UDC 004.728.3.057.4
doi:10.31799/1684-8853-2023-2-39-50
EDN: EHYAOZ

Articles

Comparison of the ways to reduce energy costs in stable massive machine-type communication systems

A. A. Burkov[a], *Assistant Professor, orcid.org/0000-0002-0920-585X, a.burkov@k36.org*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *With the development of modern communication standards, various scenarios of the Internet of things, including massive machine-type communication, are actively considered. The main requirements for such systems are a stable operation with a potentially infinite number of devices and low energy costs. Therefore, it is an urgent task to study different ways to achieve these requirements.* ***Purpose:*** *To explore and compare the effectiveness of different approaches to reducing energy costs in systems with a potentially unlimited number of user devices.* ***Results:*** *We describe a basic system model based on an ALOHA-type algorithm for the scenario of massive machine-type communication in a channel with additive white Gaussian noise. We focus on the two ways to reduce energy consumption for the system under consideration. The first way is to use hybrid automatic repeat request methods. The second way is to use a conflict resolution algorithm with a dynamic schedule. For the basic system and the proposed methods, we carry out an analysis of the lower bounds for the signal power and costs energy-per-bit under the condition of the stable operation of the system and specified requirements for spectral efficiency. It is shown that both methods make it possible to reduce energy costs compared to the basic system. At the same time, the dynamic scheduling algorithm allows to most effectively reduce the signal power from 4.5 dB at any spectral efficiency. For energy costs per bit at spectral efficiency values less than 0.42 bit/s/Hz, hybrid automatic repeat request is most effective. However, for large values, the greatest gain comes from changing the conflict resolution algorithm.* ***Practical relevance:*** *The results obtained allow us to evaluate the potential of reducing energy costs in stable systems with a large number of user devices through the use of hybrid automatic repeat request and a conflict resolution algorithm with a dynamic schedule.*

Keywords *— massive machine-type communication, internet of things, hybrid automatic repeat request, random multiple access algorithms, spectral efficiency, energy efficiency.*

For citation: Burkov A. A. Comparison of the ways to reduce energy costs in stable massive machine-type communication systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 39–50. doi:10.31799/1684-8853-2023-2-39-50, EDN: EHYAOZ

Introduction

As part of the currently unfolding 5G communication standard, as well as in the development of the next generation of 6G, Internet of Things (IoT) scenarios are actively considered [1, 2]. However, some of the most important requirements cannot be fully met within the current versions of 5G networks. The further development of the Internet of Things systems by 2030 will entail new, more stringent requirements for wireless communications in general and for the IoT in particular.

Taking into account the requirements for systems (transfer rate, delay, etc.), the following types of IoT are distinguished [3]: critical IoT, broadband IoT, industrial IoT, massive IoT.

Mass IoT is supposed to operate a huge number of simple devices that rarely transmit small amounts of data, usually powered by batteries. This system operates within a massive machine-type communication (mMTC) scenario. This scenario describes data acquisition systems with a large number of low power end devices (such as sensors) that periodically transmit a small amount of data. Examples of the mass IoT are temperature, pressure, light sensors and meters in smart home technology. The volume of transmitted data is small, but the number of IoT devices is very large [4]. The main requirements for this scenario are [5–9]:

— a large number of devices;

— stable work of the network at high values of the intensity of the appearance of messages in the system;

— low power consumption (in order to increase the service life of the device from an autonomous power source without additional maintenance);

— delivery of messages with a given reliability;

— limitation on average delay.

The paper will consider the scenario of massive machine-type communications. In the near future, IoT devices are expected to be on the order of a million per square kilometer. In addition, each of these devices will periodically transmit a small amount of data [10]. In view of the large number of devices for the scenario under consideration, it is impossible to use methods based on the static sharing of access to a common channel resource. Therefore, it is supposed to use random multiple access algorithms [11, 12].

To support emerging system requirements, it is necessary to improve existing approaches in network design. A separate task is to increase the reliability of systems by changing the structure and redundancy [13, 14]. To solve the problem of improving the energy efficiency of stable systems, two ways can be considered. The first way is to improve the methods of modulation and demodulation and the methods of error-correcting coding consistent with them [15]. For example, the use of new code modulation schemes or the use of methods for combining error-correcting coding and retransmissions, such as hybrid automatic repeat request (HARQ). The second way is to change approaches to planning and resource allocation, that is, modifying random access methods, including changing conflict resolution algorithms in the system. The use of these methods can increase the maximum of the input arrival rate to which the system operates stably, as well as reduce energy costs, which is important for devices with an autonomous power source.

In view of the foregoing, it is relevant to analyze the effectiveness of using hybrid automatic repeat request methods [16], as well as a blocked conflict resolution algorithm with a dynamic schedule [17, 18] in order to reduce the energy costs of random access systems with a potentially unlimited number of user devices.

System model

Let us introduce a system of assumptions for the system under consideration.

Assumption 1. Messages of all user devices have the same length. Each message is transmitted during the same time, let's take this time as a unit of time (slot). All system working time is divided into slots. The user devices know exactly the beginning of each slot and can only transmit at the beginning of the next slot. Slots are not pinned to user devices. The number of messages appearing in the system in each slot is distributed according to the Poisson law with the parameter Λ [messages/slot].

Assumption 2. Consider a time-continuous communication channel with additive white Gaussian noise, defined as $Y = \sum_{i=1}^{K_t} X_i + Z$ where Y is the output signal from the channel; X_i — user device signal with number i and power P; Z is additive white Gaussian noise, with zero mean and unit variance ($Z\sim(0, 1)$); K_t is the number of user devices that transmitted a signal in the slot with the number t, this value is random. Such a model corresponds to a Gaussian multiple access channel. One of three events can happen in each slot:

"Success". For a random multiple access system without HARQ methods, only one user device transmits in the slot. For a system with HRAQ methods, only one user device transmits in the slot, as well as when two or more users transmit and the collision is successfully resolved (users who transmitted at the same time affect the probability of successful message decoding due to additional message processing on the receiving side).

"Conflict". Two or more user devices transmit in one slot. And also when using HARQ methods — the collision was not resolved.

"Empty". No one passes in the window.

Remark 1. Since the noise variance in the considered system models is equal to 1, then the signal-to-noise ratio (SNR) in times is equal to $\frac{P\,[\mathrm{W}]}{1\,[\mathrm{W}]}$. Thus, the numerical value of the signal power P in watts coincides with the SNR in times. In what follows, P will mean SNR in times.

Assumption 3. A user device that has a message ready for transmission is said to be active. The number of users in the system is infinite (messages and users in the system are identical). Each active device sends a message to the channel with probability $p = \frac{1}{M}$, where M is the number of active users known to the base station and other devices. The same message will be retransmitted until it is successfully transmitted.

