: if $s_{ij}^{\ d} = 1$ then

 $\forall d'(d' \neq d, 1 \leq d' \leq D^*)$, $S_{ii}^{d'} = 0$ at the same time.

According to the definition of bipartite graph maximum weight perfect matching, it can convert multicast paging into the problem of obtaining the bipartite graph maximum weight perfect matching.

Definition 5: The complete bipartite graph model of multicast paging algorithm is defined as follows: points set $U_{\{U_1,U_2,U_3,...U_M\}}$ represents a collection of mobile users, points set $C_{\{C_1,C_2,C_3,...C_N\}}$ represents a collection of areas. $\forall u_i,c_j$, the weight w_{ij} of edge (u_i,c_j) represents the probability of staying in area c_j for user u_i and satisfies $\sum_{C_j \in C} w_{ij} = \mathbb{I}(\forall U_i \in U_j)$. In addition, in order to obtain the bipartite graph

maximum weight perfect matching, it need to add necessary virtual nodes in the bipartite graph and make them satisfy U = |C|.

The bipartite graph model of multicast paging is shown in Figure 6 (a). Assuming N < M, when building the bipartite graph matching model, we add area virtual nodes $\{C_{N+1}...C_M\}$ and set the weight of edge $(U_i,C_j)(\forall C_j \in \{C_{N+1}...C_M\})$ be $W_{ij} = 0$.

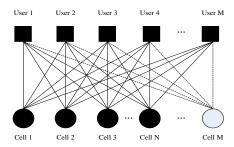


Figure 6(a). The Bipartite Graph with Weight

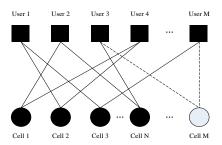


Figure 6(b). The Optimal Allocation between MS and LA

The result that multicast paging algorithm BMPS based on bipartite graph matching finally assigns paging area for the user in one paging cycle is shown in Figure 6(b). The result is gotten through the multicast paging algorithm which is

based on the bipartite graph matching, and BMPS repeatedly obtains the bipartite graph maximum weight perfect matching. From the figure we can get the paging decision of BMPS in this paging cycle, for example, user U_1 assigns area C_2 , C_3 for paging and user U_3 assigns area C_N for paging; although in the matching process, we match U_3 and C_M , but the virtual nodes $\{c_{N+1}...c_M\}$ don't participate the paging in the actual paging process.

4.2. Multicast Paging Scheme Based on Bipartite Graph Matching

- **4.2.1. Problem Description:** First, analyzing the features of the multicast paging problem:
- 1) In a paging cycle, when the available paging channels B > 1 in each paging area, in order to reduce paging delay, paging system will assign the same area C_k for a plurality of users $U_k = U_k = 0$, and assign a plurality of areas $C_k = 0$.
- 2) In different paging cycles, multicast paging algorithm can dynamically adjust the user location probability distribution by making full use of the current paging result. It is mainly divided into the following two situations:
- i) If in the d-th paging cycle, user u_i has been paged in area c_j , then after the d-th paging cycle, multicast paging algorithm doesn't need assign area for user u_i ;
- ii) If it has paged user v_i from area C_i to area C_j during the first paging cycle to the d-th paging cycle, but user v_i hasn't been paged yet; then after the d-th paging cycle, when multicast paging algorithm assigns the area for user according by the location probability, the location probability for user v_i staying in other area will increase.

Through the above analysis, the main problems for the multicast paging algorithm based on bipartite graph matching model to solve are:

- 1) How to convert the many-many mapping of multicast paging to one-one mapping of bipartite graph matching between user set U and area set C.
- 2) In the *d*-th paging cycle, based on the paging results of the first paging cycle to the (d-1)-th paging cycle, how to dynamically online modify the edge weight $e(U_i, C_j)$ of the bipartite graph to optimize the maximum weight matching between the user and the area.
- **4.2.2. Algorithm Description:** According to the analysis of the algorithm design from the view of multicast paging system as mentioned in Section 2, the process of multicast paging algorithm based on the bipartite graph matching is shown in Figure 7. The core idea of the BMPS algorithm is that first building the bipartite graph model and converting the location probability to the weight. In each paging cycle, by repeatedly obtaining the bipartite graph maximum weight perfect matching and at the same time combining user fairness, the BMPS algorithm achieves the optimal allocation scheme between the user and the paging area; then at the end of each round of the paging cycle, according to the current paging result, dynamically modifies the weight and makes it to be the bipartite graph model for the next round of paging cycle.

