IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 April 2024, accepted 13 May 2024, date of publication 20 May 2024, date of current version 29 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3403257

RESEARCH ARTICLE

LoRaWAN Optimization for Voltage Monitoring

DMITRIY KIM [1], ANDREY TURLIKOV [2], GEORGI GEORGIEV [3], AND TATYANA DEDOVA [1,4]
[1]School of Digital Technologies, Narxoz University, 050035 Almaty, Kazakhstan
[2]Higher School of Economics (HSE), HSE Campus in St. Petersburg, 194100 Saint Petersburg, Russia
[3]Center for Energy Solutions, 9003 Varna, Bulgaria (*Deceased*)
[4]EcoRisk LLP, 050026 Almaty, Kazakhstan

Corresponding author: Dmitriy Kim (dmitriy.kim@narxoz.kz)

The work of Dmitriy Kim and Tatyana Dedova was supported by the Science Committee of the Ministry of Science and Higher Education of Republic of Kazakhstan under Grant AP19680230. The work of Andrey Turlikov was supported by the Basic Research Program at HSE University.

ABSTRACT This study examines the challenges encountered when using wireless technologies based on random multiple access for voltage monitoring in low voltage (LV) electrical grids. The introduction of photovoltaic modules on the roofs of buildings creates the need to monitor the dynamics of the node voltages of the grid. We consider two related objects – an LV electrical grid and a communication network with devices in the nodes of the grid. The communication network carries out one-way transmission of monitoring data from the devices to the LV grid operator and is based on LoRaWAN technology. Transmission is carried out via a common data transmission medium. We presume that each message contains information about the node voltage. The messages from each node in the grid are sent to the grid operator at random time intervals and independently of each other. If the airtimes of messages from two or more nodes overlap, a collision occurs and none of the messages reach the operator. The operator has to monitor the random voltage behavior over time to avoid exceeding a certain level (an upper voltage limit). Under the specifics of the voltage random process, we study the problem of choosing parameters of message transmission to uniformly minimize the probability of such excess among all nodes. The work precisely formulates the optimization problem and proposes an algorithm for its solution for the particular case.

INDEX TERMS ALOHA, distributed energy sources, LoRaWAN, optimization problem, photovoltaic, probabilistic approach, random multiple access, smart grid, spreading factor allocation, voltage monitoring.

I. INTRODUCTION

Nowadays, Low Power Wide Area Networks (LPWAN) are becoming more and more widespread [1], [2]. They are used to monitor the results of measurements at objects located over a large area. Long Range Wide-Area Networks (LoRaWAN) technologies, which are part of LPWAN, enable low-cost data transmission and collection systems.

In this work we describe how the specifics of the transmitted data might be used when the object of monitoring is the low voltage (LV) power grid. At present, distributed energy resources (DER) such as photovoltaic modules (PV) on the roofs of buildings are becoming increasingly common in many countries. However, random fluctuations in distributed power generation, caused by stochastic solar generation and consumption at grid nodes, have a significant impact on the voltage profile as compared to traditional grids [3]. It becomes necessary to monitor the voltages of the grid nodes.

Thus we consider two related objects – the communication network and the LV grid. Devices of the communication network, located in the nodes of the LV grid, send their data to the base station, with the possible loss of messages. The grid operator receives data from the base station on a reliable channel without loss, so we assume the base station and the grid operator act as one unit.

We consider low-cost monitoring systems (LoRaWAN devices of Class A) [1], where the device itself initiates transmission. Approaches in which devices of other classes are used (that is, Class C, see [4]), we do not consider. Therefore, a monitoring system in which a message is transmitted when the upper voltage limit is exceeded does

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Abdur Razzaque.

not guarantee successful delivery because of the possibility of other nodes exceeding the limit at the same time because the voltages in the grid are interconnected. Sending messages at random intervals is used in similar cases in practice, which means using the ALOHA algorithm [1], [5], [6], [7], [8] in such system. So, each node in the grid sends messages to the operator with information about the current node voltage at random intervals. The operator monitors the voltage so that it does not exceed the upper voltage limit, which can happen during the interval between messages.

