UDC 004.4

Optimal Multicast Subgrouping in Mobility-Aware 5G Systems: Challenges, Modeling, and Opportunities

Vitalii A. Beschastnyi[*], **Darya Yu. Ostrikova**[*],
Alexander I. Zeifman[†‡], **Irina. A. Gudkova**[*§]

[*] *Department of Applied Probability and Informatics*
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[†] *Department of Applied Mathematics*
Vologda State University
6 Orlova St, Vologda, 160000, Russian Federation
[‡] *Institute of Informatics Problems*
Federal Research Center "Computer Science and Control"
of the Russian Academy of Sciences
44-2 Vavilov St, Moscow, 119333, Russian Federation
[§] *Faculty of Electrical Engineering and Communication*
Brno University of Technology
3058/10 Technická, 61600 Brno, Czech Republic

Email: beschastnyy_va@rudn.university, ostrikova_dyu@rudn.university, a_zeifman@mail.ru,
gudkova_ia@rudn.university

The expected growth in the mobile video (including streaming video, video downloading, conferencing, etc.) now is the key driver for rapid development of 5G wireless network technologies. This paradigm forces wireless networks to manage their resources as effectively as possible. One of the most appropriate solutions for video traffic that may provide the sufficient spectral efficiency is multicasting. Multicast sessions allows a group of users to simultaneously access the same multimedia content with the same QoE parameters. In this paper we consider different methods of multicast subgrouping. An analytical framework in the from of a queuing network is proposed to study the network performance, and a comparative study of network performance metrics under different user movement scenarios is presented.

Key words and phrases: multicast, subgrouping, 5G, analytical model, optimization.

1. Introduction

According to [1], video will exceed the 82 percent of global Internet traffic by 2022. For network operators this leaves no other option but to strive for higher spectral efficiency as video data demands both high data rates and strict Quality of Experience (QoE) requirements. One of possible solutions is to employ the existing "enhanced" multimedia broadcast multicast services (e-MBMS) architecture [2], standardized by the Third Generation Partnership Project, by adopting it to the 5G systems. This will enable 5G networks with all the opportunities of the point-to-multipoint services which are crucial to support large-scale consumption of mass multimedia services on mobile devices [3].

In this paper we consider multicast approaches of network offloading. In [4] an approach based on splitting users into multicast subgroups (subgrouping) was proposed. This approach allows to achieve better results in terms of considered metrics rather then allocating the whole pool of resources to a single multicast group. The author also propose an algorithm that allows to find a sub-optimal network configuration consisting of two sets (groups) of users and resources allocation scheme for them. The algorithm considers maximization of Aggregate Data Rate (ADR) metric, thus improving the overall throughput of the network. Due to its low complexity the algorithm satisfies computation timing constraints since it should be run every Time Transmission Interval (TTI) during which channel quality between a multicast user and the base station is considered to be constant and may be just 1 ms long. Besides, for distant users it is extremely important to take into account their spatial positions as the amount of utilized resource significantly varies [15], mainly due to the effect of interference [16, 17].

An analytical model for characterizing the impact of the users' mobility in the multicast subgroup formation problem was proposed in [5] since queuing theory [7, 18] is typically employed in description of telecommunication networks . It is shown that, under the condition of constant number of multicast users within the service area, user mobility can be modeled as a queuing network. The model allows for close analytical calculation of average experienced bit rate and maximum aggregate data rate as an optimal solution of the subgroup formation problem stated in [4].

In this paper we study the range of applicability of the algorithm proposed in [4]. We also extend the model proposed in [5] to take into account the 'ON-OFF' (active-passive) operation mode of user devices. As ADR metric is usually unfair towards end users we also consider other metrics that may be more appropriate in different scenarios.

2. System Model

Each device in a multicast group transmits periodic or aperiodic feedback signals that allow base station (BS) to define Channel Quality Indicator (CQI). On collecting the state information from all the devices inside the coverage area, the BS makes a decision on subgrouping of these devices. As the time for taking a decision is limited by one Transmission Time Interval (TTI) an effective mechanism is required for resource allocation, which can operate within very short run time.

