Optimization of Radio Resource Allocation for Multimedia Multicast in Mobile Networks
Qing Xu, Hakim Mabed, Frédéric Lassabe, Alexandre Caminada

To cite this version:
Qing Xu, Hakim Mabed, Frédéric Lassabe, Alexandre Caminada. Optimization of Radio Resource Allocation for Multimedia Multicast in Mobile Networks. VTC2013-Spring, IEEE 77th Vehicular Technology Conference, Jun 2013, Dresde, Germany. IEEE, VTC2013-Spring, IEEE 77th Vehicular Technology Conference, pp.1-5, 2013, <10.1109/VTCSpring.2013.6692680>. <hal-01304620>

HAL Id: hal-01304620
https://hal.archives-ouvertes.fr/hal-01304620
Submitted on 20 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Optimization of Radio Resource Allocation for Multimedia Multicast in Mobile Networks

Qing Xu*, Hakim Mabed§, Frédéric Lassabe* and Alexandre Caminada*

* Laboratoire Systèmes et Transports (SeT), Université de Technologie de Belfort-Montbéliard, France
Email: (qing.xu, frederic.lassabe, alexandre.caminada)@utbm.fr

§ Département d’Informatique des Systèmes Complexes (FEMTO-ST), Université de Franche-Comté, France
Email: hakim.mabed@pu-pm.univ-fcomte.fr

Abstract—In this paper we present a mathematical modeling of Radio Resource Management (RRM) for multicast service diffusion based on Multimedia Broadcast Multicast Service (MBMS) standard. In this model, a flexible allocation approach named F2R2M is proposed, combining three candidate transport channels with scalable video transmission technology. The allocation procedure is implemented based on simulated annealing algorithm with a two-dimensional optimization objective and a lexicographic order evaluation criteria. Experiments prove that, comparing with existing channel allocation approaches, F2R2M obtains allocation solution with equal QoS and lower transmission power consumption. Moreover, it reduces the possibility of achieving saturation of power or channelization codes when simulation scenarios have more users and heavy traffic load.

Index Terms—RRM; multicast; UMTS; MBMS; optimization;

I. INTRODUCTION

To support efficient distribution of multicast multimedia services over mobile network, the 3rd Generation Partnership Project (3GPP) specified the Multimedia Broadcast Multicast Service (MBMS) system for 3G network since Release 6 [1]–[3]. MBMS system is considered as a substantial platform for multicast service since it provides the diffusion of multicast multimedia services with efficient allocation of radio resource and economic resource usage. A wide range of work is investigated on MBMS for 3G network, we classify them into: a) efficient radio resource allocation; b) integration with video scalability; and c) scheduling of transmitted streams.

a) Efficient Radio Resource Allocation: In UMTS Terrestrial Radio Access Network (UTRAN) where the radio resources (power and channelization codes) are limited, the sharing of resources among numerous users per cell is constrained with more services subscriptions and higher requested traffic bandwidth. Prior research on this topic focus on i) power saving technologies [4], [5] since less power consumption brings less interferences thus larger cell capacity; ii) the selection and switching of transmission modes [6]–[8] which is crucial to the allocation efficiency, because different channels for carrying MBMS traffic have different characteristics in power and channel code consumption; iii) enhancement such as Macro Diversity and Spatio-Temporal Transmit Diversity (STTD) can also bring considerable gains [9].

b) Integration with Scalable Transmission: The multicast service quality can be improved by adapting the scalability ratio (bit rate and frame rate) of multimedia streams through different coding structures [10].

c) Scheduling of Streams: Proportional scheduling algorithms of multicast streams by using Time Division Multiplexing (TDM) are proposed for CDMA2000 [11], in which each multicast group is served by one channel. Base station schedules streams by determining the target multicast group and transmit rate per time slot.

Although MBMS RRM in 3G network has been extensively studied, multiple aspects are still not well balanced. When transmission power is saturated, should we transmit service through basic quality with full coverage or through advanced quality with smaller coverage? When channel codes are saturated, the transmission mode should be selected based on less power consumption or less occupation of channel codes? To address these demands, we propose a Flexible Radio Resource Management Model (F2R2M) combining transmission mode selection and multimedia scalability. In this model, a lexicographic-order criteria is proposed to evaluate the quality of resources allocation in terms of service satisfaction and resource consumption. Then a combinatorial optimization algorithm is presented to find the best allocation configuration with preferable balance between radio resource consumption, service coverage and service quality.