Most of the works investigating LoRaWAN optimization do not take into account the specifics of the transmitted data (see [5] and references in [1]), with the exception of a few.

The general aspects of building monitoring systems based on LoRa technology were discussed in [9]. The authors conducted experiments and noted that the stochastic property was evident and cannot be neglected in wireless environments. The features of LoRa technology were studied in [10] and it was stated that class A devices have better energy efficiency than other classes of devices. This paper describes the results of a real experiment using class A devices to monitor an electrical grid with solar energy harvesting. The article discusses the type of data being transmitted (electrical grid data), but does not discuss the stochastic nature of the data itself. The impact of the specificity of the transmitted data using LoRa technology on the architecture of the communication system was studied in [11] and [12]. These works considered the application of LoRaWAN technology for smart cities. In [11], the problem of minimizing the maximum probability of packet loss due to conflict in communication systems based on LoRa technology was formulated and solved. However, this study did not consider the specificity of transmitted data. The concepts of this study were further developed in [12]. The architecture of communication systems based on LoRa technology for monitoring high voltage electrical grids was described in [13]. A data transmission algorithm is proposed, and its energy efficiency is studied, but the stochastic nature of the transmitted data is not considered.

A survey of machine-learning methods for building networks based on LoRa technology was presented in [14]. The use of machine learning for LoRaWAN link budget analysis was studied in [15]. The works [16], [17] can provide valuable insights into the application of machine learning methodologies to improve data analysis within LoRaWAN. They can potentially improve the predictive capabilities of the system and offer the prospect of working with large-scale data, which is important for managing the extensive data generated in monitoring low-voltage grids.

In all the above-described studies of systems based on LoRa technology, the impact of the specifics of the transmitted data on the architecture of the monitoring system is only partially considered, and the influence of the quantitative characteristics of the processes monitored on the parameters of the monitoring system is practically ignored.

As a rule the problem of maximizing the probability of message delivery is formulated and solved by correctly selecting the equal average length of the intervals between transmissions without regard to the specifics of the transmitted data. In our work we formulate and solve a similar problem for the LV grid monitoring system taking into account the specifics of the transmitted data by selecting the various average lengths of the intervals between transmissions. Note that we consider only the problem of monitoring, not the management of parameters, because other technologies are used for this purpose.

Our work is organized as follows: Section II begins with a message transmission and voltage monitoring models, which assume various message sending parameters for different nodes. We choose these parameters depending on the probability that the voltage exceeds the upper voltage limit. Section III formulates the optimization problem. The solution to the optimization problem is found in Section IV for voltage described by a Brownian motion. Section V is devoted to the problem of spreading factor allocation to reduce the objective function for the case with a Brownian motion. As an illustration, a numerical example is presented in Section VI. The results of simulations are represented in Section VII.

II. SYSTEM MODEL

A. MESSAGE TRANSMISSION MODEL

Consider a communication network with $N > 0$ nodes, each of which sends messages to the operator with information about the current value of some parameter, in our case, node voltage. We will use a number of assumptions to build the model.

Assumption 1: The intervals between messages of each node have an exponential distribution with parameter $\lambda_i > 0$, $i = 1, 2, \ldots, N$, and are independent for each node.

Then the stream of messages from the ith node forms a Poisson process with intensity λ_i, $i = 1, 2, \ldots, N$. It means that the total stream of messages from all nodes is Poisson with intensity

$$\Lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_N.$$

Assumption 2: All nodes (or devices in the nodes) use the same spreading factor (SF) value with the airtime of one message $Q > 0$ and the same frequency [1], [18]. If during the time interval $(-Q, Q)$ from the moment one node sends a message, the other nodes do not send messages, then the message is successfully delivered to the operator. Otherwise a collision occurs and the messages are lost. This assumption is fulfilled in practice when all nodes are approximately the same distance from the operator. According to this distance, the spreading factor is chosen so that in the absence of collisions the messages are delivered successfully.