The simplest way to distribute multicast resources is to maintain equal data rate across all the devices. Among the many alternative solutions, the subgrouping method are given most of attention, as they can provide the highest throughput. Group-based methods consist of dividing a set of multicast objects into several subgroups consisting of receivers. Each subgroup is served with MCS that can be applied to the subgroup member with the worst channel conditions [12, 14]. Time-free and multi-layer coding techniques [6, 11] can be applied to subgroups to guarantee continuity and integrity of data transmission, and to efficiently handle subgrouping process.

The main problem in subgrouping is the computational complexity (and time) needed to find an optimal subgroup configuration (see Fig. 1). It is usually calculated using exhaustive search schemes. Run-time is critical in any wireless system where multicast

objects can experience channel fluctuations during the time interval required by BS to form subgroups according to the collected CQIs.

2.1. Subgroup Formation Problem

We refer to a single-cell coverage area with one multicast group, composed of set $A = \{1, \cdots, A\}$ of members. The BS allocates K RBs according to the Channel Quality Indicator (CQI) feedbacks [13]. Let $M = \{1, \cdots, M\}$ be the set of available CQIs, n_m be the number of members reported m-CQI, $\boldsymbol{n} = (n_1, \cdots, n_m)$. Since each CQI is associated with supported MCS, we indicate with $b_m, m \in M$ the data transmitted over one Resource Block (RB) by using the Modulation and Coding Scheme (MCS) corresponding to the m-th CQI. The BS enables the subgroup configuration, i.e., the number of subgroups with their relevant MCS and amount of allocated resources that maximizes the ADR. Let us denote $L = \{1, \cdots, L\}$, $L \in M$ the set of subgroups in the current moment, $M_l = \{m \in M : n_m > 0\}$ the set of CQIs which are reported by at least one user, and $l_m \in L, m \in M$ the affiliation of m-th CQI to l-th subgroup. Reported CQIs are divided into existing subgroups. Let k_l represent the number of resource blocks allocated to l-th subgroup, $\mathbf{k} = (k_1, \cdots, k_L)$

2.2. Mobility Model

The number of members and the values of CQIs reported by each member. Nevertheless, if we would like to know the situation not only at the current moment but also predict it in the future we could add stochastic nature in the problem formulation. In this paper we analyze the scenario of users that are moving within the circular cell of R radius . The directions the users are moving are learned from statistics of event-based simulation to find the probability that a user will change the area of the some MCS within a timeslot [5, 10]. A circular cell is divided in disjoint areas of the same CQI. We assume free space, so the areas have forms of annulus. Users are distributed according to a Poisson Point Process (PPP) on the plane and thus are uniformly deployed within a cell. Let us denote $\theta_{m,m-1}$ the probability of events when a user moves farther from the BS and changes the reported CQI from m to $m-1$; $\theta_{m,m+1} = 1 - \theta_{m,m-1}$ the probability of events when a user moves closer to the BS and changes the reported CQI from m to $m+1$; and α_m the average time when a user does not move from the area of m-MCS, i.e. he does not change its CQI from m.
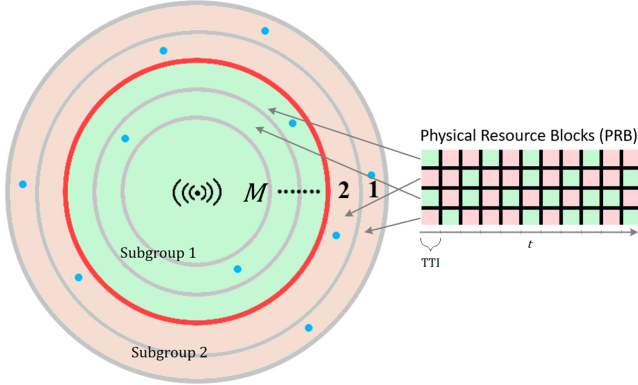
Figure 1. Subgroup configuration example

3.   Mathematical Model

In this section we first present a queuing network, where each node describes the channel quality level $m \in \mathrm{M}$, where M is the set of all channel quality indicators (CQIs). The considered network is closed, i.e. it is a network with a set of nodes without source and drain, in which a constant number N of similar arrivals are circulating [9].