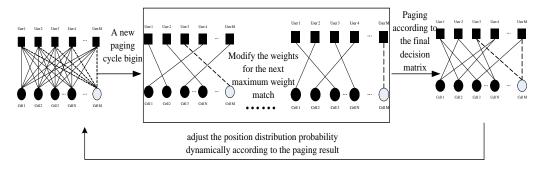


Figure 7. Multicast Search Scheme based on Bipartite Graph Matching Model

1) Optimal Probability:

In order to ensure that in each paging cycle, the total of the position probability $\sum_{1 \le i \le M} s_{ij}^{\ d} p_{ij}$ of the multicast paging algorithm based on the bipartite graph matching is the maximum, it can be achieved by the improved weighted bipartite

matching is the maximum, it can be achieved by the improved weighted bipartite graph maximum weight perfect matching Kuhn-Munkres algorithm.

Because the Kuhn-Munkres algorithm obtains the maximum weight perfect matching and it is a one-one mapping, therefore, first it need solve the problem 1 of the mentioned multicast paging algorithm: how to convert the many-many mapping of multicast paging to one-one mapping of bipartite graph matching between user set U and area set C.

Figure 7 indicates that, through repeatedly obtaining the maximum weight matching $\{M_1, M_2, ..., M_b\}$, the multicast paging algorithm based on bipartite graph matching gains the optimum paging decision $s^d = \sum_{1 \le i \le b} M_i * sgn(W)$ and the maximized $\sum_{1 \le i \le M} s_{ij}^d p_{ij}$. Before obtaining the (k+1)-th maximum weight matching, it need modify the weight according to the k-th matching result. If $e(U_i, C_j) \in M_k$, when obtaining M_{k+1} , it sets weight $M_{ij} = 0$, that is, $e(U_i, C_j)$ exits next matching. This can avoid assigning the same area C_i for the user U_i to waste the paging bandwidth.

2) Collision Avoidance:

Since the available bandwidth of each paging area is B, then in a paging cycle, when multicast paging strategy assigns the number of users for the area c_j over B, it will cause user conflict.

Before the multicast paging algorithm based on bipartite graph matching obtaining the maximum weight matching, if $\sum_{U_i \in U} s_{ij}^{\ \ d} = B$, we set the value of the

bipartite graph weight be $w_{ij} = 0 (\forall U_i \in U)$ which can ensure the area without available assigned bandwidth exit next maximum weight matching.

3) User Fairness:

In order to reduce the paging delay, that is to reduce the total paging cycles for the last user being paged, multicast paging system tries to assign the same number areas for each user.

In the *i*-th round of the paging cycle, the expected paging number for each user is:

$$n_{NP} = \frac{N_i}{d}. \tag{10}$$

 N_i represents the area need be paged in the i-th paging cycle. If before the (i-1)-th paging cycle, there already has paged the total users of M in area C_i , then area C_i is no longer to do paging. d_i represents the left paging cycles, $d_i = D - i + 1$.

In the *i*-th round of the paging cycle, the average paging channel assigned for each user is:

$$n_{BP} = \frac{N_i * B}{M_i} \tag{11}$$

M, indicates that hasn't paged the user yet.

Definition 6: in the i-th paging cycle, the average channel assigned by multicast paging algorithm for the unpaged users is:

$$Bu = \min(n_{NP}, n_{BP}) = \min(\frac{N_i}{d_i}, \frac{N_i * B}{M_i})$$
 (12)

4) Bandwidth Saving:

In a paging cycle, after one perfect matching, it will update the corresponding weight matrix. If $e(U_i, C_j) \in M_k$, when obtaining M_{k+1} , we set weight $W_{ij} = 0$ avoid assigning the same area C_i to user U_i .

In different paging cycle, if in the d-th round of the paging cycle, the paging decision is $s_{ij}^{d} = 1$, then according whether user u_{ij} is paged in the area c_{ij} in the d-th round of the paging cycle, the algorithm dynamically modifies the weight.

