# Математическая теория телетрафика и сети телекоммуникаций

UDC 621.39:519.872.4

## An Analytical Model of Load Distribution Schemes in LTE Heterogeneous Networks

**T. V. Efimushkina, K. E. Samuylov**

*Telecommunication Systems Department*
*Peoples' Friendship University of Russia*
*Miklukho-Maklaya str., 6, Moscow, Russia, 117198*

The resource allocation problem in Orthogonal Frequency Division Multiple Access relay-enhanced heterogeneous cellular networks is studied. An analytical model of the downlink channel in discrete time is suggested. We derive and analyze various resource allocation algorithms. In order to evaluate the role of different resource allocation schemes we obtain blocking probabilities and other performance metrics of interest.

**Key words and phrases:** LTE-Advanced, OFDMA, relay station, analytical model, performance metrics.

## 1. Introduction

The evolving 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) standard is based on the Orthogonal Frequency Division Multiple Access (OFDMA), which is known for its high spectral efficiency and inherent robustness against frequency-selective fading [1, 2]. Moreover, the multihop relay concept has been introduced into LTE-Advanced based cellular networks to provide ubiquitous high-data-rate coverage. In this paper we consider the fixed infrastructure usage model of multihop relay stations (RSs), which is shown in Fig.1. Among the advantages of relay stations deployment are, first of all, the improvement of system capacity and coverage by dividing one long channel path into several shorter links and by offering alternative paths to users located in shadow areas. Secondly, taking into account the high costs of the Base Stations (BSs) due to the necessity of a very high density for provision of the sufficient coverage, the deployment costs of the cellular systems with relays decrease. Thirdly, RS does not need a wired backbone access [3]. All in all, the flexibility in relay positioning allows a faster network construction [4, 5].

Resource allocation in OFDMA relay-enhanced cellular networks has become a popular topic recently [6–8]. By introducing multihop relaying to OFDMA cellular networks, larger capacity and coverage can be expected; however, there are still lots of crucial issues that should be taken into account. Firstly, full compatibility with manufactured LTE based devices is essential. Secondly, resource allocation on different hops should be cooperated to avoid data shortage or overflow in relay nodes [3]. Thirdly, such technical issues as relay placement, resource allocation and scheduling are extremely important. The proper allocation of the channel resources can cooperatively reduce the wastes of radio resources, and thus, increase network efficiency. All in all, noting the high variation of the traffic distribution between the Mobile Stations (MSs), it is essential to analyze dynamic resource allocation schemes that assign different amounts of resources to RSs according to various traffic demands and topologies.

Before considering the resource assignment problem, the frame structure and various OFDMA-based resource allocation architectures are to be considered. Resources

Received 28$^{\text{th}}$ August, 2012.

This work was supported in a part by Russian Foundation for Basic Research (Grant №12-07-00108-a.

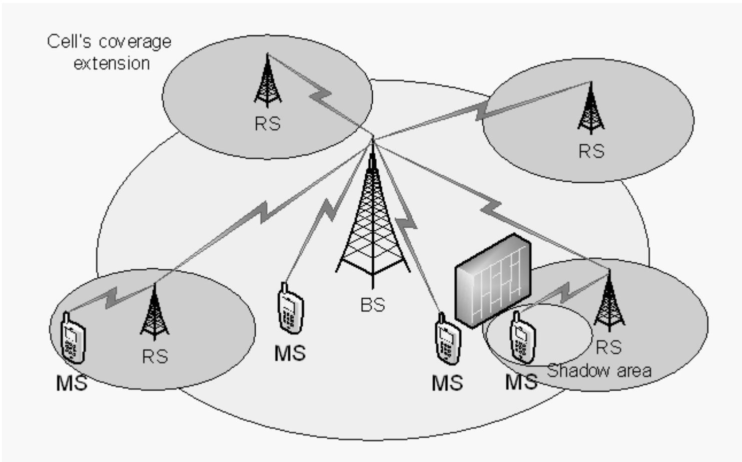


**Figure 1. The fixed infrastructure of relay-enhanced LTE-Advanced cellular networks**

in wireless communication systems usually refer to time, spectral and power. To address the downlink resource allocation problem in OFDMA relay-enhanced cellular networks we ignore the control messages and focus on the downlink data subframe shown in Fig. 2, which contains $S$ time slots in the time domain and $N$ subchannels in the frequency domain. The basic unit for data frame scheduling can be a slot which is a time-frequency unit represented by $(n, s)$ with $n \in \{1, 2, ..., N\}$ and $s \in \{1, 2, ..., S\}$. Slots in the first subframe can be assigned to transmissions in BS-RS and BS-MS links, whereas those in the second subframe can be allocated to transmissions in BS-MS and RS-MS links.