3.1.   Case of Active Users

Earlier in [5] the queuing network was considered only with "active" users, that is, the subscribers which use the mobile network continuously. The transition digram for the considered queuing network is schematically presented in Fig. 2.
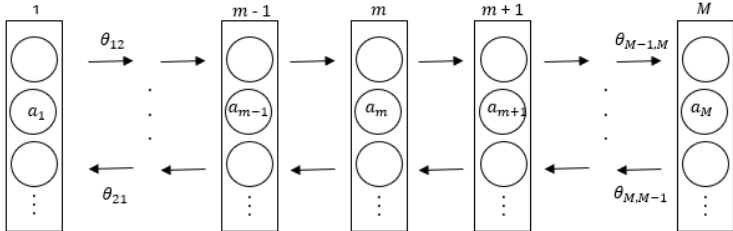

Figure 2. Transition diagram for queuing network with 'active' users only

The state space of the system $X(\mathrm{M}, N)$ satisfies the relation (1):

$$X(\mathrm{M}, N) = \{\boldsymbol{n} : n_i = 0, \cdots, N; \sum_{i=1}^{M} n_i = N\}. \tag{1}$$

In this case, the stationary probabilities $p(\boldsymbol{n})$ of the states of the system are calculated as (2).

$$p(\boldsymbol{n}) = G^{-1}(\mathrm{M}, N) \prod_{m=1}^{M} \frac{(h_m a_m)^{n_m}}{n_m!}. \tag{2}$$

where $G(\mathrm{M}, N) = \sum_{\boldsymbol{n} \in X(\mathrm{M},N)} \prod_{m=1}^{M} \frac{(h_m a_m)^{n_m}}{n_m!}$ and h_m, $m \in \mathrm{M}$ is calculated as (3).

$$\begin{cases} h_1 = h_2(1 - \theta_{12}); \\ h_m = h_{m-1}\theta_{mm-1} + h_{m+1}(1 - \theta_{mm+1}), 2 \le m \le M - 1; \\ h_M = h_{M-1}\theta_{MM-1}. \end{cases} \tag{3}$$

3.2.   Case of Active and Passive Users

Similarly, we may describe the model for devices operating in two modes (the 'ON-OFF' model), where the 0-node describes the 'passive' (or 'OFF') mode, when devices do not exchange the data with the BS (Fig. 3).

Here passive node is denoted as 0-node, and N is the total number of customers in the system in both active and passive nodes. θ_{nm}, $m = \overline{1, M}$ indicate the probabilities of transitions of active users between adjacent nodes, while θ_{0m}, where $m = \overline{1, M}$, indicate the probabilities of transitions from the passive node to the active ones. The average time until the user moves from the area m is similarly indicated with α_m. In order to prove that a given physical system exists, we show that passive node is in one of the available states, that is, satisfies the normalization condition (4):

$$\sum_{m=1}^{M} \theta_{0m} = 1. \tag{4}$$

The state space of the system $X(\mathrm{M}, N)$ satisfies the relations (1) and (5):

$$|X(\mathrm{M}, N)| := \binom{M + N - 1}{M - 1}, M \ge 0, N \ge 0. \tag{5}$$

In this case, the stationary probabilities $p(\boldsymbol{n})$ of the states of the system are calculated as (2), where h_m, $m \in \mathrm{M}$ is calculated as (6).