Case 1: user U_i has been paged in area C_j , then in the (d+1)-th round of the paging cycle, the algorithm modifies $P_{ij} = O(\forall C_j \in C)$ to make user U_i exit paging allocation.

Case 2: user u_i hasn't been paged in area c_j , then in the (d+1)-th round of the paging cycle, the algorithm modifies $P_{ij} = 0$. In addition, the improved Kuhn-Munkres algorithm utilizes the feasible numeral to obtain equal subgraph and guarantees compared to other users, the probability of user u_i staying in the left area increase, then to obtain the maximum weight perfect matching.

Algorithm 1: multicast paging algorithm based on bipartite graph matching

Input: in multicast paging, the number of mobile users is M, the number of areas is N, the maximum paging delay is D, the available paging channel of each paging area is B, the probability of user Ui staying in area C_j is $P=\{P_{ij}\}$, the actually staying area of user is $L=\{L_{ij}\}$, $L=\{L_{ij}\}$ represents user Ui is in area C_j .

Output: paging delay, paging cost

In the following steps, R represents the multicast paging result, S represents the paging decision matrix, S_{ij} =1 represents user Ui will be paged in area C_j , MA represent the maximum weight perfect matching gained by the Kuhn-Munkres algorithm.

Step 1: Initialization: R=0, d=D; delay=0; cost=0;

Step 2: A new round paging starts. Initialization: W=P, S=0, MA=0, $Bu=min(N_i/d_i, N_i*B/M_i)$; delay++;

Step 3: Use the improved Kuhn-Munkres algorithm to obtain the maximum weight perfect matching MA;

- Step 4: Set paging decision matrix: from the first user, i=1;
- Step 5: If MA(i)=j, $W(i,j)\neq 0$ and $R_i=0$, then set S(i,j)=1, that is, in the current paging cycle, it will page user Ui in area C_i .
 - Step 6: Consider the next user, i++, if i<M, then turn to step 5.
 - Step 7: modify weight matrix: for 1 < j <= N, If $\sum S(i,j) = B$, W(i,j) = 0 (i = 1,...,M);
- Step 8: Bu=Bu-1;If Bu>0, turn to step 3. Otherwise, according paging decision matrix to page, comparing S and actual location matrix L, update paging result is R, update paging cost is cost=cost+ $\sum_{1 \le i \le M} \sum_{1 \le i \le N} s_{ij}$;
- Step 9: By the paging result, the update location probability matrix is: for 1 < i < M, If R(i)=1, P(i,j)=0 (j=1,...,N); then, when S(i,j)=1, P(i,j)=0.
- Step 10: d=d-1; if d>0 and there exists I, $R_i=0$, that is, there are users who haven't been paged yet, then turn to step 2; otherwise, the algorithm ends.

4.2.3. Algorithm Correctness and Complexity Analysis:

Proposition 1: In the condition of satisfying both constraint 1 and constraint 2, the multicast paging algorithm based on bipartite graph matching ensure that $\sum_{1 \le i \le M} s_{ij}^{\ \ d} p_{ij}$

is maximization in each paging cycle. Constraint 1: $\forall d (1 \le d \le D^*), C_j \in C, \sum_{U, \in U} S_{ij}^d \le B$;

Constraint 2: if
$$s_{ii}^{d} = 1$$
, then $\forall d (d \neq d, 1 \leq d \leq D^{*})$ and $s_{ii}^{d} = 0$.

Proof: Assuming the multicast paging algorithm based on bipartite graph matching generates paging decision sequence $\{s^1, s^2 ... s^{n^2}\}$ in each paging cycle; assuming in the d-th paging cycle, paging decision matrix s^a is constituted by the Bu times maximum weight matching.

1) First, to prove the multicast paging algorithm based on bipartite graph matching satisfies:

Constraint 1:
$$\forall d (1 \le d \le D^*), C_j \in C, \sum_{U_i \in U} S_{ij}^d \le B$$

Set s^{di} represents the decision matrix $s^{dk} = \sum_{1 \le i \le k} M_i * sgn(W)$ which is generated until the k-th maximum weight matching;

Because in each matching M, $\forall U_j \in U$ and there is only one side having saturation point U_j in M, so $\forall U_j \in U$ and $\sum_{C_j \in C} S_{ij}^{-d(k+1)} - \sum_{C_j \in C} S_{ij}^{-dk} \le 1$.