This paper is structured as follows. In section II we introduce related work on RRM for MBMS. The model description and optimization strategies are presented in section III. In section IV we illustrate the experiments along with a comparative evaluation with the other existing algorithms. Final conclusions are drawn in section V.

II. RELATED WORK

The basic MBMS introduction followed by an analysis of the related work are given in this section.

1) MBMS Transmission Modes: The MBMS service over UTRAN interfaces could be carried by PTM and PTP mode. In PTM mode, MBMS data is carried by a forward access channel (FACH) covering the whole cell. Each FACH needs one channel code serving large amount of users, but may waste power when there is small number of users or users are very close to Node B [5]. The PTP mode uses the dedicated channel (DCH) or shared channel (HS-DSCH), each DCH needs one channel code serving one dedicated user; the shared channel
occupies up to 15 channel codes for users. PTP mode controls link quality better than PTM but the served user number is limited due to power and channel code restriction [12], [13].

2) **MBMS Power Counting**: 3GPP defines MBMS power counting (MPC) mechanism [13] aiming to minimize Node B’s power requirements during transmission. Before data transfer, when the estimated power consumption of MBMS service in a cell is under an operator-defined threshold, network will establish PTP connections. The switch from PTP to PTM occurs when power exceeds the threshold, and vice versa. MPC has limited flexibility because it only considers delivering service for all users with full service quality, and does not support PTP and PTM for one service concurrently. Therefore, when MBMS transmission power in one cell is near saturation, MPC does not provide alternate allocation scheme (e.g. reduce power consumption by decreasing service’s quality) allowing new service or new users to access into network.

3) **Dual Transmission Mode**: Dual transmission mode (DTM) allows the co-existing usage of PTP and PTM mode for one MBMS service [14]. It adapts FACH coverage for users with better link quality, while the users near the cell edge are served by DCHs. FACH coverage is dynamically adapted by changing transmission power, meanwhile the DCH connections are released or established. The advantage of DTM is obvious during handover for single user. But it does not take into account HS-DSCCH, which can increase the power efficiency for MBMS [5], [8].

4) **Scalable FACH Transmission**: Scalable transmission is a potential power saving technology for MBMS [9], [15]. With scalable video coding, multicast service can be divided into single layer (SL) and multiple layer (ML) transmission schemes [10]. ML service can split into several streams with lower bit rate hence lower QoS requirement compared with a non-scalable stream. (e.g. 256 kbps service has two 128 kbps flows). Scalable FACH transmits flows through common channels with predefined geographical coverage [9]. The basic flow is sent to all subscribers (95% coverage) to guarantee service reception, the advanced flow is sent to users within 50%. Basic flow’s transmission power is reduced with lower bit rate, and so do the advanced flows with smaller coverages. Scalable FACH is not optimized in terms of flexibility. On the one hand, when cell power is ample, saving power consumption by reducing service quality is not necessary; on the other hand, when service demand is too high to be satisfied with full service quality due to power saturation, the trade-off between service quality and power is not efficient with fixed coverages.

5) **Dynamic Power Setting**: Dynamic Power Setting (DPS) for PTM mode was initially raised in [16]. Instead of fixing the FACH power to cover the whole cell, based on the dynamic and periodic report from user, RNC dynamically adjust FACH power to just achieve the worst users. DPS is utilized in our work and integrated with the other advanced mechanisms.

III. **DESCRIPTION OF FLEXIBLE RADIO RESOURCE MANAGEMENT MODEL (F2R2M)**

We propose here a complete mathematical model allowing to numerically evaluate the quality of radio resource allocation solution for every cell controlled by a given RNC.

A. **Model Phases**

The procedures of our model is divided into three phases:

1) **Parameter Collect Phase**: RNC collects MBMS service and user information.
2) **Estimation Phase**: RNC searches for the optimum allocation of radio resource through optimization procedure.
3) **Resource Allocation Phase**: when MBMS session start, RNC establishes the transport channels for selected users and allocates the planned channelization codes and power for channels.

Before and during data transfer, any change of MBMS session state (e.g. user mobility, new service) will trigger the estimation phase, in which, RNC collects following variables:

- \( T(c) = \{ t_1, ..., t_k \} \). Set of mobile terminals in cell \( c \).
- \( C(c) = \{ (x_1, y_1), ..., (x_k, y_k) \} \), \( t_k \in T(c) \). Set of instantaneous coordinates of terminals.
- \( S(c) = \{ s_1, ..., s_{N_s} \} \). Set of services in cell \( c \).
- \( F(s_i) = \{ f_{s_i,0}, f_{s_i,1}, f_{s_i,2}, f_{s_i,3} \} \), \( s_i \in S(c) \). Set of flows (and their bandwidth) of each service, 
- \( Dist(s_i) = \{ t_{k-1}, t_k, ..., t_k \} \), \( s_i \in S(c), t_k \in T(c) \). Multicast group of service.