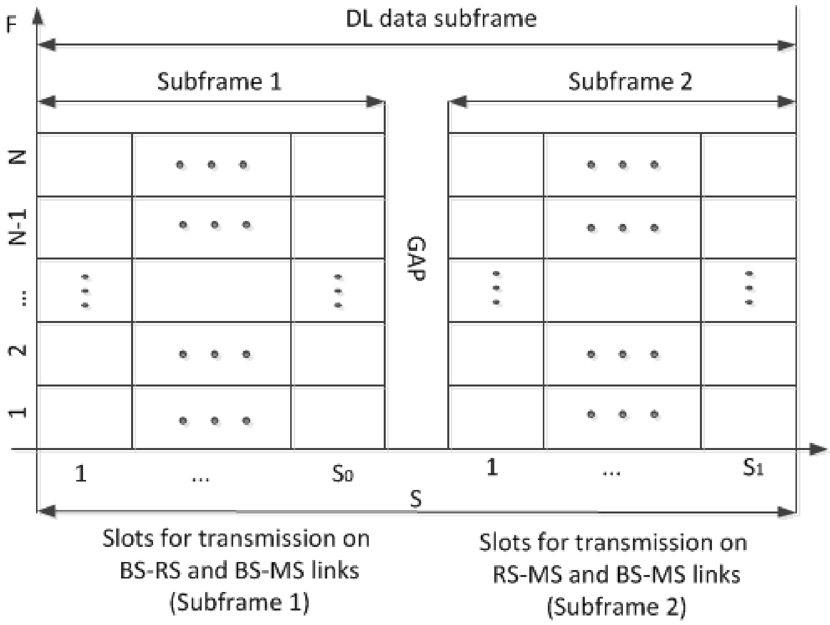


**Figure 2. Dowlink data subframe for OFDMA relay-enhanced networks**

There are two types of resource allocation architectures for relay-enhanced networks named centralized and semi-distributed architectures [3, 9, 10]. In centralized allocation, BS is responsible for allocating the available resources to all links. To perform efficiently, BS needs to be aware of the Channel State Information (CSI) of each link and the queue length on every RS. The centralized allocation can reduce the complexity of RSs, but it consumes more resources for control message exchange. In semi-distributed allocation, BS assigns each RS a number of slots, which are allocated their associated users by using its own scheduler. In this way, system overhead for information exchange between BS and RSs as well as the computational complexity of the BS are reduced.

Therefore, in this paper we formulate and solve the analytical model, which considers the downlink channel with semi-distributed architecture of the OFDMA-based LTE network. We analyze various resource allocation schemes, which take a slot as a scheduling unit. In the next section, the formulation of the proposed model is provided.

## 2. Analytical model of heterogeneous LTE-based network with two types of nodes: BS and RSs

In this section the analytical model of the heterogeneous LTE-Advanced network, with two types of nodes: BS and RSs, is formulated for downlink channel. Various channel resource assignment schemes are analyzed, among which the proportional one is the candidate which allows to improve the bandwidth utilization. Performance measures that enable the evaluation of the proposed resource allocation algorithms are derived.