By the algorithm that is shown in Fig.9: if
$$\sum_{U_i \in U} S_{ij}^{dk} = B$$
, then set $W_{ij} = 0 (\forall U_i \in U)$, then $\forall k', (k \le k' \le Bu)$ and $S^{dk'} = \sum_{1 \le i \le Bu} M_i * \operatorname{sgn}(W) = \sum_{U_i \in U} S_{ij}^{dk} = B$. So, $\forall d (1 \le d \le D^*), C_j \in C$, $\sum_{U_i \in U} S_{ij}^{d} \le B$.

2) Next, to prove the multicast paging algorithm based on bipartite graph matching satisfies:

Constraint 2: if
$$S_n^d = 1$$
, then $\forall d (d \neq d, 1 \leq d \leq D^*)$ and $S_n^d = 0$.

By the algorithm that is shown in Fig.9: if $s_{ij}^{d} = 1$, set $P_{ij} = 0$ then $w_{ij} = 0$. By $s^{d} = \sum_{1 \le i \le Bu} M_i * sgn(W)(d > d)$, it can be known $s_{ij}^{d} = 0$. So, this algorithm satisfies constraint 2.

3) To prove in the condition of satisfying both constraint 1 and constraint 2, the multicast paging algorithm based on bipartite graph matching ensure that $\sum_{1 \le i \le M} s_{ij}^{\ \ \ \ \ \ } p_{ij}$ is maximization in each paging cycle.

Lemma 1: Set i be a viable vertex label of graph G. If the equality sub-graph has a perfect matching M^* , then M^* is the maximum weight perfect matching of G.

Lemma 1 is proved as follows:

Due to the equal sub graph G_i is generated by graph G, so the perfect matching M^* of G_i is the perfect matching of G; furthermore,

$$w(M^*) = \sum_{e \in M^*} w(e) = \sum_{v \in V} l(v)$$

On the other hand, any perfect matching M of G has:

$$w(M) = \sum_{e \in M} w(e) \le \sum_{v \in V} l(v)$$

So, $w(M^*) \ge w(M)$, that is M^* is the maximum matching of G.

Therefore, the algorithm ensures that $\sum_{1 \le i \le M} s_{ij}^{\ d} p_{ij}$ is maximization in each paging cycle in the condition of satisfying both constraint 1 and constraint 2.

The time complexity of the Kuhn-Munkres algorithm is $O(n^3)$ and $n = \max(|M|, |N|)$. It can be known by algorithm 2 that in each paging cycle, the improved Kuhn-Munkres algorithm is called up to Bu times. If the anticipant paging delay of multicast paging is D^* , then the time complexity of the whole multicast paging algorithm is $O(D^*Bun^3)$, that is $O(n^3)$.

5. Experimental Results and Performance Analysis

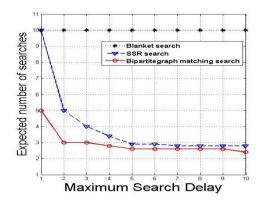
In order to verify and analyze the performance of the BMPS algorithm, we simulate and compare BMPS, the BP (Blanket Paging) algorithm used by GSM and SSR algorithm mentioned in reference [16]. The main idea of the BP algorithm is: if the available paging channel number of each area is B, then the BP algorithm pages the user number of B in all areas at the same time. SSR algorithm bases on the distribution probability of users in each area to assign the area with greater probability to users for paging in each paging cycle. For user conflict, that is, the assigned user number of one area exceeds the available paging channel number of the area, the solution of SSR is reassigning the users with smaller staying probability in the area to the area with suboptimal probability.

The simulations use MATLAB 7.7.0 to achieve the three multicast paging algorithms. Multicast paging experiment scene is in 10 paging areas, using BP algorithm, SSR algorithm and BMPS algorithm to page 10 mobile users, respectively. The following will from the effect of the maximum paging delay on paging overhead and delay, the performance of the algorithm in the situation of user conflict, the user fairness and bandwidth cost to simulate and analyze the BPMS algorithm.

