Network Selection For One-to-Many Services In 3G-Broadcasting Cooperative Networks

Luan Huang, Kar Ann Chew, Rahim Tafazolli

Mobile Communications Research Group
Center for Communication System Research
CCSR, University of Surrey
Guildford, UK
{L. Huang, K. Chew, R. Tafazolli }@surrey.ac.uk

Abstract—**The beyond 3G or 4G mobile systems envisions heterogeneous infrastructures comprising diverse wireless systems, e.g., 2G, 3G, DVB, WLAN, and various transmission approaches, e.g., "one-to-one" and "one-to-many". In this context, a *network selection* (NS) problem emerges regarding determining the appropriate Access Network (AN), as users are reachable through several different ANs. This paper addresses the issue of provisioning "one-to-many" services over heterogeneous wireless networks in terms of how to choose the AN that satisfies the bandwidth requirement of services, while maximizing the *system profit* obtained in the combined network. A heterogeneous network comprising Multicast Broadcast Multimedia Service (MBMS) of the third generation mobile terrestrial network and the digital video broadcasting transmission system for handheld terminals (DVB-H) is adopted in this study. Both networks cooperate and complement each other to improve the resource usage and to support "one-to-many" services with their multicast and broadcast transmission capabilities. Based on this architecture, an algorithm framework is defined and proposed to solve the NS problem for the "one-to-many" services. Six schemes based on the algorithm framework are then evaluated by simulation.**

Keywords- network selection; heterogenerous; QoS; multicast; broadcast;

I. INTRODUCTION

A recent effort of the cellular infrastructure enhancement is towards "Beyond 3G (B3G)", or "composite radio" (CR) [1]. These concepts assume heterogeneous networks infrastructure comprising different wireless access systems such as GSM/GPRS, UMTS, DVB and WLAN. The study presented in this paper considers the CR of MBMS and DVB-H networks as the research target for "B3G", due to their inherent "one-to-many" or multicast/broadcast transmission capabilities, which have been regarded as resource efficient to deliver multimedia services over bandwidth-limited wireless systems [2]. MBMS is a technique under investigation within the 3GPP framework as an enhancement to the current UMTS [3], while DVB-H a standard that is given additional features based on current DVB to support IP-based broadcast services for handheld portable and mobile terminals [4].

One of the principle research challenges involved in heterogeneous networks is the *network selection* (NS) problem regarding determining the appropriate AN among composite ANs when users are reachable through several ANs. Recently, there are research efforts on addressing the NS problem for "one-to-one" or unicast service. [5] [6] [7] proposed a user-centric mechanism in which every mobile user chooses the most profitable AN according to the QoS profile, user preference, price model and network conditions. However, a new problem emerges when multicast based "one-to-many" services are introduced in heterogeneous networks. We term this new problem as Network Selection for "One-to-Many" Services (NS-OMS). In this context, the user-centric NS for unicast services is not applicable for multicast services due to the diverse preferences and requirements of users. Thus, our work expands the earlier work by handing over the decision-making responsibility from each user to the network that chooses the most appropriate AN for the multicast service desired by a group of users. To our best knowledge, this paper is the first attempt to address the problem of NS-OMS. A fundamental algorithm framework is first defined and proposed. Six schemes based on this algorithm framework are then evaluated by simulations. The implementation issues of the proposed algorithm are addressed in [8] in terms of architecture, functional entities, network interfaces, signaling and associated parameters. The application to be considered in this NS-OMS problem is the Group-based Mobile Multimedia-on-Demand (GMMoD) service, which is a real-time streaming service. GMMoD aims to stream low-cost multimedia services by sending one copy of data through one multicast channel to subscribed mobile users. These users are referred to as a group [2].

The rest of the paper is structured as follows. In section II, our algorithm framework to solve the NS-OMS problem is discussed in detail. The simulation model and results obtained are demonstrated and analyzed in section III. Section IV concludes the paper with directions for the future work.