B. **Decision Variables**

Unlike MPC and DTM, our model performs channel allocation for each flow composing a service, hence each service can be transmitted either in scalable mode \( f_{s,1} \) and the advanced flows) or non-scalable mode (the original content \( f_{s,0} \)).

F2R2M allows the combination of PTM and PTP modes for a given flow. The possible assignments of transport channel include: i) pure PTP or pure PTM mode, i.e. the conventional modes in MBMS standard; ii) mix of PTP mode: co-existing of dedicated and shared channel to transfer the same flow to different users; and iii) co-existing of PTP and PTM modes.

Consequently, for each flow \( f_{s,j} \) of a service \( s \), our algorithm partitions the multicast group into four disjointed sets: users covered by a FACH UE \( UE_{FACH}(f_{s,j}) \); users served through
DCHs UE\textsubscript{dch}(f\textsubscript{s,j}) users sharing HS-DCH UE\textsubscript{hs}(f\textsubscript{s,j}) and not served users UE\textsubscript{noch}(f\textsubscript{s,j}). In addition, F2R2M decides to diffuse the original or the scalable flows of service.

According to the user sets UE\textsubscript{type}(f\textsubscript{s,j}) for flows and the requested bandwidth, RNC performs a deterministic procedure to associate available channel code(s) (according to OVSF allocation scheme [17]) with each nonempty set. When no channel is available for a given user belonging to UE\textsubscript{type}(f\textsubscript{s,j}), the user is switched to UE\textsubscript{noch}(f\textsubscript{s,j}).

Once user and channel codes allocation are determined, the power allocated to transport channels is implicitly determined. Here we describe the power calculation for each channel:

a) $P_{FACH}$: We apply DPS to FACH, then its downlink transmission power level is different depending on the various cell coverage [18], i.e. the user distributions in UE\textsubscript{fach}(f\textsubscript{s,j}).

b) $P_{DCH}$: Equation 1 shows the total DCH transmission power required for $n$ users in a cell [19],

$$P_{DCHs} = P_p + \sum_{i=1}^{n} L_{p,i} \cdot \frac{P_{p} + x_i}{E_b/N_0 x_i \cdot R_{p,i}} + p$$

(1)

where $P_p$ is the power for common control channel, $P_n$ the background noise. $L_{p,i}$ is the path loss of user, $W$ is the bandwidth in UMTS environment, $R_{b,i}$ is the $i$th user transmit rate, $E_b/N_0$ is the target experienced signal quality of user. $p$ the orthogonality factor, $x_i$ is the intercell interference observed by the $i$th user, expressed by $x_i = \sum_{j=1}^{M} P_{r,j}$. $P_{T,j}$ is the transmission power in neighboring cell $c_j$ ($j = 1...M$). $L_{ij}$ is the path loss from $i$th user to the $j$th cell.

c) $P_{HS-DSCH}$: There are two options for HSDPA power allocation: RNC allocates a fixed amount of power for HSDPA transmission per cell, or Node B adjusts any unused power in cell for HSDPA. In this work we focus on the second method in order to provide only required amount of power to satisfy users in UE\textsubscript{hs}(f\textsubscript{s,j}). Equation 2 expresses the required transmit power to guarantee a minimum HS-DSCH throughput [20],

$$P_{HS-DSCH} \geq SINR \times \left[ p - G^{-1}\frac{P_{own}}{SF_{16}} \right]$$

(2)

Where $P_{own}$ is the own cell interference experienced by user, $SF_{16}$ the spreading factor. $G$ is the geometry factor defined by $G = \frac{P_{own}}{P_{th}}$, related with the user position. In the macro-cell (hexagonal layout with 1000 m base station spacing), users within 80% coverage experience a geometry factor of $-2.5$dB or better, within 95% a geometry factor at least $-5.2$dB [21]. With the target BLER and the channel quality information (CQI) from users, we obtain the Signal to Interference Noise Ratio (SINR) from the analytic formulation driven by link-level simulation results in [22]. The CQI is obtained through the target bandwidth of HS-DSCH and mapping table of MAC-hs Bit Rates versus CQI in [23]. Then $P_{HS-DSCH}$ is calculated by applying SINR and $G$ into Equation 2.