5.1. Effect of Maximum Paging Delay on Algorithm Performance

In order to ensure the multicast paging algorithm can page all the M users within the maximum paging delay D, according to theorem 2, the available paging channels B of each area is set $as \lceil M / D \rceil$ in the simulation experiment. The specific numerical

of simulation parameters are set as follows: M=10 and N=10; Considering when D changes from 1 to 10, the changes of paging overhead and actual delay of the three algorithms.



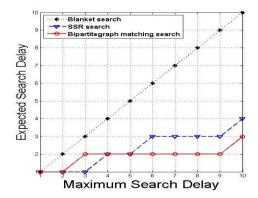


Figure 10. Paging Cost

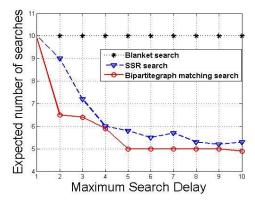
Figure 11. Paging Delay

Under the different maximum paging delay, the average paging overhead for each user of three multicast paging algorithm is shown in Figure 10. BS algorithm has the worst overhead performance, because BS algorithm need page each user in all the areas, doesn't consider the location probability distribution of users, so the average overhead of BS algorithm is the total number N of areas. As we expected, in Figure 10, the overhead performance of the BMPS algorithm is the best. With the increase of the maximum paging delay D, the average overhead paging of BMPS algorithm and SSR algorithm show a decreasing trend, and obviously are less than that of the BS algorithm. Compared with the SSR algorithm, the average overhead of BMPS algorithm reduces by 5%~50%. This is because when SSR algorithm assigns the area with larger location probability to each user, it may cause user conflict and affects the performance of the algorithm. However, the BMPS algorithm based on bipartite graph model maximum perfect matching ensures the total of the location probability for each paging decision is the maximum.

Figure 11 shows under the different maximum paging delay, the experienced delay for successfully paging the last user of BP algorithm, SSR algorithm and BMPS algorithm. With the maximum delay D increasing, the available paging channels B of each area are decreasing, so the actual paging delay of three algorithms are on the rise. BMPS algorithm has the minimum delay, SSR algorithm has the second. It can be seen from the figure that the BMPS algorithm can realize the optimization of the paging overhead and the whole performance.

5.2. Effect of User Conflict on Algorithm Performance

In order to study the effect of user conflict on paging overhead and delay of the three algorithms, we simulate the following multicast paging experiment scene: M=10;N=10; 4 mobile users are in area 3 while the other 6 users are in area 4. When D changes from 1 to 10, under the situation of user conflict, we consider the changes of paging overhead and actual delay of the three algorithms.



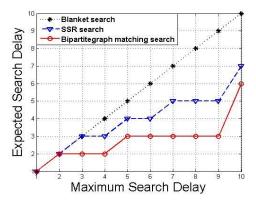


Figure 12. Paging Cost under Collision

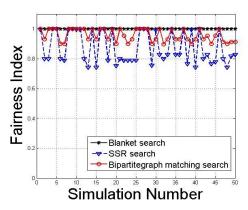
Figure 13. Paging Delay under Collision

Figure 12 and 13 shows with the user conflict, the average paging overhead and actual paging delay of the three algorithms. Comparing to the general situation of mobile users distribution as shown in Figure 10 and Figure 11, with the user conflicts, both the paging cost and the delay performance of SSR algorithm and BMPS algorithm are decreased. But the performance of the BMPS algorithm is better than SSR algorithm, as shown in Figure 12, compared to the SSR algorithm, BMPS algorithm can reduce maximum 40% of the paging delay. This is because the BMPS algorithm uses bipartite graph matching model to assign the area for users from the perspective of the overall paging system, and ensure to achieve the optimization of the overall paging overhead and delay under the situation of user conflict.

5.3. User Fairness

The paging process can't finish until the multicast paging successfully page the last user. Therefore multicast paging strategy must take into account the user fairness, as far as possible assigning equal paging times for each user.