$$\begin{cases} h_1 = h_2\theta_{21} + h_0\theta_{01}; \\ h_m = h_{m-1}\theta_{mm-1} + h_{m+1}\theta_{mm+1} + h_0\theta_{0m-1}, 2 \le m \le M - 1; \\ h_M = h_{M-1}\theta_{MM-1} + h_0\theta_{0M}. \end{cases} \tag{6}$$
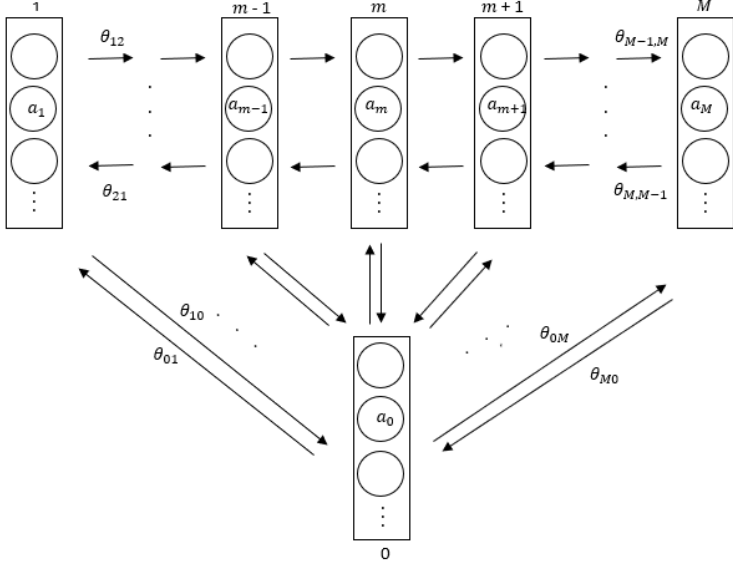
Figure 3. Transition diagram for queuing network with 'active' and 'passive' users

4.   Optimization Problems

The problem of calculating the subgroup configuration for a wireless multicast network can be considered in the context of two possible scenarios: in the first scenario, the total data rate is found in order to achieve maximum throughput of the whole network. In the second scenario, the rates are somehow evenly distributed among all devices. In this section, the main goal for scenario 1 is to analyze the range of applicability of the algorithm proposed in [4], while for the second scenario is to study the fairness of resource distribution between devices by the Jain's fairness index [8] for the considered utility functions.

4.1.   Aggregate Data Rate Maximization

Aggregate Data Rate shows the overall size of data that can be downloaded within one TTI by the devices what is extremely important for applications that involve on-line video streaming and thus require higher throughput.

To calculate the Aggregate Data Rate we use an objective function (7) and the algorithm proposed in [4].

$$U\big(\sigma = \{M_1, \cdots, M_L\}, \boldsymbol{k} = (k_1, \cdots, k_L)\big) =$$
$$= \sum_{l=1}^{L} \Big(b_{\min_{m \in M_l}(m)} \cdot k_l \cdot \sum_{m \in M_l} n_m\Big) := ADR \to max. \tag{7}$$

In Fig. 4 relative difference of results obtained with use of the considered algorithm and brute force analysis is shown. The results include 3 possible scenarios: uniform, centered and boundary trend of users' disposition.
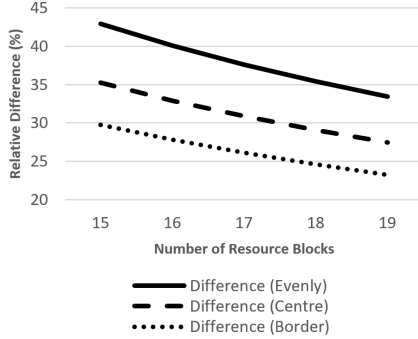

Figure 4. Relative difference between the ADR maximization algorithm and brute force analysis

The presented plot shows that the algorithm offers good results only for large number of RBs, which are applicable to wide bandwidth (15, 20 MHz), but might be ineffective in case of narrow bands (1.4 – 10 MHz).