II. NS-OMS ALGORITHM FRAMEWORK

An algorithm framework is developed and proposed in this paper to solve the NS-OMS problem. It provides a solution in terms of essential functionalities, interfaces and procedure for network operators or service providers to find out the appropriate AN for the required QoS guarantees of multicast services. Fig. 1 illustrates the algorithm framework. This algorithm framework consists of three main parts, namely, *constrains and goal*, *service scheduling* and *resource allocation*. Constrains and goal identify constrains to obey and

the objective to achieve when solving the NS-OMS problem, each adjustable according to specific requirements of different network operators and service providers. Service scheduling determines in what time and which requests are ready for the resource allocation that will in turn, determine the most appropriate AN for these requests. By using different mechanisms in service scheduling and resource allocation, this algorithm framework is flexible to enable future enhancements.
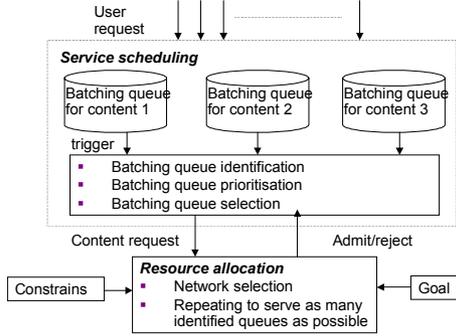

Figure 1. NS-OMS Algorithm Framework

Each part involved in the NS-OMS algorithm framework is described in depth below.

A. Constrains and goal:

This paper considers wireless resources in the downlink as constrains due to the following reasons. Firstly, the scarce wireless resource is the major obstacle to the cost effective provisioning of high bandwidth multimedia services in wireless networks. Secondly, the wireless multimedia service envisions asymmetric transmission with more resource required in downlink than in uplink. Finally, the property of GMMoD application does not allow the control function during an application transmission, such as, pause, rewind etc., which alleviates the resource stress further in the uplink.

MBMS has WCDMA as the air interface in which the total transmission power of the downlink is shared among all the users. In order to maintain a stable and efficient system, the downlink operation is usually controlled through the transmission power rather than through the cell load factor, as in the uplink. Therefore, the total transmission power in a Node B of UMTS network needs to be constrained below an available but finite transmission power budget before allocating resource to a new session. The power constrain is given below [9]:

$$P_{total\_old} + \Delta P_s < P_{threshold} \qquad (1)$$

where, P_{total_old} is the total downlink transmission power before the admittance of the new session s and is the aggregation of the downlink transmission power required by each active connection; $P_{threshold}$ is the volume of the maximum available transmit power adjustable by network operators; ΔP_s is the estimated transmission power required by the new session s. MBMS does not use the fast power control mechanism that applies to determine the transmission power for each active receiver in UMTS. Instead, the proportion of the available Node B power allocated to a particular session with a certain

bandwidth requirement will essentially be fixed to cover certain coverage.

In DVB-H, IP datagrams are encapsulated into the multiprotocol encapsulation (MPE) section and then multiplexed with TV data in a DVB national/regional multiplex (MUX) [4]. The admittance of a new IP session mainly depends on the size of the residual space available after reserving enough space for TV data in the MUX. Therefore, similar to the power constraint in MBMS, in DVB-H, the resource can be allocated to the new IP session only if the maximum available bandwidth budget is not exceeded, i.e.,

$$B_{total\_old} + \Delta B_s < B_{threshold} \qquad (\Delta B_s = R_s) \qquad (2)$$

where, B_{total_old} is the volume of the total occupied bandwidth by IP active sessions before the admittance of the new session s; ΔB_s is the bandwidth of the new session s and is estimated by its information rate R_s (i.e., the guaranteed bit rate) according to its QoS level; $B_{threshold}$ is the volume of the maximum available transmission bandwidth adjustable by network operators and should be configured by taking the space of the reserved resource for TV data into account.

The goal of NS-OMS algorithm is to select the most appropriate service set-up time and AN for each user's GMMoD content request so that user's QoS requirement, in particularly the bandwidth, is satisfied and resources constrains are met while maximizing the *System Profit (SP)*. *SP* is the difference between the revenue earned in serving users and the resource cost consumed in transmitting services. By assuming the volume-based charging model, *SP* is derived as follows:

$$SP = \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{j=1}^{M} b_i d_i (w_i n_{kij} - \frac{r_{kij}}{C_k}) \qquad (3)$$

where, K is the number of inter-worked networks; N is the number of provided contents in a network; M is the number of established groups for a content in a network; b is the content bandwidth and d is the content duration; w is revenue per bit for a content earned from a group; n is the group size in the number of requests; r/C is the resource cost per bit for a group, where r is the amount of resources consumed by a group and C is the amount of available resources in a network. Note that the resource units to be considered are different in various networks, i.e., power for MBMS and bit rate for DVB-H.

B. Service scheduling

Service Scheduling involves the following four steps:

1) Batching: Batch or aggregate multiple users' requests for the same content in a *batching queue* before the service set-up. In the batching, a user could leave from the batching queue due to the lack of the patience to wait, which is referred to as *reneging* [10]. Users in each batching queue are treated as a specific group requiring a particular content. The requests in one group are to be served simultaneously over a multicast channel in order to save wireless resources. More detailed discussion regarding baching can be found in [2].