Assumption 4. On the transmitting side, a checksum is added to all transmitted messages, after which the received data undergoes error-correcting coding and is transmitted to the channel, where noise is added to the signal. At the receiving side, error correction decoding is performed first, and then the checksum is checked. If the checksum is correct, a acknowledge is sent, if the checksum is incorrect, a negative-acknowledgement is sent. When the user receives an acknowledgement, the transmitter assumes a "Success" event occurs on the channel, when the user receives a negative-acknowledgement — "Conflict". It is assumed that the receipt is received by the transmitter in the same slot in which the message was transmitted and cannot contain errors.

Assumption 5. Transmission is carried out using an error-correcting code of sufficient length with a code rate $R = W \log_2(1 + P) - \varepsilon$, where $\varepsilon \to +0$.

Remark 2. According to Shannon's theorem (direct coding theorem), there is such a code-modulation scheme with speed R such that the decoding error probability P_e can be arbitrarily small $\to 0$ if the inequality $R < C = W \log_2(1 + P)$ is satisfied. If this inequality is not satisfied, then according to Shannon's theorem (inverse coding theorem), the error probability $\to 1$ [19].

Analysis the systems will consider signal strength and energy cost per bit for the required spectral efficiency T [bit/s/Hz].

Analysis of the basic random multiple access system

Under the basic system is meant the operation of the ALOHA-based random access algorithm in a noisy channel.

Let's introduce the concept of signal-to-interference-plus-noise ratio (SINR), which is defined as

$$\gamma_{i,t} = \frac{\alpha_{i,t} E_i}{\sum\limits_{\substack{j=1 \\ j \neq i}}^{K_t} \alpha_{j,t} E_j + N_0}, \tag{1}$$

where $\gamma_{i,t}$ is the SINR for the i-th user device in slot number t; $\alpha_{i,t}$ is the signal energy loss in the channel of the i-th user device in slot number t; E_i is the signal energy of the i-th user; K_t is the set of user devices transmitting in the slot t; N_0 is the noise energy.

Taking into account the definition of SINR, we can formulate the following statement.

Statement 1. The SINR ratio in the considered systems can be calculated as follows:

$$\gamma_{i,t} = \frac{1}{(K_t - 1) + P^{-1}},$$

where P is the signal strength.

Proof: The model of the system without attenuation and Rayleigh fading is considered, therefore for all i and t in expression (1) $\alpha_{i,t} = 1$. In accordance with the system of assumptions, the signal power of all user devices is the same (and the energy too), and the number of users transmitting in the slot t, taking into account the target one, is equal to K_t, then

$$\gamma_{i,t} = \frac{E}{(K_t - 1)E + N_0}.$$

Divide the numerator and denominator by the signal energy, then

$$\gamma_{i,t} = \frac{1}{(K_t - 1) + \frac{N_0}{E}}.$$

Consider separately $\frac{N_0}{E}$:

$$\frac{N_0\ [\mathrm{J}]}{E\ [\mathrm{J}]} = \frac{P_{noise} \cdot T\ [\mathrm{W \cdot s}]}{P \cdot T\ [\mathrm{W \cdot s}]} = \frac{P_{noise}\ [\mathrm{W}]}{P\ [\mathrm{W}]}.$$

Taking into account the Assumption 2 $P_{noise} = 1$, therefore $\frac{N_0}{E} = \frac{1}{P} = P^{-1}$. By substituting this value into the previous expression, we obtain Statement 1, which was required to be proved.

Let us describe the operation of a basic system with an ALOHA-type random access algorithm for a channel with AWGN, as well as give clarifications and examples.

Remark 3. Taking into account Assumption 5, the direct coding theorem (see Remark 2), and independent attempts to transmit a message, the following events can be defined in the system under consideration:

— successful decoding of the message occurs if $W\log_2(1 + \gamma_{i,t}) > R$;

— message decoding error occurs if $W\log_2(1 + \gamma_{i,t}) \le R$.

Taking into account the introduced system model, Statement 1 and Remark 3, let us specify the events occurring in the channel (see Assumption 2) for the basic system:

"Success". Only one user i transmitted in the slot, and in this case $\gamma_{i,t} = \frac{1}{P^{-1}} = P$, and, consequently, $W\log_2(1 + \gamma_{i,t}) > R$, which corresponds to successful decoding of the message (see Remark 3).

"Conflict". More than one user $K_t > 1$ was transmitting in the slot. In this case, it is also easy to show that $P > \gamma_{i,t}$ for $K_t > 1$. Therefore, $W\log_2(1 + \gamma_{i,t}) \le R$, which corresponds to a message decoding error (see Remark 3).

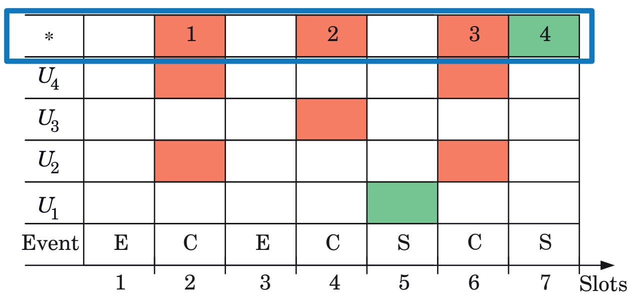
"Empty". Nobody transmitted messages and, therefore, the power at the input of the receiver are low and the base station understands that the channel is empty.

Let's consider an example of the operation of the base system from the side of the target user device (marked in Fig. 1 by the symbol *) in the presence of other user devices in the system (U_i means the i-th user device). The events in the figure are designated as follows: "S" — success, "C" — conflict and "E" — empty. Let the signal power value in this example be $P = 1$.

Let's describe the events in each slot.

— First slot. Nobody sent messages — "Empty" event.

— Second slot. Passed the target (first attempt), 2nd and 4th users:



■ ***Fig. 1.*** An example of the operation of the basic system on slots

• user * has the following SINR value $\gamma_{*,2} = \frac{1}{2+P^{-1}} < P$, where 1/3 < 1 means the message is not decoded;

• user 2: $\gamma_{2,2} = \frac{1}{2+P^{-1}} < P$, decoding error;

• user 4: $\gamma_{4,2} = \frac{1}{2+P^{-1}} < P$, decoding error.

The messages of all three users were not decoded, the base station reports the "Conflict" event.

— Third slot. Nobody sent messages — "Empty" event.

— Fourth slot. Passed target (second attempt) and 3rd users:

• user *: $\gamma_{*,4} = \frac{1}{1+P^{-1}} < P \rightarrow 1/2 < 1$, decoding error;

• user 2: $\gamma_{2,4} = \frac{1}{1+P^{-1}} < P$, decoding error.

A "Conflict" event has occurred in the slot.

— Fifth slot. Transmitted by only 1st user:

• user 1: $\gamma_{1,5} = \frac{1}{P^{-1}} = P$, successfully decoded the message.

The message is successfully decoded, a "Success" event occurs.