### 2.1. Description of the model

We consider semi-distributed resource allocation scenario of a cellular network with one BS and $K$ RSs, $K < \infty$. The downlink subframe shown in Fig.2 is divided into $S$ channels (Chs), which occupy the smallest unity of channel bandwidth both in frequency and time bands. All the $S$ Chs are distributed between the BS and $K$ RSs to transmit the packets in the direction of MSs. In the model description we frequently refer to the request, which has a physical meaning of a packet. Let us suppose that there are $K+1$ types of requests in the cell. Here, $k$-request corresponds to the $k$-type, which is to be transmitted to the MS in the coverage area of the BS in case $k = 0$, or in the coverage area of the $k$-RS ($\mathrm{RS}_k$), $k = \overline{1, K}$. The arriving requests at the BS and RSs are stored in the buffers, the capacity of which for BS is $r_0, r_0 < \infty$ and for $\mathrm{RS}_k$ corresponds to $r_k$, $r_k < \infty$, $k = \overline{1, K}$. Moreover, we presume that the arrived requests at the queuing system with fully occupied buffers are lost and do not influence its functioning.

The system functions in discrete time measured in time slots with the constant length $h$, which equals the duration of the subframe in LTE network. Assume that all the changes in the system occur at time moments $nh$, $n = 1, 2, ....$ Therefore, during the slot $n$ , or the time interval $[nh, (n + 1)h)$, we assume that the following events take place in sequential order:

- the requests are serviced by the Chs of $\mathrm{RS}_k$ and the space they took in the buffers of $\mathrm{RS}_k$ is released;
- the requests are serviced by the Chs of the BS and the corresponding buffer space is released;
- the serviced $k$-requests arrive to the buffers of the corresponding $\mathrm{RS}_k$, $k = \overline{1, K}$, while the space they took in the buffers of BS is released;
- new requests arrive to the buffer of the BS;
- all the $S$ Chs are reallocated between the BS and RSs;
- the state fixing.

Let $\eta_n$, $\eta_n \in \{0, 1\}$ is a number of group arrival of requests during the slot $n$, taking into account that all of the $\eta_n$, $n \geqslant 0$, — are independent identically distributed random variables (RVs) with a generation function (GF):

$$A(z) = Mz^{\eta_n} = 1 - a + az, \quad | z | \leqslant 1, \quad a = P\{\eta_n = 1\}, \quad 0 < a < 1, \quad n \geqslant 0.$$

The number of requests $\chi_n$ in the arrived group is an independent from $n$ RV with the GF:

$$G(z) = Mz^{\chi_n} = \sum_{i\geqslant 1} g_i z^i, \quad |\, z \,|\leqslant 1, \quad G(1) = 1, \quad g_i = P\{\chi_n = i\}, \quad n \geqslant 0.$$

Thus, the arrival requests follow a group geometric distribution $\mathrm{Geom}^G$ as the time duration between the group arrivals conforms a geometric distribution with the mean $1/a$ and is characterized by the GF:

$$A(G(z)) = 1 - a + aG(z) = \sum_{i\geqslant 1} a_i z^i, \quad |\, z \,|\leqslant 1, \quad a_0 = \overline{a} = 1 - a, \quad a_i = ag_i, \quad i \geqslant 1.$$

Assume that each request from the arrival group belongs to type $k$ ($k$-request) with the probability $c_k$, $k = \overline{0, K}$, $c_\bullet = 1$. Here and further, dot in place of index means a full sum of the variable. All in all, the arrival rate can be characterized by the $(K + 1)$ — dimensional group geometrical distribution. In order to simplify the modeling, the service time is considered to follow the deterministic law with the duration of the request's service equal to one slot. Hence, every request is serviced during one slot, and releases the occupied buffer space. The described system can be denoted in the following way: $[\mathrm{Geom}^G_{K+1} \mid D = 1 \mid \mathrm{dif}(s_\bullet) = S \mid \overline{r}$. The notation $\mathrm{dif}(s_\bullet) = S$ indicates that the number of Chs varies from slot to slot and constitutes to the overall $S$ Chs, which allows to investigate different resource allocation schemes. The structure of the designed system is presented in Fig. 3.