2) Queue Idenitification: Identify and mark queues for the resource allocation where the reneging has happened. This step

is triggered by a user just about to leave. Note that if this user can not be served in the resource allocation procedure due to the lack of resouces, this user reneges and is lost, otherwise, this user along with the other users in this queue is served.

3) Queue Prioritization: Sort all marked queues identified in step 2 in descending order depending on certain proprity creteria. The higher order corresponds to the higher priority to be considered for the resource allocation. Two creteria are considered in this paper:

a) First Come First Serve (FCFS): Queues are sorted according to t_c-t_1, where t_c is the current time at the scheduling point and t_1 is the arrival time of the first user staying in the queue.

b) Maximum Queue Length (MQL): Queues are sorted according to n, where n is the number of users staying in the queue.

4) Queue Selection: select one queue from all identified queues, at a time, in the order made in step 3;

C. Resource allocation

Resource allocation involves the following two steps:

1) Network Selection: select and assign an appropriate AN to the queue chosen in step of *Queue Selection*, according to a certain policy. The selected AN is used to transmit the content requested. Three policies are examined in this paper:

a) Fixed AN assigment (FANA): Assign users to the only AN that provides the content required. This AN is refered to as an originating network. The other potential ANs are referred to as cooperating networks. If there is no enough resource to satisfy the bandwidth requirement in the originating network, users' requests are batched in the queue until they renege or free resources are released;

b) Retry-based AN assignment (RANA): First, assign users to their originating networks. The cooperating network will be tried to set up the connection whenver there is no enough resource available in the originating network;

c) Objective-based AN assignment (OANA): Each AN is treated equally. The chosen AN is that gives the hightest *SP* for the content desired, which is calculated based on equation (3).

2) Repeating: repeat the step of *queue selection* and in turn *resource allocation* for the chosen queue to accomondate as many queues as possible into networks.

III. PERFORMANCE EVALUATION

Six schemes (based on two *Queue Prioritization* criteria, three *Network Selection* policies) have been evaluated to investigate the proposed algorithm framework for the NS-OMS problem by a simulator constructed by OPNET [11]. Heterogeneous networks containing MBMS with seven cells and DVB-H with one cell are constructed, both cover one service area *A*. Two service classes are assumed. Each service class is characterized by operator, request arrival rate, mean duration and bandwidth. Each class includes five different contents. The duration of content is given by a random variable, which is determined by an Exponential distribution and is

updated every one hour. For schemes evaluated in this paper, the revenue per bit of each content is set to 0.004 in order to give a positive *SP* and for simplification to present schemes. Requests for each service class is generated with Poisson distribution, and for each service class request, the probability of choosing a specific content of this service class is given by the Zipf distribution [12], i.e., the probability of choosing the *k*th content is given by:

$$p_k = \frac{c}{k^{(1-\theta)}} \qquad (4)$$

$$c = 1 \Big/ \sum_{k=1}^{M} \frac{1}{k^{(1-\theta)}} \qquad (5)$$

where, *c* is a normalization constant; *M* is the number of contents provided by a service class; θ is the parameter that determines the shape of the distribution. A small θ corresponds to a more severe discrimination of requests among contents and indicates that some contents are desired more frequently than the others. As in [13], θ=0.271 is used.

A deterministic reneging model is used [10]. In this model, after waiting a certain period of time given by the *waiting threshold w*, the user's remaining time until departure is given by the random variable *R*, which is determined by an Exponential distribution with a *mean waiting time u* that elapses after *w* until the user leaves the system. Thus the user's overall wait time is given by *W=w+R(u)*.

A fluid flow mobility model is adopted [14]. In this model, the mobile terminals are uniformly distributed over a given cell. Mobile terminal have a mean velocity of *V* and their direction of movements are uniformly distributed over [0, 2π]. The residing time of a mobile within a cell is exponentially distributed and the mean dwell time can be estimated by $1/\eta$, where η is the cell crossover rate. The crossover rate for the circular cell of radius *r* is given by:

$$\eta = \frac{2V}{\pi r} \qquad (6)$$

Simulation parameters used in this study are listed in Table 1 and Table 2.