— Sixth slot. Transmitted target (third attempt), 2nd and 4th users:

• user *: $\gamma_{*,6} = \frac{1}{2+P^{-1}} < P$, while 1/3 < 1, decoding error;

• user 2: $\gamma_{2,6} = \frac{1}{2+P^{-1}} < P$, decoding error;

• user 4: $\gamma_{4,6} = \frac{1}{2+P^{-1}} < P$, decoding error.

Event "Conflict".

— Seventh slot. Passed only by the target user *:

• user *: $\gamma_{*,7} = \frac{1}{P^{-1}} = P$, successful decoding.

The message is successfully decoded, a "Success" event occurs.

Let us consider how to calculate the spectral efficiency T [bit/s/Hz] if the critical input arrival rate Λ_{cr} of messages/slot is known, up to which the system is stable for the systems under consideration.

Definition 1. The critical input arrival rate Λ_{cr} is related to the spectral efficiency T by the following expression:

$$T = \frac{\Lambda_{cr} R}{W}. \quad (2)$$

Remark 4. For systems with an ALOHA-type algorithm, the average number of transmissions $E[S]$ is related to the critical input arrival rate $\Lambda \le \Lambda_{cr}$ by the following expression $E[S] = 1/\Lambda$ [20].

Taking into account what was written earlier, we introduce the following Assumption.

Statement 2. The spectral efficiency for the base system is calculated as

$$T = \frac{\log_2(1+P) - \varepsilon}{e}, \quad (3)$$

where $\varepsilon \rightarrow 0$.

Proof: According to Definition 1 and Remark 4, the spectral efficiency can be redefined as follows:

$$T = \frac{R}{WE[S]}. \quad (4)$$

The average number of message transmissions is calculated as

$$E[S] = \sum_{s=1}^{\infty} s \Pr\{S = s\},$$

where $\Pr\{S = s\} = P_e(s-1) - P_e(s)$ and $P_e(s)$ is the probability that s transmissions in a row were unsuccessful. In accordance with the definition of the concept of successful message decoding from Remark 3, we can write the following:

$$P_e(s) = \prod_{i=1}^{s} \Pr\left\{R \ge W \log_2\left(1+\gamma_{*,i}\right)\right\},$$

where $\gamma_{*,i}$ is the SINR for the target user device *. Taking into account the fact that the transmissions are independent and the distribution $\gamma_{*,i}$ does not change with time, this expression can be rewritten as follows:

$$P_e(s) = \left(\Pr\left\{R \ge W \log_2\left(1+\gamma_{*,1}\right)\right\}\right)^s.$$

In accordance with the introduced model of the system $R = W \log_2(1 + P) - \varepsilon$, where $\varepsilon \rightarrow +0$ (see Assumption 5):

$$P_e(s) = \left(\Pr\left\{W \log_2\left(1+P\right) - \varepsilon \ge W \log_2\left(1+\gamma_{*,1}\right)\right\}\right)^s,$$

where $\varepsilon \rightarrow 0$.

Since the band when calculating the speed and throughput are the same, then in the condition $\Pr\{W\log_2(1 + P) - \varepsilon \ge W\log_2(1 + \gamma_{*,s})\}$, it can be omitted. Provided that the logarithm is a monotonically increasing function and $\varepsilon \rightarrow 0$, then the logarithm can be omitted in the condition and its arguments can be compared. After some simplifications, write:

$$P_e(s) = \left(\Pr\left\{P > \gamma_{*,1}\right\}\right)^s.$$

The calculation of SINR for users transmitting at time t, taking into account the system of assumptions, was defined earlier (see Statement 1). When adding a target user to the channel, SINR for it is defined as $\gamma_{*,s} = \frac{1}{K_s^* + P^{-1}}$, where K_s^* is the number of users who transmitted together with the target during the s-th transmission. It can be easily seen that $\frac{1}{K_s^* + P^{-1}} < P$, only for $K_s^* > 0$. Since the system is analyzed in saturation mode, then K_s^* is distributed according to the Poisson law with parameter 1. Then:

$$\Pr\left\{P > \hat{\gamma}_{*,s}\right\} = \Pr\left\{K_s^* > 0\right\} = 1 - e^{-1}.$$

Whence it follows that

$$P_e(s) = \left(1 - e^{-1}\right)^s.$$

Let's write $\Pr\{S = s\}$:

$$\Pr\left\{S = s\right\} = \left(1 - e^{-1}\right)^{s-1} - \left(1 - e^{-1}\right)^s = \left(1 - e^{-1}\right)^{s-1} e^{-1}.$$

It is easy to show that in this case:

$$E[S] = \sum_{s=1}^{\infty} s\left(1 - e^{-1}\right)^{s-1} e^{-1} = e.$$

Substituting the values of the code rate R in accordance with Assumption 5 and $E[S]$ calculated earlier into expression (4), we obtain Statement 2, which was required to be proved.

Remark 5. As is known, the critical input arrival rate Λ_{cr} for the basic algorithm of the ALOHA-type is equal to e^{-1}, substituting this value into expression (2) we get a result similar to Statement 2, which is explained by Remark 4.

Let's show how to calculate the lower bounds for signal power P and energy per bit $\frac{E_b}{N_0}$.

Statement 3. The lower bound for the signal power P in the base system for a given spectral efficiency T is defined as

$$P > 2^{eT} - 1. \tag{5}$$

Proof: This statement follows directly from expression (3) if we express P from it.

The value of energy per bit can $\frac{E_b}{N_0}$ be obtained by recalculating from the signal power. For this, we introduce the following definition.

Definition 2. The energy per bit is related to the signal power P as follows:

$$\frac{E_b}{N_0} \triangleq PE[S]\frac{W}{R},$$

where $E[S]$ is the average number of transmissions before the successful delivery of one message; W is the channel bandwidth,

For the base system, the lower bound for energy cost per bit $\frac{E_b}{N_0}$ is calculated according to the following statement.

Statement 4. The value of the lower bound on energy costs per bit $\frac{E_b}{N_0}$ for a given spectral efficiency T for the basic system can be found as

$$\frac{E_b}{N_0} = \frac{2^{eT} - 1}{T}. \tag{6}$$

Proof: It follows from Remark 4 that, for the system under consideration, the spectral efficiency can be rewritten as $T = \frac{R}{WE[S]}$.

With this in mind, it can be seen that the definition of energy per bit (see Definition 2) can be written as follows:

$$\frac{E_b}{N_0} = \frac{P}{T},$$

where

$$P = \frac{E_b}{N_0}T.$$

Substituting this value into expression (3) from Statement 2, we get

$$T = \frac{\log_2\left(1 + \frac{E_b}{N_0}T\right) - \varepsilon}{e}.$$

Let us express from the given expression $\frac{E_b}{N_0}$, we will obtain Statement 4, which was required to be proved.

Analysis of the system using HARQ with Chase combining

Let us describe the operation of a system with a random access algorithm of the ALOHA-type when using HARQ with Chase combining, as well as give clarifications and examples.