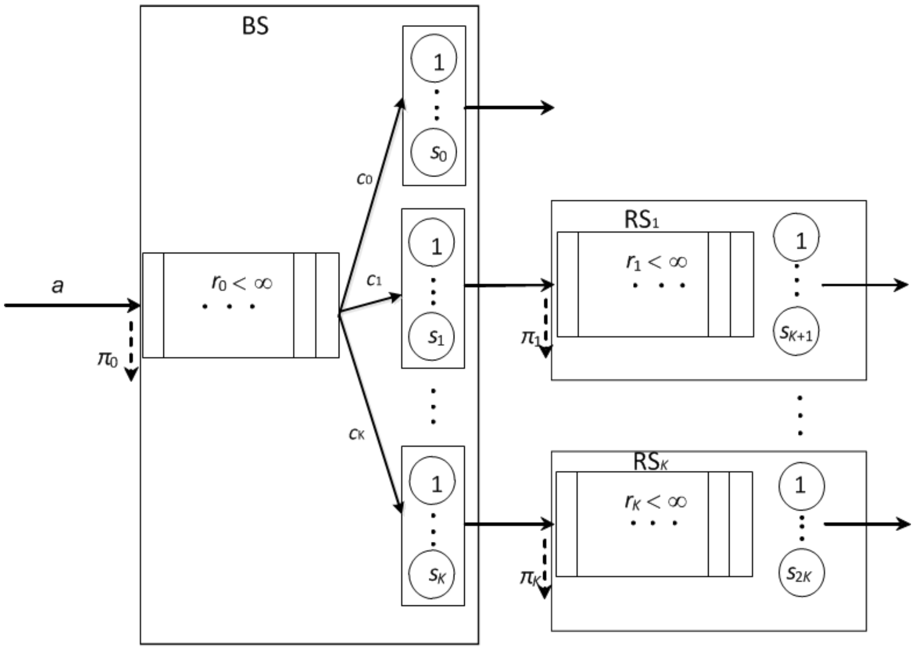


**Figure 3. The structure of the model with overall $S$ Chs, which are distributed between BS ans RSs, and $K + 1$ types of incoming requests**

### 2.2. Balance equations

The requests in the buffer of the BS do not differentiate in types, whereas the type selection is conducted using the polynomial scheme with the probabilities $c_0, ..., c_K$. Note that this preposition is valid for the analyzed cellular network with deterministic resource allocation schemes and also the systems, which consider only the total number of packets in the buffer of the BS without differentiation in types.

The functioning of the network system is described by the homogeneous Markov chain $\xi_n$ at time moments $nh + 0$, $n \geqslant 0$, with the state space:

$$X = \{\vec{x} = (x_0, x_1, ..., x_K)^{\mathrm{T}} : x_k = \overline{0, r_k}, \quad k = \overline{0, K}\}, \quad |\, X \,| = \prod_{k=0}^{K} (r_k + 1),$$

where $x_k$ — is a number of $k$-requests in the buffer of the corresponding BS or $RS_k$. In order to take into consideration the resource distribution at slot $n$, it is essential to input the vector $\vec{s}^n = (s_0^n, s_1^n, ..., s_{2K}^n)^{\mathrm{T}} = (f_0^n(\vec{x}), f_1^n(\vec{x}), ..., f_{2K}^n(\vec{x}))^{\mathrm{T}} = \vec{f}^n(\vec{x})$, where $\vec{f}^n(\vec{x})$ — is a function, which defines the strategy of resource distribution, and the values of which depend on the system state at slot $n$.

The function $\vec{f}^n(\vec{x})$ can indicate the following resource allocation algorithms (assume that $\vec{f}^n(\vec{x}) = \vec{f}(\vec{x}), n \geqslant 0$):

1. Deterministic 1 (do not depend on $\vec{x}$):

$$s_k = \left\lfloor \frac{S}{2K+1} \right\rfloor, \quad k = \overline{1, 2K}, \quad s_0 = S - \sum_{k=1}^{2K} s_k.$$

2. Deterministic 2 (do not depend on $\vec{x}$) is investigated in [3]:

$$s_k = \left\lfloor \frac{S - \lfloor \frac{S}{2} \rfloor}{K+1} \right\rfloor, \quad k = \overline{1, K}, \quad s_{K+k} = \left\lfloor \frac{\lfloor \frac{S}{2} \rfloor}{K} \right\rfloor, \quad k = \overline{1, K-1},$$

$$s_0 = S - \left\lfloor \frac{S}{2} \right\rfloor - \sum_{k=1}^{K} s_k, \quad s_{2K} = \left\lfloor \frac{S}{2} \right\rfloor - \sum_{k=1}^{K-1} s_{K+k}.$$