TABLE I. SIMULATION PARAMETERS

Simulation Parameters	Value	
	Service 1	*Service 2*
Operator of service 1	MBMS	DVB-H
Request arrival rate (requests/hour/cell)	300	1500
Mean duration (minutes)	10	15
Bandwidth (kbps)	64	128
MBMS cell radius	1400 m	
Available resources in MBMS	18W/cell	
Available resources in DVB-H	2Mbps/cell	
Mean mobility velocity	3 km/hour	
Mean waiting time	2 minutes	

Simulation Parameters	Value	
	Service 1	*Service 2*
Waiting threshold	Variable	

TABLE II.   REQUIRED FRACTION OF CELL TX POWER-3KMH [15]

Date Rate	Fraction of cell Tx power for 1% BLER and –3dB Geometry
64kbps	17.8%
128kbps	38%

The performance is evaluated by varying the *waiting threshold*. Fig. 2 and Fig. 3 present the *SP* per unit time and the *satisfactory ratio* with varying *waiting threshold*. *Satisfactory ratio* is defined as the ratio of satisfied requests to offered requests. A satisfied request is one that did not renege from the waiting queue before being served, and that is not dropped during a cell-to-cell horizontal-handover in a network
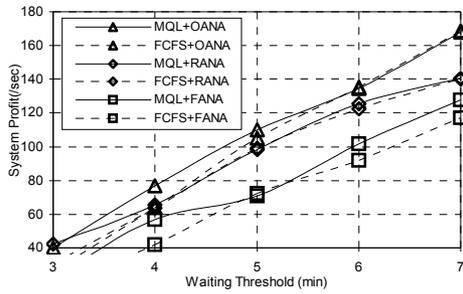

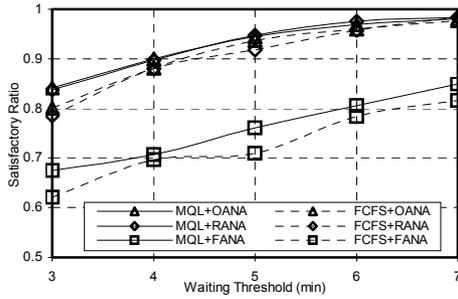Figure 2.   System profit versus waiting threshold

.

Figure 3.   Satisfactory ratio versus waiting threshold

When the *waiting threshold* increases, all schemes give higher *SP* and *satisfactory ratio* because longer *waiting threshold* corresponds to the higher possibility of users waiting for the content set-up without reneging, which in turn reduces the request arrival rate for the resource allocation, thereby releasing more free resources. Of all the schemes, OANA and RANA tend to obtain the performance achievable by satisfying all the users, i.e., 100% *satisfactory ratio* in maximum. On the contrary, the separated resource usage in FANA prevents it from reaching the maximum performance.

Furthermore, it can be seen that the MQL prioritization criterion gives more *SP* and higher *satisfactory ratio* than the FCFS prioritization criterion with the same network selection policy. This is because MQL is able to schedule more number of users into one multicast channel, which increases the resource utilization.

In addition, under the same queue prioritization criteria, OANA and RANA give similar *satisfactory ratios,* which are higher than FANA. This is because both RANA and OANA are able to redirect traffic into cooperating ANs whenever there are no enough resources, ensuring that resources are utilized fully and efficiently. On the contrary, FANA policy only uses resources in the originating network for the content desired, which results in an unbalanced network status, i.e., overloading in a network, whereas underloading in the other. Moreover, in terms of *SP*, OANA obtains better performance over other network selection policies because OANA always transmits the traffic through the network that gives the higher profit.

In general, from results displayed in Fig. 2 and Fig. 3, it can be concluded that MQL+OANA is the optimum scheme in terms of *SP*. To further assess the performance of schemes considered, in particular the impact of different *Queue Prioritization Criteria*, we look at another important performance metric, i.e., system unfairness. The degree of the system unfairness is defined by the variance of the *satisfactory ratio* of the contents, which is in turn divided by the total number of contents provided in the system, i.e.,

$$SystemUnfairness = \sqrt{\frac{\sum_{k=1}^{M}(SR_k - \overline{SR})^2}{M}} \qquad (7)$$

where, SR_k is the *satisfactory ratio* for a content k and M is the total number of contents provided by the system. Clearly, the system unfairness increases when there is a discrimination of requests among contents corresponding to the variety of content popularities.
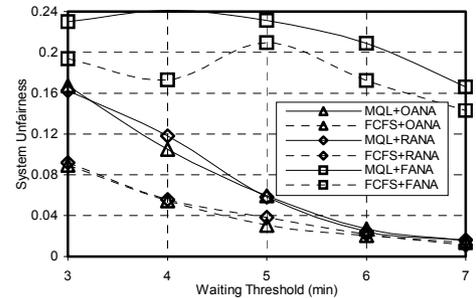

Figure 4.   System unfairness versus waiting threshold

Fig. 4 presents the system unfairness for varying the *waiting threshold*. It can be seen that the FCFS queue prioritization criterion gives lower system unfairness than the MQL with the same network selection policy. This confirms the fairness capability of FCFS because FCFS treats each content equally regardless of its popularity. In addition, the FANA network selection policy gives the higher system unfairness over the

other policies due to the unbalanced resource utilization in two networks.