Taking into account Assumption 5, the direct coding theorem and the fact that the HARQ with Chase combining can be considered as signal accumulation at the receiver [16]. Then we can rewrite the calculation of SINR, taking into account Chase combining, as $\tilde{\gamma}_{i,n} = \sum_{j=1}^{n} \gamma_{i,j}$, where $\tilde{\gamma}_{i,n}$ is the SINR of the i-th user during the n-th attempt to send a message, taking into account the accumulation of energy; n is the number of attempts to send a message to the i-th user.

Remark 6. Then, taking into account the dependence that has arisen when trying to send a message, you can redefine the following events:

— successful decoding of the message. Occurs if

$$W\log_2(1+\tilde{\gamma}_{i,n}) = W\log_2\left(1+\sum_{j=1}^{n}\gamma_{i,j}\right) > R;$$

— message decoding error. Occurs if

$$W\log(1+\tilde{\gamma}_{i,n}) = W\log\left(1+\sum_{j=1}^{n}\gamma_{i,j}\right) \le R.$$

Taking into account the model introduced earlier, and the accumulation of energy due to HARQ, we will clarify the events occurring in the channel (see Assumption 2) for a system with HARQ:

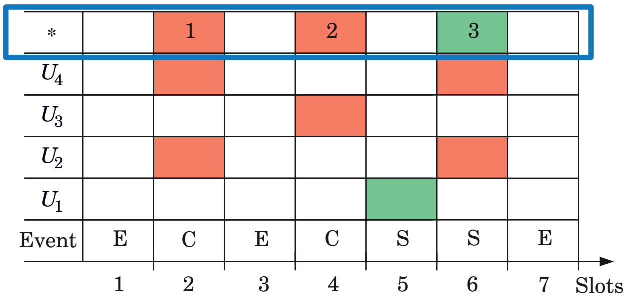
"Success". Only one user i transmitted in the slot, and in this case $\tilde{\gamma}_{i,n} = \sum_{j=1}^{n-1}\gamma_{i,j} + \frac{1}{P^{-1}} = \sum_{j=1}^{n-1}\gamma_{i,j} + P$, and therefore $W\log(1+\tilde{\gamma}_{i,n}) > R$, which corresponds to successful decoding of the message (see Remark 6). And also when transmitting by several users, if for one of the users (*) in the current collision it was fulfilled $W\log\left(1+\sum_{j=1}^{n}\tilde{\gamma}_{*,j}\right) > R$, which also leads to successful decoding.

"Conflict". More than one user $K_t > 1$ was transmitting in the slot, and for all users in the collision $W\log\left(1+\sum_{j=1}^{n}\tilde{\gamma}_{*,j}\right) \le R$, which corresponds to a message decoding error for all users in the mixture (see Remark 6).

"Empty". Nobody transmitted messages, and, therefore, the power at the input of the receiver is low and the base station understands that the channel is empty.

Let's consider an example of the operation of a system with HARQ from the side of the target user device (marked in Fig. 2 by the symbol *) in an example similar to the previous one. As before, other user devices are designated as U_i — the i-th user device. Taking into account the fact that attempts to decode one message during retransmissions are now dependent, we will consider each transmission in more detail. Let the value of the signal strength, as before, $P = 1$.

— First slot. Nobody sent messages is "Empty" event.



■ ***Fig. 2.*** An example of the operation of the system when using HARQ with Chase combining

— Second slot. Passed the target (first attempt), 2nd and 4th users:

- user *: $\tilde{\gamma}_{*,1} = \gamma_{*,2} = \frac{1}{2+P^{-1}} < P$, $1/3 < 1$, hence decoding error;
- user 2: $\tilde{\gamma}_{2,1} = \gamma_{2,2} = \frac{1}{2+P^{-1}} < P$, decoding error;
- user 4: $\tilde{\gamma}_{4,1} = \gamma_{4,2} = \frac{1}{2+P^{-1}} < P$, decoding error.

Event "Conflict".

— Third slot. Nobody sent messages is "Empty" event.

— Fourth slot. Passes target (second try) and 3rd users:

- user *: $\tilde{\gamma}_{*,2} = \gamma_{*,2} + \gamma_{*,4} = \frac{1}{2+P^{-1}} + \frac{1}{1+P^{-1}} = \frac{3+P^{-1}}{2+3P^{-1}+P^{-2}} < P$, $4/6 < 1$, hence decoding error;
- user 3: $\tilde{\gamma}_{3,1} = \gamma_{3,4} = \frac{1}{1+P^{-1}} < P$, $1/2 < 1$, decoding error.

Both messages are not decoded, "Conflict" event.

— Fifth slot. Transmitted by only 1st user:

- user 1: $\tilde{\gamma}_{1,1} = \gamma_{1,5} = \frac{1}{P^{-1}} = P$, successful decoding.

User 1's message is successfully decoded, a "Success" event occurs.

— Sixth slot. The target (third attempt), 2nd and 4th users were sent again:

- user *: $\tilde{\gamma}_{*,3} = \gamma_{*,2} + \gamma_{*,4} + \gamma_{*,6} = \frac{1}{2+P^{-1}} + \frac{1}{1+P^{-1}} + \frac{1}{2+P^{-1}} = \frac{4+3P^{-1}}{2+3P^{-1}+P^{-2}} > P$, $7/6 > 1$, hence the successful decoding of the message;
- user 2: $\tilde{\gamma}_{2,2} = \gamma_{2,2} + \gamma_{2,6} = \frac{1}{2+P^{-1}} + \frac{1}{2+P^{-1}} = \frac{2}{2+P^{-1}} < P$, $2/3 < 1$, decoding error;
- user 4: $\tilde{\gamma}_{4,2} = \gamma_{4,2} + \gamma_{4,6} = \frac{1}{2+P^{-1}} + \frac{1}{2+P^{-1}} = \frac{2}{2+P^{-1}} < P$, decoding error.

The target user has accumulated enough power to successfully decode, that is, $\tilde{\gamma}_{*,3} > P$, hence $W\log(1+\tilde{\gamma}_{i,n}) > R$, which means that on the third attempt, a successful decoding occurs for the message of the target user. "Success" event for user device *.

— Seventh slot. Nobody sent messages is "Empty" event.

During the demonstration of the example, it is shown that the use of HARQ can complicate the decoding procedure, but at the same time, in a specific example, it has reduced the number of transmissions for the target user, which also reduces power consumption and delay in the system.

Let us introduce a statement for calculating the spectral efficiency of a system with HARQ.