4.2.   Fairness Spectrum Allocation

In order to equalize distribution of data rates among devices, we consider three objective functions: Fair Data Rate (FDR), Balanced Data Rate (BDR), and Proportional Fairness (PF) [6].

FDR tries to minimize difference in average rates (8):

$$U\big(\sigma = \{M_1, \cdots, M_L\}, \boldsymbol{k} = (k_1, \cdots, k_L)\big) = \max_{1 \le l \le L} \Big(b_{\min_{m \in M_l}(m)} \cdot k_l\Big) -$$
$$\min_{1 \le l \le L} \Big(b_{\min_{m \in M_l}(m)} \cdot k_l\Big) := FDR \to min. \tag{8}$$

BDR maximizes the minimal rate without taking into account higher rates (9):

$$U\big(\sigma = \{M_1, \cdots, M_L\}, \boldsymbol{k} = (k_1, \cdots, k_L)\big) =$$
$$= \min_{1 \le l \le L} \Big(b_{\min_{m \in M_l}(m)} \cdot k_l\Big) := BDR \to max. \tag{9}$$

PF maximizes the total network bandwidth and evenly distribute the rates (10):

$$U\big(\sigma = \{M_1, \cdots, M_L\}, \boldsymbol{k} = (k_1, \cdots, k_L)\big) =$$
$$= \sum_{l=1}^{L} \bigg(\log\big(b_{\min_{m \in M_l}(m)} \cdot k_l\big) \cdot \sum_{m \in M_l} n_m \bigg) := PF \to max. \tag{10}$$

Input data for numerical experiment

Parameter	Value
M	15
N	50
K	16
$\boldsymbol{n}^{\text{Uniform}}$	$(4, 3, 3, 3, 3, 4, 3, 3, 4, 3, 4, 3, 3, 4, 3)$
$\boldsymbol{n}^{\text{Centered}}$	$(1, 1, 1, 1, 1, 1, 13, 15, 10, 1, 1, 1, 1, 1, 1)$
$\boldsymbol{n}^{\text{Boundary}}$	$(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 10, 13, 15)$

Numerical results for data rate equalization functions

User Distribution Type	FDR	BDR	PF
Uniform	0.95	0.89	0.77
Centered	0.98	0.96	0.89
Boundary	0.75	0.96	1

We analyze the proposed functions using the Jane's fairness index, which is calculated as (11):

$$J = \frac{\left(\sum_{l=1}^{L} \left(b_{\min_{m \in M_l}}(m) \cdot k_l \cdot \sum_{m \in M_l} n_m \right) \right)^2}{N \cdot \sum_{l=1}^{L} \left(\left(b_{\min_{m \in M_l}}(m) \cdot k_l \right)^2 \cdot \sum_{m \in M_l} n_m \right)}, 0 \leq J \leq 1. \tag{11}$$

where '0' indicates that the rates are distributed poorly, and 1 in the best possible way in term of fairness.

For input data presented in Table 1 we present numerical results (Table 2). The obtained results show that in case of uniform and centered distribution of users over the service area, the FDR function allows to find the finest network configuration. But in case the users are distributed mainly by the cell's boundaries, it is much more effective to use PF function. If users behave in the way that they are always either closer to the center or to boundaries, the preferred decision should be to select BDR function.

5. Conclusions

In this paper we considered different methods of multicast subgrouping. We introduced an analytical model for multicast-enabled network cell that includes user mobility and in-off operation modes. We also provided the numerical analysis of the range of applicability for the ADR maximization algorithm. Finally, we presented three objective functions that allow for data rate equalization among users with different channel quality conditions, and provided the numerical analysis based on Jane's fairness index to reveal the scenarios that are applicable for the functions.