C. Decision Principles

User sets are selected according to flow level:

- $\cup UE(f_{s,j}) \cup UE_{dch}(f_{s,j}) \cup UE_{hs}(f_{s,j}) = Rt(f_{s,j})$.
- $Rt(f_{s,j}) = Dist(s), j = 0, 1$. To guarantee service coverage, all users of multicast group should be selected to receive $f_0$ or $f_1$, unless channel codes are saturated.
- $Rt(f_{s,j}) \leq Rt(f_{s,j-1}), j \geq 2$. The advanced flow is only sent to users which also receive lower flow.

Then, the repartition of users should be in accord with channel characteristics:

1) Considering FACH can be listened by all multicast users within its coverage. $\cup UE(f_{s,j})$ includes the nearest users in $Dist(s)$ (under a distance threshold $d_{thr}, d_{thr}$ is determined during optimization.

2) The other users in $Rt(f_{s,j})$, farther than $d_{thr}$, are assigned to $UE_{dch}(f_{s,j})$ or $UE_{hs}(f_{s,j})$.

3) UE(ch\textsubscript{m}, f\textsubscript{s,j}) $\cap$ UE(ch\textsubscript{n}, f\textsubscript{s,j}) = $\phi$, user sets for each flow must not overlap.

D. Optimization Strategies

At the beginning of estimation phase, RNC initializes an allocation solution for whole cell. Then it searches for the better solution based on simulated annealing (SA) algorithm. SA is chosen as it is simple to implement and adaptable to a variety of problems including telecommunications [24]. During the search procedure, iterations are performed until stopping criterion is met (i.e.temperature declines to zero). In each iteration, a new solution is generated by reselecting decision variables. We define a two-dimensional cost function to measure the quality of each solution, then reject or accept it based on a proposed evaluation criteria. When estimation phase stops, RNC starts resource allocation phase.

1) Initialization: We use MPC to initialize solution for each flow: users are served by a pure transmission mode which costs the minimum power. When initial power consuming is over budget, the farthest users for advanced flows will be rejected until a feasible solution is obtained.

2) Cost Function: We aim at finding optimum solution to guarantee the QoS requirement in terms of the bandwidth of allocated channels, and minimize the transmission power while avoiding power saturation. A cost function reflecting these two aspects is defined, to calculate the lost of throughput $Tb(c)$ and the power consumption $Po(c)$.

a) Throughput Optimization: The lost throughput of whole cell is expressed as follows.

$$Th(c) = \sum_{s_i \in S(c)} Th(s_i)$$

(3)

$$Th(s_i) = \sum_{f_j \in F(s_i)} \sum_{t_u \in Dist(f_{s,j})} \max([-\Delta_{s,u}, 0])$$

(4)

$\Delta_{s,u}$ is the bandwidth difference between allocated channel and required service for all users in multicast group. The channel bit rate is determined by its OVSF code(s) [17].

b) Power Optimization: MBMS transmission power in the same cell is calculated as following:

$$Po(c) = \sum_{s_i \in S(c)} Po(s_{c,i})$$

$$Po(s_{c,i}) = \sum_{f_j \in F(s_i)} \sum_{c_{hi}} P_{f,j,c_{hi}}$$

(5)

(6)
3) Evaluation Criteria: Once a new solution \( x'_i \) is generated by modifying the current solution \( x_i \), we evaluate \( x'_i \) in lexicographic order: \( x' \) is accepted when \( Th(x') = Th(x) \) and \( Po(x') \leq Po(x) \), or \( Th(x') < Th(x) \). Otherwise, to avoid being trapped on the local optima, a random value \( p \in (0, 1] \) is generated, and \( x' \) is accepted when \( p < \frac{1}{Th(x')} \).

IV. EXPERIMENTS AND RESULTS

We consider one cell in a hexagonal structure of 19 cells, where each Node-B covers three cells, and only multicast services are transmitted in this cell. Table I presents the system simulation parameters. The maximum power for MBMS transmission (19 w) in one cell is the total transmission power (43 dBm or 20 w) minus 30 dBm (1 w) for common channels.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular layout</td>
<td>19 Cells</td>
<td>Orthogonally factor</td>
<td>0.5</td>
</tr>
<tr>
<td>Node B Tx. power</td>
<td>43 dBm</td>
<td>Site to site distance</td>
<td>3 km</td>
</tr>
<tr>
<td>Common Channel power</td>
<td>30 dBm</td>
<td>Background noise</td>
<td>-100 dBm</td>
</tr>
<tr>
<td>Background noise</td>
<td>-100 dBm</td>
<td>COI's</td>
<td>CQI 1-6</td>
</tr>
<tr>
<td>Power of neighbor cell</td>
<td>37 dBm</td>
<td>Propagation models</td>
<td>Cost 231</td>
</tr>
</tbody>
</table>