Statement 5. The spectral efficiency T of the system when using HARQ with Chase combining is defined as

$$T = \frac{\log\left(1+P\right)-\varepsilon}{S(P)}. \tag{7}$$

The function $S(P)$, which is included in expression (7), determines the average number of transmissions of the message before success and is calculated as follows:

$$S(P) = \sum_{s=1}^{\infty} s\left(P_e(s-1)-P_e(s)\right),$$

where $P_e(s) = \Pr\left\{P > \sum_{i=1}^{s}\frac{1}{K_i^* + P^{-1}}\right\}$, and $K_1^*, K_2^*, \ldots$ is a sequence of independent random variables identically distributed according to the Poisson law with parameter 1.

Proof: The proof is similar to the previous one, with some differences. For this type of HARQ, $P_e(s)$ is defined as [16]

$$P_e(s) = \Pr\left\{R \geq W\log\left(1+\sum_{i=1}^{s}\hat{\gamma}_{*,i}\right)\right\}.$$

This condition can be simplified similarly to the condition in the analysis of the basic system:

$$P_e(s) = \Pr\left\{P > \sum_{i=1}^{s}\hat{\gamma}_{*,i}\right\}.$$

Taking into account the addition of the target user device to the system $\hat{\gamma}_{*,i} = \frac{1}{K_i^* + P^{-1}}$, where K_i^* is the number of users who transmitted with the target user device when the i-th message was sent. Since the system is analyzed in the saturation mode, then $K_1^*, K_2^*, \ldots$ is a sequence of independent random variables identically distributed according to the Poisson law with parameter 1. Whence Assumption 5 follows and follows.

The paper does not consider the issue of numerical calculation of the function $S(P)$, since this is a difficult task from the field of probability theory. Given that the number of transfers is a random variable that has a finite mathematical expectation and variance for any $P > 0$ (within the framework of the system model under consideration), then in the work, when obtaining numerical results, the Monte Carlo method will be used to calculate this function.

If possible, a way to calculate lower bounds for signal power P and energy per bit $\frac{E_b}{N_0}$ in a HARQ system.

It follows directly from Statement 5.

Statement 6. The lower bound for the signal power P in the system with HARQ Chase combining for a given spectral efficiency T is defined as a solution to the equation

$$P = 2^{S(P)T} - 1. \tag{8}$$

Proof: This statement directly follows from expression (7) if we express P from it.

The value of energy per bit $\frac{E_b}{N_0}$, as before, can be obtained by recalculation from the signal strength.

Statement 7. The value of the lower bound costs of energy per bit $\frac{E_b}{N_0}$ for a given spectral efficiency T for a system with the first type of HARQ can be found as a solution to the equation

$$\frac{E_b}{N_0} = \frac{2^{S\left(\frac{E_b}{N_0}T\right)T} - 1}{T}. \tag{9}$$

Proof: The proof is similar to Statement 4, so we will show briefly. It follows from Remark 4 and Statement 5 that for the system under consideration the spectral efficiency can be rewritten as

$$T = \frac{R}{WS\left(P\right)},$$

where

$$P = \frac{E_b}{N_0}T.$$

Substituting this value into expression (3) from Statement 2, we get

$$T = \frac{\log_2\left(1 + \dfrac{E_b}{N_0}T\right) - \varepsilon}{S\left(\dfrac{E_b}{N_0}T\right)}.$$

After a number of simplifications, we obtain Statement 7, which was to be proved.

Analysis of a system with a blocked random access algorithm with a dynamic schedule

Consider changing the random multiple access algorithm in a noisy channel. As an algorithm, a blocked conflict resolution algorithm with a dynamic schedule will be used [17, 18]. The algorithm works in sessions. The session starts with a "Conflict" of some multiplicity (multiplicity 0 — "Empty" event, multiplicity — "Success" event, multiplicity > 1 — "Conflict" event) and lasts in accordance with the conflict resolution algorithm. Let's change Assumption 3.

Assumption 3*. There are an infinite number of unique preambles, arranged in such a way that when superimposed on the channel, the base station can determine the list of preambles in the mixture, but cannot recover the user data. When a message occurs at the user device, it transmits it in the next slot, if the system is not in scheduled mode, otherwise it postpones transmission until this mode is completed. On a collision event, when > 1 preamble is detected, the base station generates a schedule for the system to switch to the time division mode, where the preamble list determines the slot order for the user device to retransmit the message on the given schedule.

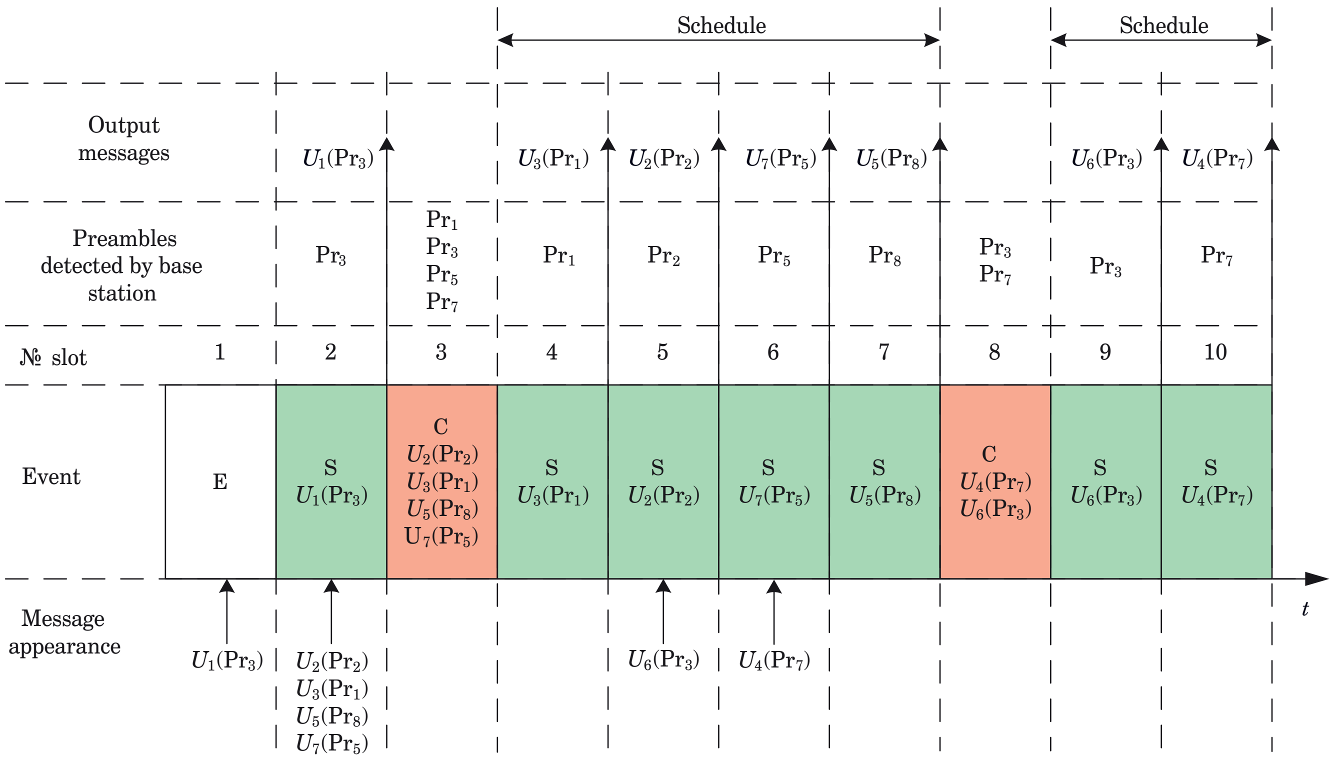
An example of the operation of the algorithm is shown in Fig. 3. Let's describe the events in the system by slots:

— First slot. At the beginning of operation, the system is empty, that is, there are no user devices with messages ready to be transmitted. The slot has an "Empty" event. During the slot a message appears to the first user and he selects the third preamble $U_1(\mathrm{Pr}_3)$.