Acknowledgments

References

1. Cisco Visual Networking Index: Forecast and Methodology, 2017–2022.
2. 3GPP, TS 22.246, Multimedia Broadcast/Multicast Service (MBMS) user services // Stage 1 , Rel. 15, July 2018.
3. J. Montalban, P. Scopelliti, M. Fadda, E. Iradier, C. Desogus, P. Angueira, M. Murroni, and G. Araniti, Multimedia Multicast Services in 5G Networks: Subgrouping and Non-Orthogonal Multiple Access Techniques // IEEE Point-to-multipoint communications and broadcasting in 5G, March 2018.
4. G. Araniti et al., "A Solution to the Multicast Subgroup Formation Problem in LTE Systems," IEEE Wireless Commun. Letters, vol. 4, no. 2, April 2015, pp. 149–52.
5. D. Ostrikova, F. Rinaldi, V. Beschastnyi, I. Gudkova, L. Militano, G. Araniti, A. Iera, K. Samouylov, Analytical Model for Multicast Subgrouping in 5G-Mobile eMBMS Environment // 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2017, pp. 13-19.
6. A. Bobbio, M. Gribaudo, M. Telek, Analysis of large scale interacting systems by mean field method // IEEE computer society, `DOI10.1109/QEST.2008.47`, 2008.
7. G.P. Basharin, Yu.V. Gaidamaka, and K.E. Samouylov, Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Automatic Control and Computer Sciences, 47(2), pp. 62–69, 2013
8. T. Girici, C. Zhu, J.R. Agre, and A. Ephremides, Proportional Fair Scheduling Algorithm in OFDMA-Based Wireless Systems with QoS Constraints // Journal of Communications and Networks, Vol. 12, No. 1, February 2010.
9. K.E. Samouylov, I.A. Gudkova, and D.Y. Ostrikova, Modelling and performance analysis of multicast file repair in 3GPP LTE networks // Lecture Notes in Computer Science. Vol. 9247, pp. 383–392, 2015.
10. M.C. González, C.A. Hidalgo, and A.L. Barabási, Understanding individual human mobility patterns // Nature 453, June 2008.
11. 3GPP, TS 36.300, Evolved Universal Terrestrial Radio Access (EUTRA) and Evolved Universal Terrestrial Radio AccessNetwork (EUTRAN), Rel. 13, January 2016.
12. W. Rhee, and J. Cioffi, Increase in capacity of multiuser OFDM systems using dynamic subchannel allocation // Proc. IEEE 51st "Efficient frequency domain packet scheduler for point-to-multipoint transmissions in LTE networks Vehicular Technology Conference (VTC-Spring), May 2000.
13. T.P. Low, M.O. Pun, Y.W.P. Hong, and C.C.J. Kuo, Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks // IEEE Transaction on Wireless Communications, vol. 9, no.2, pp. 791-801, September 2009.
14. G. Araniti, I. Bisio, M. D. Sanctis, F. Rinaldi, A. Sciarrone, Joint Coding and Multicast Subgrouping Over Satellite-eMBMS Networks // IEEE Journal on Selected Areas in Communications, vol. 36, no.5, pp. 1004-1016, 2018.
15. V. Naumov, K. Samouylov, Analysis of multi-resource loss system with state-dependent arrival and service rates // Probability in the Engineering and Informational Sciences, 31 (4), pp. 413-419, 2017.
16. A. Samuylov, D. Moltchanov, Yu. Gaidamaka, S. Andreev, Ye. Koucheryavy, Random Triangle: A Baseline Model for Interference Analysis in Heterogeneous Networks // IEEE Transactions on Vehicular Technology, 65 (8), art. no. 7275184, pp. 6778-6782, 2016.
17. V. Begishev, R. Kovalchukov, A. Samuylov, A. Ometov, D. Moltchanov, Yu. Gaidamaka, S. Andreev, An analytical approach to SINR estimation in adjacent

rectangular cells // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9247, pp. 446-458, 2015.

18. Yu. Gaidamaka, A. Pechinkin, R. Razumchik, K. Samouylov, E. Sopin, Analysis of an MG1R queue with batch arrivals and two hysteretic overload control policies // International Journal of Applied Mathematics and Computer Science, 24 (3), pp. 519-534, 2014.