3. Proportional:

$$s_{K+k} = \left\lfloor \frac{x_k S}{x_\bullet} \right\rfloor, \quad S' = S - \sum_{k=1}^{K} s_{K+k},$$

$$s_k = \left\lfloor \frac{\lfloor x_0 c_i \rfloor S'}{x_0} \right\rfloor, \quad k = \overline{1, K}, \quad s_0 = S' - \sum_{k=1}^{K} s_k.$$

If $0 < a < 1$ the Markov chain $\xi_n, n \geqslant 0$ is aperiodic, and there exists a stationary probability distribution $[\vec{x}], \vec{x} \in X$.

To simplify the target equation, let us input the notations with short descriptions, presented below:

1. $s_k^{\min} = \min(x_k, s_{K+k})$ — a number of the serviced requests.

2. $s_\bullet^{\min} = \sum\limits_{k=1}^{K} s_k^{\min}$ — an overall number of the serviced requests by the Chs of $\mathrm{RS}_1$,...,$\mathrm{RS}_K$.

3. $s_0^{\min} = \min(s_0, x_0 - s_\bullet^{\min})$ — a number of the serviced 0-requests at the BS.

4. $r_0' = x_0 - s_\bullet^{\min} - s_0^{\min}$ — a number of requests left in the buffer of the BS after the 0-requests are serviced, taking into account that the requests, which left the Chs of $\mathrm{RS}_1$,...,$\mathrm{RS}_K$ arrive at the same quantity to their buffers from the corresponding Chs of the BS.

5. $n_k = \min\left(s_k - s_k^{\min}, r_0' - \sum\limits_{i=1}^{k-1} \delta(x_i, r_i) n_i\right)$ — a number of requests serviced at the BS and headed for the $\mathrm{RS}_k$, but which do not have any free space in the buffers of the $\mathrm{RS}_k$. This particular case is considered when the number of $k$-requests coincide with the buffer's capacity.

6. $s_{0,q}^{\min} = \min\left(s_0, x_0 - \sum\limits_{k=1}^{K}\left(s_k^{\min} - q_k\right)\right)$ — a number of the serviced 0-requests during the slot, taking into account the variable $q_k$, which takes both negative and positive values.

7. $r_{0,q}' = x_0 - \sum\limits_{k=1}^{K}\left(s^{\min}{}_k - q_k\right) - s_{0,q}^{\min}$ — a number of requests left in the buffer of the BS after the 0-requests are serviced, taking into account that the requests, which left the Chs of $\mathrm{RS}_1$,...,$\mathrm{RS}_K$ arrive at the quantity $q_k$ to their buffers from the corresponding Chs of the BS.

8. $n_{k,q} = \min\left(s_k - s_k^{\min} + q_k, r_{0,q}' - \sum\limits_{i=1}^{k-1} \delta(x_i, r_i) n_{i,q}\right)$, $k = \overline{1, K}$ — a number of requests serviced at the BS and headed for the $\mathrm{RS}_k$, but which do not have any free space in the buffers of the $\mathrm{RS}_k$, taking into account that $x_k$ changes according to $q_k$. This particular case is considered when the number of $k$-requests coincide with the buffer's capacity.

9. $\Omega_0 = \{\vec{x} : x_k \leqslant s_{K+k}, k = \overline{1, K}\}$ for $\forall \vec{x} \in X \backslash \vec{0}$ — a state space, which satisfies the condition: the number of requests in the buffer of the $\mathrm{RS}_k$ does not exceed the number of allocated Chs. Note that the Chs at the BS are not taken into account.

10. $\Omega_1 = \left\{\vec{x} : x_0 \geqslant s_\bullet^{\min}, s_k \geqslant s_k^{\min}, k = \overline{1, K}\right\}$ for $\forall \vec{x} \in X \backslash \vec{0}$ — a state space, which satisfies the condition: the number of allocated Chs at the BS for the requests that head to the $\mathrm{RS}_k$ exceeds the number of serviced requests at $\mathrm{RS}_k$. The following condition guarantees that the number of serviced requests at $\mathrm{RS}_k$ will be able to occupy the released buffers.