— Second slot. At the start of the second slot, the first user has a message ready and successfully sends it. A "Success" event has occurred in the slot. During the slot, messages appear for 2, 3, 5 and 7 users who have selected 2, 1, 8 and 5 preambles respectively — $U_2(\mathrm{Pr}_2)$, $U_3(\mathrm{Pr}_1)$, $U_5(\mathrm{Pr}_8)$, $U_7(\mathrm{Pr}_5)$.

— Third slot. 2, 3, 5 and 7 users decided to send the message, a "Conflict" occurred. The base station detected 2, 1, 8 and 5 preambles. According to the list of preambles, a schedule is built for the next fourth slot.

— Fourth slot. In accordance with the schedule transmits $U_3(\mathrm{Pr}_1)$, the event "Success".



■ ***Fig. 3.*** An example of the operation of a blocked conflict resolution algorithm with a dynamic schedule

— Fifth slot. In accordance with the schedule transmits $U_2(\mathrm{Pr}_2)$, the event “Success”. During the slot, a message appeared for the 6th user who selected the 3rd preamble — $U_6(\mathrm{Pr}_3)$. Since the system is in scheduled mode, it postpones sending the message until the mode ends.

— Sixth slot. Transmits $U_7(\mathrm{Pr}_5)$, “Success” event as scheduled. During the slot, a message appeared for the 4th user who chose the 7th preamble — $U_4(\mathrm{Pr}_7)$. Since the system is in scheduled mode, it postpones sending the message until the mode ends.

— Seventh slot. Transmits $U_5(\mathrm{Pr}_8)$, “Success” event as scheduled.

— Eighth slot. By the beginning of the eighth slot, the system ends in schedule mode, users 6 and 4, who postponed the transmission of messages, can now transmit. A “Conflict” situation appears in the slot. The base station detected preambles 3 and 7, respectively, and generated a new schedule.

— Ninth and tenth slots. Users $U_6(\mathrm{Pr}_3)$ and $U_4(\mathrm{Pr}_7)$ successfully transmit messages in accordance with the schedule.

For the considered algorithm, the papers [17, 18] provide an analysis of the critical input arrival rate, and it is shown that $\Lambda_{cr} = 1$. Based on this and Definition 1, we introduce the following statement.

Statement 8. The lower bound for the signal power P for a given spectral efficiency T for a blocking algorithm with a dynamic schedule for an infinite number of preambles is defined as

$$P = 2^T - 1. \tag{10}$$

Proof: Assumption 5 implies that $R = W\log(1 + P) - \varepsilon$, where $\varepsilon \to +0$, taking into account Definition 1, we can write

$$T < \Lambda_{cr}\log(1+P),$$

where should

$$P > 2^{\frac{T}{\Lambda_{cr}}} - 1.$$

As noted earlier, it follows from [17, 18] that for the considered algorithm $\Lambda_{cr} = 1$, then

$$P > 2^T - 1.$$

Which is what needed to be proven.

Now let’s write a statement for calculating the lower bound costs for energy per bit $\dfrac{E_b}{N_0}$ in the system under consideration.

Statement 9. The lower bound on the energy per bit $\dfrac{E_b}{N_0}$ for a given spectral efficiency T for a blocking algorithm with an infinite number of preambles is defined as

$$\frac{E_b}{N_0} > \frac{2\left(2^T - 1\right)}{T}. \tag{11}$$

Proof: From expression (10) it follows that $P = 2^T - 1$. In accordance with the definition of one, we can write $R = \dfrac{WT}{\Lambda_{cr}}$, since $\Lambda_{cr} = 1$, then $R = WT$. Taking into account Definition 2, we obtain

$$\frac{E_b}{N_0} > E[S]\frac{2^T - 1}{T}.$$

Let us describe the calculation of the average number of retransmissions $E[S]$ for a system with a blocked conflict resolution algorithm. Consider the average number of transmissions for a target user device added to the system at a random time (this approach does not violate the system performance). If no one appeared during the previous schedule, then the user will make only 1 transfer. Otherwise, the user will get into conflict on the first transfer and will transfer a second time according to the schedule. In accordance with the system of assumptions and the features of the operation of the blocked algorithm, the user’s message will be successfully delivered when scheduled. When considering the operation of the system in saturation mode, the probability that no one appeared in the system during the time schedule tends to zero. Therefore, the subscriber will send the message twice, i.e. $E[S] = 2$. Means

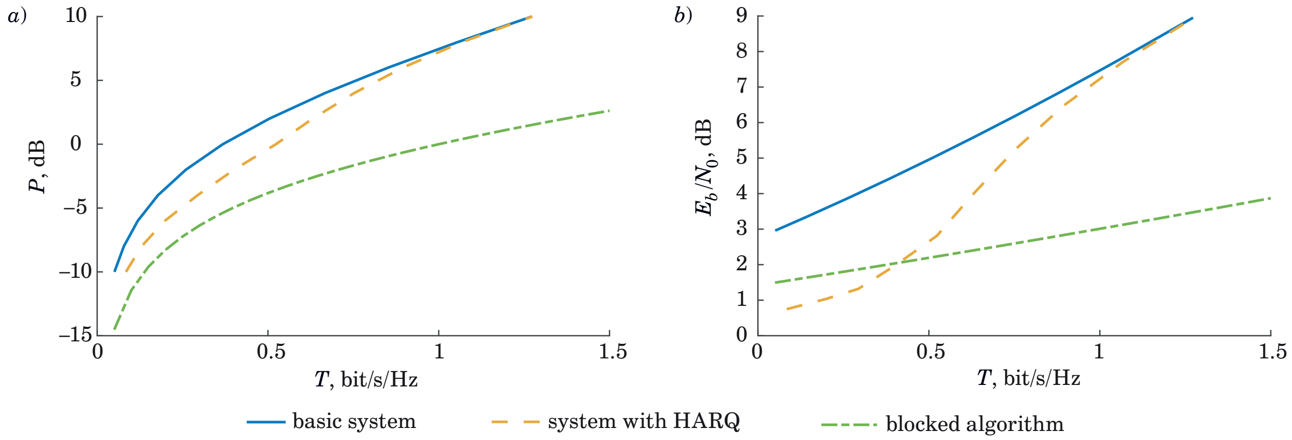
$$\frac{E_b}{N_0} > \frac{2(2^T - 1)}{T}.$$

Which is what needed to be proven.