### A. Experiment Scenarios

Benchmark is not existing since it is the first time that scalable transmission technology is combined with combinational channel assignment for MBMS. To access the performance of our algorithm, we implement F2R2M and competing allocation approaches on the same platform, they are “MBMS Power Counting” (MPC), “Dual Transmission Mode” (Dual Tx) and “Scalable FACH Transmission” (S-FACH). Besides, to prove the advantage of layer based channel allocation, we applied MPC for each flow (S-MPC). A comparative experiment is then conducted with following scenarios (Figure 2).

We create six problem instances with different traffic loads and user distributions. Two couples of scenarios (2s-30u-SS and 2s-50u-SN; 3s-80u-SNN and 3s-80u-SSN) have the same user and service setting, but service \( s_2 \) is transmitted as two flows of 64 kbps and one flow of 128 kbps respectively. For MPC and Dual Tx, services are transmitted in non-scalable mode. S-FACH allocates common channel for flow with fixed coverage [9]: 95% for \( f_1 \) (or \( f_0 \)), 50% for \( f_2 \) and \( f_3 \) 33%. The solutions of MPC, Dual Tx, S-FACH and S-MPC are determined based on minimum power consumption.

### B. Experimental Results

The experimental results in Table II are presented in two aspects: the lost of throughput transferred in percentage and the consumed power of all MBMS multicast services within one cell. For example in 1s-20u, 20 users request total bandwidth of 256 x 20 = 5120 kbps, S-FACH loses 25% (1280 kbps). Solutions with power less than 19w are feasible and emphasized in boldface.

When service transmission is non-scalable mode (i.e. MPC and Dual Tx), only 2s-30u can be transmitted through feasible solution with MPC, while Dual Tx saturates power, the reason is that it does not consider the utilization of HS-DSCH. Such inefficiency is confirmed in the allocation for 3s-80u-SNN (Table III), where MPC consumes less power than Dual Tx because users of \( s_2 \) and \( s_3 \) receive services through HS-DSCH. When traffic load is heavier, e.g.2s-50u and 3s-80u, MPC achieves saturated transmission power since it does not consider multimedia scalability.

S-FACH solves the power saturation problem of MPC for the first four scenarios. It reduces coverage for advanced flows hence consuming less power while providing service coverage (all service can be transmitted). However, when power is not saturated, such QoS sacrifice is unnecessary. For 2s-30u, both S-FACH and MPC solutions are feasible, S-FACH costs less power than MPC but loses more than half of bandwidth (58.3%) due to the smaller coverage for \( f_{s_1,2} \). Hence, according to our evaluation criteria, MPC is better than S-FACH for 2s-30u. Besides, when service demand is even higher (i.e. 3s-80u), S-FACH does not flexibly balance the service quality and power with the fixed flow coverage.
Therefore for the last two scenarios, it still achieves power saturation, which actually could be avoided by decreasing users for advanced flows.

The results of S-MPC reveal that scalable transmission costs less power than non-scalable scheme thus achieve feasible solution. From the results of S-MPC for 2s-50u-SN/SS and 3s-80u-SN/SS, with the same user distribution and total traffic load, the scalable transmission of S-MPC consumes less power. However, for scenarios having more users (2s-50u-SS and 3s-80u-SSN), S-MPC increases the possibility of channel codes saturation because it allocates only pure transmission mode for each flow, that may results huge consumption of channel code when DCH users are numerous.

Our algorithm outperforms the other algorithms for all scenarios. For 2s-30u, when conventional approaches could allocate radio resources properly, F2R2M consumes less power (47% of MPC solution) with coordinated QoS thanks to layered channel allocation. For 2s-50u-SN/SS, our algorithm avoids unneeded QoS decrease by flexibly allocating users for each flow.

### V. Conclusion

We present a flexible radio resource allocation algorithm for multimedia multicast service based on metaheuristic approach. In this model, we design a two-dimensional cost function that reflects both the service quality and radio resource consumption. Then a lexicographic-order evaluation criteria is proposed allowing to find the best solution satisfying the QoS requirement of multicast service and minimize the power consumption with limited radio resource. Experimental results show that our algorithm balances the power consumption and service quality by applying layered channel allocation, and reduces the possibility of radio resource saturation by adapting combinational channel assignment.

### REFERENCES