11. $\Omega_2 = \left\{\vec{x} : x_0 \geqslant \sum\limits_{k=1}^{K}\left(s_k^{\min} - q_k\right) \geqslant 0\right\}$, $q_k = \overline{-x_k, r_k - x_k}$, $k = \overline{1, K}$ for $\forall \vec{x} \in X \backslash \vec{0}$ — a state space, which satisfies the condition: the number of allocated Chs at the BS for the requests that head to the $\mathrm{RS}_k$ exceeds the number of serviced requests at $\mathrm{RS}_k$, taking into account the variable $q_k$.

12.
$$a' = \begin{cases} \sum\limits_{i=1}^{\infty} a_{s_0^{\min} + s_\bullet^{\min} + \sum\limits_{k=1}^{K} \delta(x_k, r_k) n_k + i}, & \text{if } \delta(x_0, r_0) = 1 \\ a_{s_0^{\min} + s_\bullet^{\min} + \sum\limits_{k=1}^{K} \delta(x_k, r_k) n_k}, & \text{otherwise} \end{cases}$$

— the probability of the arrival of the number of requests to the buffers of the BS during the slot to preserve exactly the same state of the system

13.
$$a_q' = \begin{cases} \sum\limits_{i=1}^{\infty} a_{s_{0,q}^{\min} + \sum\limits_{k=1}^{K}(s_k^{\min} - q_k) - q_0 + \sum\limits_{k=1}^{K} \delta(x_k, r_k) n_{k,q} + i}, & \text{if } \delta(x_0, r_0) = 1 \\ a_{s_{0,q}^{\min} + \sum\limits_{k=1}^{K}(s_k^{\min} - q_k) - q_0 + \sum\limits_{k=1}^{K} \delta(x_k, r_k) n_{k,q}}, & \text{otherwise} \end{cases}$$

— the probability of the arrival of the number of requests to the buffers of the BS during the slot to preserve the state of the system, taking into account the change $q_k$.

Therefore, the steady-state probability distribution $[\vec{x}], \vec{x} \in X$ is found from the balance equations [11, 12]:

$$a[\vec{0}] = \overline{a} \sum_{\Omega_0} c_0^{x_0} [\vec{x}],$$

$$1 - \sum_{\Omega_1} c_0^{s_0^{\min}} \prod_{k=1}^{K} c_k^{s_k^{\min} + \delta(x_k, r_k)(1+\ldots+n_k)} a'[\vec{x}] =$$

$$= \sum_{\Omega_2} c_0^{s_{0,q}^{\min}} \prod_{k=1}^{K} c_k^{s_k^{\min} - q_k + \delta(x_k, r_k)(1+\ldots+n_{k,q})} a'_q \left[\vec{x} + \sum_{k=0}^{K} q_k \vec{e}_k\right],$$

and normalizing equation

$$\sum_{\vec{x} \in X} [\vec{x}] = 1,$$

where $\delta(a,b) = \begin{cases} 0, & a \neq b; \\ 1, & a = b. \end{cases}$

### 2.3. Performance measures

The blocking probability as well as other performance characteristics can be obtained from the steady-state probability distribution.

1. The blocking probability $\pi_k$ of the k-requests, $k = \overline{0, K}$:

$$\pi_0 = \sum_{g = r_0 - r'_0 + 1}^{\infty} a_g \sum_X [\vec{x}], \quad \pi_k = \sum_{s'_k = r_k - x_k + s_k^{\min}}^{n_k} c_k^{s'_k} \sum_{\vec{s}: s_k > r_k - x_k + s_k^{\min}} [\vec{x}], \quad k = \overline{1, K}.$$

2. The overall blocking probability $\pi$ of the requests: $\pi = 1 - \prod_{k=0}^{K} \pi_k$.
3. The mean number $N_k$ of the $k$-requests in the system, $k = \overline{0, K}$: $N_k = \sum_X x_k [\vec{x}]$.
4. The mean number $N$ of the requests in the system: $N = N_\bullet$.
5. The mean number $S_k^{\mathrm{mean}}$ of the Chs in the system:

$$S_k^{\mathrm{mean}} = \sum_X \left\lfloor \frac{x_k S}{x_\bullet} \right\rfloor [\vec{x}], \quad k = \overline{1, K},$$

$$S_0^{\mathrm{mean}} = \sum_X \left( \left( S' - \sum_{k=1}^{K} s_k \right) + \sum_{k=1}^{K} \left\lfloor \frac{x_k S'}{x_\bullet} \right\rfloor \right) [\vec{x}], \quad S' = S - \sum_{k=1}^{K} s_{K+k}.$$