Numerical results

Using the formulas obtained earlier for the lower bounds of energy consumption, we will compare the considered systems.

Figure 4, *a* shows the lower bounds for the signal power obtained by formula (5) for the basic system, by formula (8) for a system with HARQ (using the Monte Carlo method for the average number of transmissions) and by formula (10) for a system with a blocked conflict resolution algorithm due to a dynamic schedule. It follows from the figure that the use of HARQ makes it possible to reduce the lower bound for the signal power P at spectral efficiency values less than 1 bit/s/Hz;



■ ***Fig. 4.*** The lower bound of the signal power (*a*) and on the energy per bit (*b*) for the systems under consideration on the required spectral efficiency

at higher values, the signal power begins to coincide with the base system. The maximum energy gain of HARQ relative to the base system is about 2 dB in the region of spectral efficiency $T = 0.5$ bit/s/Hz. However, the blocked algorithm makes it possible to further reduce the lower bound for the signal power, gaining from 4.5 dB relative to the base system, and from 3.4 dB relative to the system with HARQ. This gain grows with the increase in the values of the spectral efficiency of the system.

Figure 4, *b* shows the lower bounds for energy costs per bit obtained by formula (6) for the base system, by formula (9) for a system with HARQ (using the Monte Carlo method for the average number of transmissions) and by formula (11) for systems with a blocked conflict resolution algorithm due to a dynamic schedule. It follows from the figure that the use of HARQ makes it possible to reduce the cost per bit compared to the basic system, however, as the required spectral efficiency increases, the gain decreases. The maximum energy gain is about 2.5 dB at the value of the spectral efficiency $T = 0.28$ bit/s/Hz. The minimum gain of a system with a blocked algorithm in relation to the basic one is 1.5 dB and grows with an increase in the spectral efficiency. However, when comparing the blocked algorithm with HARQ, it can be seen that at spectral efficiency values below a certain value ($T = 0.42$ bit/s/Hz), the blocked algorithm has higher costs energy per bit. This is due to the specifics of the algorithm's collision resolution method, which increases the average number of message transmissions in the system. But after the value of the spectral efficiency $T = 0.42$ bit/s/Hz, the energy per bit for the blocked algorithm turns out to be lower than systems with HARQ, and the energy gain increases.

Conclusion

The paper analyzed and compared the effectiveness of various approaches to reduce energy costs in massive machine-type communication systems with a potentially unlimited number of users. For this purpose, a basic system model was formulated and described based on the ALOHA-type algorithm for the massive machine-type scenario in a channel with additive white Gaussian noise. Two ways to reduce energy costs for were also described.

The first way to use it is to apply hybrid automatic repeat request methods based on Chase combining. For this approach, within the framework of the introduced model, an example of the operation of the system was shown with a description of the main features.

The second way is to change the random multiple access algorithm to a conflict resolution algorithm with a dynamic schedule. For this algorithm in the base model, assumption 3 from the base model was changed and an example of the system operation is given.

For all considered systems, an analysis of the lower bounds for signal power and costs energy per bit during stable working and specified requirements for spectral efficiency is given. It is shown that both methods make it possible to reduce energy costs compared to the basic system. When calculating the lower bounds in the system with hybrid automatic repeat request, the Monte Carlo method was used to determine the average number of message transmissions before successful decoding.

When analyzing the dependencies of the lower bound for the signal power on the spectral efficiency of the system, the following results were obtained. The use of HARQ allows the signal power at spectral efficiency values (less than 1 bit/s/Hz), at higher values the signal power begins to match the base system. The maximum energy gain from using hybrid automatic repeat request relative to the base system is about 2 dB with a spectral efficiency of about 0.5 bit/s/Hz. The blocked algorithm reduces the lower bound more efficiently than hybrid automatic repeat request. Gain relative to the base system from 4.5 dB, and relative to the system with HARQ from 3.4 dB. This gain grows with the increase in the values of the spectral efficiency of the system.

When analyzing the dependence of costs energy per bit on the spectral efficiency of the system, the following results were obtained. At values of spectral efficiency less than 0.42 bit/s/Hz, the use of hybrid automatic repeat request most effectively reduces energy consumption (energy gain of the order of 2.5 dB at a spectral efficiency of 0.28 bit/s/Hz). However, for large values of the spectral efficiency, the energy per bit for the blocked algorithm is lower than for a system with HARQ.

The results obtained allow us to evaluate the potential for reducing energy costs in stable systems with a potentially unlimited number of user devices through the use of hybrid automatic repeat request and a algorithm conflict resolution with dynamic schedule. However, when analyzing the base system and the hybrid automatic repeat request, an unrealistic assumption was assumed — the system and users know exactly the number of active users. This assumption can be taken into account and corrected by using methods for estimating the number of active users (see, for example, [21, 22]).

Financial support

The paper was prepared with the financial support of the Russian Science Foundation, project No. 22-19-00305 “Spatial-temporal stochastic models of wireless networks with a large number of users”.