The simulation parameters are set as follows: M=10;N=10;D=2. Figure 14 shows the user fairness of BP algorithm, SSR algorithm, BMPS algorithm in the 50 times simulation experiments. Horizontal axis represents the number of simulations, the vertical axis indicates the user fairness index during each multicast paging process, the calculation method is proposed by definition 4. This value is closer to 1, the better user fairness is. Because the BP algorithm assigns N areas to each user for paging, so its user fairness index is 1, that is user fairness is best. It can be seen from Figure 14, the user fairness index of the BMPS algorithm is between 0.9~1, and the user fairness index of the SSR algorithm is around 0.7~1, so the user fairness of the BMPS algorithm is obviously better than SSR algorithm.



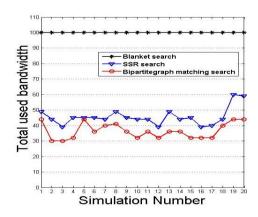


Figure 14. Fairness of Mobile Users

Figure 15. Bandwidth Overhead

5.4. Bandwidth Overhead

In the bandwidth limited wireless communication, the control channel is a very valuable resource, so the multicast paging algorithm should as far as possible to save the wireless channel resources under the condition of guarantee the paging delay. The simulation parameters are set as follows: M=10;N=10;D=3. Figure 15 shows the bandwidth overhead of BP algorithm, SSR algorithm, BMPS algorithm in the 20 times simulation experiments. The ordinate indicates the total number of paging channels which are used in the whole multicast paging process. Figure 15 shows that the bandwidth overhead of BP algorithm is the maximum and the value is always M*N=100. The bandwidth overhead of SSR algorithm is around 40~60. The bandwidth overhead of BMPS algorithm is the minimum, and the average value is about 35.

6. Conclusion

For the problem of limited bandwidth and delay of the multicast paging, this paper proposed an effective multicast paging scheme to optimize the performance of paging overhead and delay. First, in order to predict the location probability distribution of users, the paper proposed the multicast paging system model. For the system target, this paper proposed the multicast paging algorithm BMPS based on bipartite graph matching. The BMPS algorithm chose the optimal allocation scheme between users and paging areas by obtaining bipartite graph perfect matching. The paper proved the correctness of BMPS algorithm and analyzed the complexity of BMPS algorithm. Finally, simulation results indicated that BMPS algorithm could achieve optimization performance of paging overhead and delay, could effectively solve the problem of user conflict, could guarantee user fairness and save bandwidth overhead.

Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 61201231) and Beijing Natural Science Foundation Major Project (No. 61222103).

References

- [1] Y. Zhou, J. Wang, T. S. Ng, K. Higuchi and M. Sawahashi, "OFCDM: a promising broadband wireless access technique", IEEE Communications Magazine, vol. 46, (2008) March, pp. 39-49.
- [2] Y. Zhou, H. Liu, Z. Pan, L. Tian and J. L. Shi, "Spectral and Energy Efficient Two-Stage Cooperative Multicast for LTE-A and Beyond", IEEE Wireless Magazine, (2014) April, pp. 34-41.