From results presented above, it can be concluded that to achieve our goal, i.e., maximizing the *SP*, the scheme of MQL+OANA is the optimum scheme. MQL can increase the number of satisfied users by always scheduling the longest queue, whereas, OANA can help to increase the system profit for each selected queue by assigning the queue to the most profitable access network. However, the high *SP* gained by MQL+OANA is at the price of the high system unfairness that can be improved by using FCFS+OANA. This result indicates that the optimum scheme varies with various goals corresponding to different requirements of particular network operators and service providers.

IV. CONCLUSIONS

As mobile communications is enhanced towards "composite radio", multicast transmission is attracting more and more attention because it offers a more cost-effective provisioning of multimedia services. This paper proposed, for the first time, an algorithm framework to solve the problem of the network selection for the multicast traffic, which is termed as Network Selection for "One-to-Many" Services (NS-OMS) problem, by studying the scenario of MBMS and DVB-H cooperative networks. The goal is to satisfy users' QoS requirements of the content desired, while maximizing the *system profit*. This algorithm framework provides functionalities, interfaces and procedure to solve the NS-OMA problem. The flexibility of this algorithm framework enables future enhancements. Six schemes based on the algorithm framework have been developed and evaluated by simulation. The results show that the MQL+OANA scheme obtains better performances than other schemes in terms of the *system profit*.

In terms of future work, we plan to investigate more schemes based on the algorithm framework by considering different optimization objectives and to find out the most appropriate scheme in different environments. We also plan to enhance our algorithm framework by considering the adaptive QoS to the time-varying network conditions and resource availability.

V. ACKNOWLEDGEMENTS

[1] V. Gazis, N. Housos, A. Alonistioti, and L. Merakos, "Generic system architecture for 4G mobile communications", IEEE VTC, April 2003

[2] L. Huang, K. A. Chew, S. Thilakawardana, R. Tafazolli, "Optimal approach of achieving group-based mobile multimedia-on-demand services," 3G, Oct., 2004.

[3] 3GPP, "Introduction of multimedia broadcast/multicast service (MBMS) in the radio access network (RAN); stage 2" TS 25.346, v6.1.0, June 2004

[4] ETSI, "Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)," EN 302 304 v1.1.1, June, 2004

[5] E. Alexandri, G. Martinez, and D. Zeghlache, "An intelligent approach to partition multimedia traffic onto multiple radio access networks."

[6] V. Gazis, N. Houssos, N. Alonistioti and L. Merakos, "On the Complexity of "Always Best Connected" in 4G Mobile Networks," IEEE VTC, Oct. 2003

[7] J. Noonan, P. Perry and J. Murphy, "Client Controlled Network Selection," 3G, Oct., 2004.

[8] L. Huang, K. A. Chew, and R. Tafazolli, "Distributed resource management for interworking of cellular and digital broadcasting networks," IST Summit, June, 2004.

[9] H. Holma and A. Toskala, "WCDMA for UMTS. Radio Access for Third Generation Mobile Communications," John Wiley & Sons, Ltd., England, 2001

[10] H. Shachnai, P. S. Yu, "Exploring wait tolerance in effective batching for video-on-demand scheduling," Multimedia Syst. J., vol. 6, no. 6, Dec. 1998

[11] www.opnet.com

[12] Y. S. Chen, "Mathematical modeling of empirical laws in computer application: A case study," Comp. Math. Applicat., pp.77-87, Oct. 1992.

[13] S. H. Gary Chan, and S. H. Ivan Yeung, "Broadcasting video with the knowledge of user delay preference," IEEE Transactions on Broadcasting, vol. 49, no. 2, June. 2003.

[14] B. Jabbari and W. F. Fuhrmann, "Teletraffic modeling and analysis of flexible hierarchical cellular nertworks with speed-sensitive handoff strategy," IEEE Journal on Selected Areas in Communications, vol. 15, no. 8, Oct. 1997.

[15] 3GPP, "S-CCPCH Performance for MBMS," TR 25.803, v1.3.0, March 2004