References

1. Letaief K. B., Chen W., Shi Y., Zhang J., & Zhang Y. J. A. The roadmap to 6G: AI empowered wireless networks. *IEEE Communications Magazine*, IEEE, 2019, vol. 57(8), pp. 84–90. doi:10.1109/MCOM.2019.1900271
2. Piran M. J., & Suh D. Y. Learning-driven wireless communications, towards 6G. *2019 Intern. Conf. on Computing, Electronics & Communications Engineering (iCCECE)*, IEEE, 2019, pp. 219–224. doi:10.1109/iCCECE46942.2019.8941882
3. Zaidi A., Hussain Y., Hogan M., & Kuhlins C. Cellular IoT evolution for industry digitalization. *Ericsson White Paper*, GFMC, 2019, vol. 19, p. 17.
4. Vaezi M., Azari A., Khosravirad S. R., Shirvanimoghaddam M., Azari M. M., Chasaki D., & Popovski P. Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G. *IEEE Communications Surveys & Tutorials*, IEEE, 2022, vol. 24(2), pp. 1117–1174. doi:10.1109/COMST.2022.3151028
5. Borisovskaya A. V., & Turlikov A. M. Reducing energy consumption in the IoT systems with unlimited number of devices. *2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2021, pp. 1–6. doi:10.1109/WECONF51603.2021.9470596
6. Stusek M., Stepanov N., Moltchanov D., Masek P., Mozny R., Turlikov A., & Hosek J. Optimizing NB-IoT communication patterns for permanently connected mMTC device. *2022 IEEE Wireless Communications and Networking Conf. (WCNC)*, IEEE, 2022, pp. 1413–1418. doi:10.1109/WCNC51071.2022.9771847
7. Stepanov N. V., Tyurlikov A. M., & Begishev V. O. Optimal device associations in machine-to-machine networks of NB-IoT and LTE-M standards. *Systems and Means of Informatics*, 2022, vol. 32, no. 3, pp. 63–70 (In Russian). doi:10.14357/08696527220306
8. Borisovskaya, A. V., Turlikov A. M. Estimation of the average age of information in random access systems with multiple departure. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 1, pp. 51–60 (In Russian). doi:10.31799/1684-8853-2023-1-51-60
9. Galinina O. S., Andreev S. D., Tyurlikov A. M. On multiple access of a large number of machine-type devices in cellular networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 4, pp. 105–114 (In Russian) doi:10.31799/1684-8853-2018-4-105-114
10. Chen S., Liang Y. C., Sun S., Kang S., Cheng W., & Peng M. Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed. *IEEE Wireless Communications*, IEEE, 2020, vol. 27(2), pp. 218–228. doi:10.1109/MWC.001.1900333
11. Clazzer F., Munari A., Liva G., Lazaro F., Stefanovic C., & Popovski P. From 5G to 6G: Has the time for modern random access come? *arXiv preprint arXiv:1903.03063*, 2019.
12. Wu Y., Gao X., Zhou S., Yang W., Polyanskiy Y., & Caire G. Massive access for future wireless communication systems. *IEEE Wireless Communications*, IEEE, 2020, vol. 27(4), pp. 148–156. doi:10.1109/MWC.001.1900494
13. Bogatyrev V. A., Bogatyrev S. V., & Bogatyrev A. V. Model and interaction efficiency of computer nodes based on transfer reservation at multipath routing, *2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2019, pp. 1–4. doi:10.1109/WECONF.2019.8840647

14. Bogatyrev V. A., Bogatyrev A. V., & Bogatyrev S. V. Redundant multi-path service of a flow heterogeneous in delay criticality with defined node passage paths. *Journal of Physics: Conference Series*, IOP Publishing, 2021, vol. 1864, no. 1, p. 012094. doi:10.1088/1742-6596/1864/1/012094
15. Bachevsky S. V., Dvornikov S. V., Dvornikov S. S., Tsarelungo A. B. Radio speed proposals. *Questions of Radio Electronics. Series: Television Technology*, Radio electronics issues, 2021, no. 4, pp. 40–45 (In Russian).
16. Burkov A. A., Turlikov A. M. An upper estimate of the spectral efficiency for systems with hybrid automatic repeat request with a restriction on the type of modulation. *Questions of Radio Electronics. Series: Television Technology*, Radio electronics issues, 2020, vol. 1, pp. 74–83 (In Russian).
17. Burkov A., Frolov A., & Turlikov A. Contention-based protocol with time division collision resolution. *2018 10th Intern. Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, IEEE, 2018, pp. 1–4. doi:10.1109/ICUMT.2018.8631217
18. Burkov A. A., Saveliev A. A. Comparison of blocked and nonblocked algorithm random multiple access with a dynamic schedule. *Scientific Session of the SUAI Collection of Reports*, Saint-Petersburg, 2019, vol. 1, pp. 306–313 (In Russian).
19. Cover T. M., & Thomas J. A. *Elements of Information Theory*. John Wiley & Sons, 2012.
20. Burkov A., Shneer S., & Turlikov A. An achievability bound of energy per bit for stabilized massive random access Gaussian channel. *IEEE Communications Letters*, IEEE, 2020, vol. 25(1), pp. 299–302. doi:10.1109/LCOMM.2020.3023461
21. Yu J., Zhang P., Chen L., Liu J., Zhang R., Wang K., & An J. Stabilizing frame slotted ALOHA-based IoT systems: A geometric ergodicity perspective. *IEEE Journal on Selected Areas in Communications*, IEEE, 2020, vol. 39(3), pp. 714–725. doi:10.1109/JSAC.2020.3018795
22. Liu J., Seo J. B., Jin H. Online transmission control for random access with multipacket reception and reservation. *IEEE Internet of Things Journal*, IEEE, 2022, vol. 9, iss. 22. doi:10.1109/JIOT.2022.3188280

УДК 004.728.3.057.4
doi:10.31799/1684-8853-2023-2-39-50
EDN: EHYAOZ

Сравнение способов снижения затрат энергии в стабильных системах массовой межмашинной связи

А. А. Бурков[a], ассистент, orcid.org/0000-0002-0920-585X, a.burkov@k36.org
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: при развитии современных стандартов связи активно рассматриваются различные сценарии интернета вещей, в том числе массовой межмашинной связи. Основными требованиями к таким системам являются стабильная работа при потенциально бесконечном числе устройств и низкие затраты энергии, поэтому исследование способов достижения данных требований актуально. **Цель:** исследовать и сравнить эффективность различных подходов к снижению затрат энергии в системах с потенциально неограниченным числом пользователей. **Результаты:** описана модель базовой системы на основе алгоритма типа АЛОХА для сценария массовой межмашинной связи в канале с аддитивным белым гауссовым шумом. Описаны два способа снижения энергозатрат для рассматриваемой системы: использование методов гибридной решающей обратной связи и алгоритма разрешения конфликтов с динамическим расписанием. Для базовой системы и предложенных способов проведен анализ нижних границ для мощности сигнала и затрат энергии на бит при стабильной работе системы и заданных требованиях к спектральной эффективности. Показано, что оба способа позволяют уменьшить затраты энергии по сравнению с базовой системой. При этом алгоритм с динамическим расписанием дает наибольший выигрыш по сравнению с базовой системой для мощности сигнала от 4,5 дБ при любой спектральной эффективности. Для затрат энергии на бит при значениях спектральной эффективности меньше 0,42 бит/с/Гц наиболее эффективно применение гибридной обратной связи. Однако при больших значениях наибольший выигрыш дает изменение алгоритма разрешения конфликтов. **Практическая значимость:** полученные результаты позволяют оценить потенциальные возможности уменьшения затрат энергии в стабильных системах с большим числом пользовательских устройств за счет применения гибридной обратной связи и алгоритма разрешения конфликтов с динамическим расписанием.

Ключевые слова — массовая межмашинная связь, интернет вещей, гибридная решающая обратная связь, алгоритмы случайного множественного доступа, спектральная эффективность, энергоэффективность.

Для цитирования: Burkov A. A. Comparison of the ways to reduce energy costs in stable massive machine-type communication systems. *Информационно-управляющие системы*, 2023, № 2, с. 39–50. doi:10.31799/1684-8853-2023-2-39-50, EDN: EHYAOZ
For citation: Burkov A. A. Comparison of the ways to reduce energy costs in stable massive machine-type communication systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 39–50. doi:10.31799/1684-8853-2023-2-39-50, EDN: EHYAOZ