- [3] Y. Zhou, H. Liu, Z. Pan, L. Tian, J. L. Shi and G. Yang, "Two-Stage Cooperative Multicast Transmission with Optimized Power Consumption and Guaranteed Coverage", IEEE JSAC on SEED, vol. 32, no. 2, (2014) February, pp. 274-284.
- [4] A. Bar-Noy and G. Malewicz, "Establishing wireless conference calls under delay constraints", The 21th ACM Symp, on Principles of Distributed Computing (PODC), (2002), pp. 41-50.
- [5] A. Bar-Noy, I. Kessler and S. Moshe, "Mobile users: to update or not to update?", Wireless Network, (1995), pp. 175-195.
- [6] Y.-H. Zhu and V. C. M. Leung, "Optimization of Sequential Paging in Movement-Based Location Management Based on Movement Statistics", IEEE Transaction on vehicular technology, vol. 56, no. 2, (2007) March, pp. 955-964.
- [7] B. Hajek, K. Mitzel and S. Yang, "Paging and Registration in Cellular Networks: Jointly Optimal Policies and an Iterative Algorithm", IEEE Transactions on information theory, vol. 54, no. 2, (2008) February, pp. 608-622.
- [8] J. Petzold, A. Pietzowski, F. Bagci, W. Trumler and T. Ungerer, "Prediction of Indoor Movements Using Bayesian Networks", In Location- and Context-Awareness (LoCA 2005), Oberpfaffenhofen, Germany, (2005) May.
- [9] L. Vintan, A. Gellert, J. Petzold and T. Ungerer, "Person Movement Prediction Using Neural Networks", In First Workshop on Modeling and Retrieval of Context, Ulm, Germany, (2004) September.
- [10] D. Ashbrook and T. Starner, "Using GPS to learn significant locations and predict movement across multiple users", Personal and Ubiquitous Computing, vol. 7, no. 5, (2003), pp. 275-286.
- [11] S. Lee and K. C. Lee, "Context-prediction performance by a dynamic bayesian network: Emphasis on location prediction in ubiquitous decision support environment", Expert Systems with Applications, vol. 39, no. 5, (2012), pp. 4908-4914.
- [12] J. Petzold, F. Bagci, W. Trumler and T. Ungerer, "Comparison of Different Methods for Next Location Prediction", Lecture Notes in Computer Science, Euro-Par 2006 Parallel Processing, pp. 909-918.
- [13] M. Maitra, D. Saha, P. Sarathi Bhattacharjee and A. Mukherjee, "An Intelligent Paging Strategy Using Rule-Based AI Technique for Locating Mobile Terminals in Cellular Wireless Networks", IEEE Transactions on vehicular technology, vol. 57, no. 3, (2008) May, pp. 1834-1845.
- [14] Y. Xiao, H. Chen and M. Guizani, "Non-Blocking Pipeline Paging with Known Location Probabilities for Wireless Systems", IEEE Transactions on wireless communications, vol. 6, no. 10, (2007) October, pp. 3632-3640.
- [15] L. Epstein and A. Levin, "A PTAS for delay minimization in establishing wireless conference calls", Discrete Optimization, vol. 5, (2008), pp. 88-96.
- [16] A. Bar-Noy and Z. Naor, "Efficient multicast search under delay and bandwidth constraints", Wireless Netw, vol. 12, (2006), pp. 747-757.
- [17] A. Bhattacharya, S. K. Das, "LeZi-Update: An Information-Theoretic Framework for Personal Mobility Tracking in PCS Networks", Wireless Networks, (2002), pp. 121-135.
- [18] A. Boytsov and A. Zaslavsky, "Extending context spaces theory by proactive adaptation", Smart Spaces Next Gener Wired/Wireless Network, pp. 6294, (2010), pp. 1-12.
- [19] A. Boytsov and A. Zaslavsky, "Context prediction in pervasive computing systems: achievements and challenges", Supporting Real Time Decis. Making, vol. 13, (2011), pp. 35-63.

Authors



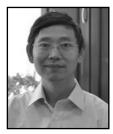
Yuan Yao, he received the Dr. Eng. degree from Institute of Computing Technology, Chinese Academy of Sciences in 2010, and received the M.S. degree from Beijing Jiaotong University in 2007. He is currently working at the Chinese Academy of Sciences. His research interests include wireless communications, system architecture, networking protocol.



Zhang Dalin, he received the Dr. Eng. degree from the University of Fukui, Japan in 2013, and received the B.S. degree in Electrical Information Engineering and M.S. degree in Computer Application Technology from Beijing University of Chemical Technology, China, in 2006 and 2009, respectively. He is currently working as a postdoctor in the Institute of Computing Technology, Chinese Academy of Sciences. His research interests include additive array antenna, MIMO system, mobile network architecture, etc.



Wang Qing, she received M.S. degree in the Institute of Computing Technology, Chinese Academy of Sciences. Her research interests include wireless communication theory and Embedded Operating System.



Shi Jinglin, he (sjl@ict.ac.cn) currently serves as the director of the Wireless Communication Technology Research Center of ICT/CAS. He is also a visiting professor at Beijing University of Posts and Telecommunications, the University of Sydney, the University of Wollongong, and Macquarie University. His research interests include wireless communications system architecture and management, wireless signal processing theory, and wireless communications baseband processor design.

International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.8, No.11 (2015)