6. The mean number $U_k$ of the serviced $k$-requests during a slot:

$$U_k = \min\left(N_k, S_k^{\mathrm{mean}}\right), \quad U_0 = \min\left(N_0, S_0^{\mathrm{mean}}\right).$$

7. The mean spending time $T_k$ at the BS:

$$T_0 = \frac{N_0}{a_0(1 - \pi_0)}, \quad a_0 = a \sum_{i=1}^{\infty} i g_i.$$

## 3. Conclusion

In this paper we investigate the resource allocation problem in OFDMA relay-enhanced heterogeneous cellular networks by means of analytical modeling. We consider various resource allocation algorithms. In order to evaluate the role of different resource allocation schemes we derive the blocking probabilities and other performance metrics of interest.

## References

1. Long Term Evolution (LTE): Overview of LTE Air-Interface. White Paper. — Motorola, 2007.
2. REPORT ITU-R M.2134. Requirements Related to Technical Performance for IMT-Advanced Radio Interface(s). — 2008.
3. *Wang L.* Resource Allocation in OFDMA Relay-Enhanced Cellular Networks: Ph.D. thesis / School of Multidisciplinary Sciences The Graduate University for Advanced Studies (SOKENDAI). — 2010.
4. A Survey on 3GPP Heterogeneous Networks / A. Damnjanovic, J. Montojo, W. Yongbin et al. // IEEE Wireless Commun. — 2001. — Pp. 10–21.
5. *Osman A., Mohammed A.* Performance Evaluation of a Low-Complexity OFDM UMTS-LTE System. — 2008.
6. Performance Analysis of Reuse Partitioning-Based Subchannelized OFDMA Uplink Systems in Multicell Environments / E. Oh, M.-G. Cho, S. Han et al. // IEEE Transactions on Vehicular Technology. — 2008. — Pp. 2617–2621.
7. *Wongand I. C., Evans B. L.* Optimal Downlink OFDMA Resource Allocation with Linear Complexity to Maximize Ergodic Rates // IEEE Transactions on Wireless Communications. — 2008. — Pp. 962–971.
8. *Wengand L., Murch R. D.* Cooperation Strategies and Resource Allocations in Multiuser OFDMA Systems // IEEE Transactions on Vehicular Technology. — 2009. — Pp. 2331–2342.
9. *Kwak R., Cioffi J.* Resource-Allocation for OFDMA Multi-Hop Relaying Downlink Systems // Proc. IEEE Globecom. — 2007.
10. *Kaneko M., Popovski P.* Adaptive Resource Allocation in Cellular OFDMA Systems with Multiple Relay Stations // Proc. IEEE VTC-Spring. — 2007.
11. *Башарин Г. П.* Лекции по математической теории телетрафика. — 2009. [*Basharin G. P.* Lekcii po matematicheskoyj teorii teletrafika. — 2009. ]
12. *Iversen V. B.* Teletraffic Engineering Handbook // ITU-D, SG 2 Q 16/2. — 2008.

УДК 621.39:519.872.4

## Аналитическая модель схем распределения нагрузки в сетях LTE с разнородными узлами связи

**Т. В. Ефимушкина, К. Е. Самуйлов**

*Кафедра систем телекоммуникаций*
*Российский университет дружбы народов*
*ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Проведено исследование технологии беспроводного доступа LTE на основе ортогонального частотного мультиплексирования OFDM. Разработана и проанализирована аналитическая модель функционирования гетерогенной сети LTE для нисходящего канала. Получены формулы для основных ВВХ функционирования модели гетерогенной сети.

**Ключевые слова:** LTE-Advanced, OFDMA, ретрансляционная станция, аналитическая модель, вероятностно-временные характеристики.