According to the standard [18] (see also [2]) airtime Q must not exceed 1 % of the time between sent messages of each node. To account for this constraint, we introduce the following assumption on the average time between messages:

Assumption 3: Let

$$\frac{1}{100}\frac{1}{\lambda_i} \geq Q, \; i = 1, 2, \ldots, N. \tag{1}$$

Based on these assumptions and by performing a series of arguments [19] we consider a simplified model for total stream of successfully delivered messages and use the properties of the Poisson process. According to this model the stream of successfully delivered messages from ith node might be defined as Poisson with intensity $\lambda_i p$, where $p = e^{-2Q\Lambda}$, and the total stream of successfully delivered messages from all nodes is Poisson with intensity Λp.

Remark 1: If $\lambda_i = \lambda, \; i = 1, 2, \ldots, N$, then the intensity of the total stream of successfully delivered messages Λp reaches its maximum at $\Lambda = \frac{1}{2Q} \min\left\{1, \frac{N}{50}\right\}$. It easy to note that with

$$\lambda = \frac{1}{2Q}\frac{1}{\max\{N, 50\}} \tag{2}$$

the minimal average time between two successfully delivered messages to the operator from every node is equal to

$$2Q \max\{N, 50\} e^{\min\left\{1, \frac{N}{50}\right\}}. \tag{3}$$

B. VOLTAGE MONITORING MODEL

Consider a message transmission model for monitoring voltages of LV electrical grids with DERs using the LoRa technology. In addition to the uncertainty due to random intervals between successfully delivered messages, the voltages themselves depend on stochastic factors: solar radiation, users' energy consumption, etc. Therefore, we assume that the node voltage is a random process $U_i(t), \; t \geq 0, \; i = 1, 2, \ldots, N$. For brevity, we will sometimes omit the node index i. Let $U(t)$ be a Markov random process with paths in the space of functions without second-order discontinuities.

The grid operator is required to monitor the risk of the voltage exceeding the upper voltage limit (UL). So at each moment of successful message delivery, the operator needs to decide whether to take costly voltage reduction measures, for example disconnecting the DER from the grid, or to wait until the next successfully delivered message arrives. The decision depends on the current voltage value of the node and the probability that an excess UL will occur before the next successfully delivered message from the same node arrives. At the moment of receiving a successfully delivered message from a particular node, it does not matter to the grid operator what voltage was previously in this node. Therefore, due to the Markov property, we can take this moment as the initial $t = 0$ and assume that $U(0) < UL$. Let τ be an exponentially distributed random variable with parameter λp, for some $\lambda > 0$. By "risk" we mean the probability that the path $U(t)$ crosses the threshold UL during interval τ between two successfully delivered messages:

$$\mathbf{P}\left(\overline{U}\left(\tau\right) > B\right), \tag{4}$$

where $\overline{U}(t) = \sup_{0 \leq s \leq t} U(s)$, $B = UL - U(0)$.

We argue that instead of relying on equal intensity λ, each node sends messages with various parameters λ_i, $i = 1, 2, \ldots, N$, which we can choose based on minimizing the maximum risk (probability of exceeding the upper voltage limit among all nodes given they have the same initial value). Formulating such an optimization problem is not straightforward. Therefore, we first formulate the problem under some constraints before finding its solution for the particular case.

III. FORMULATING THE OPTIMIZATION PROBLEM

It is known that obtaining an explicit distribution of $\overline{U}(t)$ is a very difficult problem (see [20]). However, under some general constraints, it is more feasible to obtain only its Laplace transform over the time component. Note that the relation (4) can be rewritten as:

$$\mathbf{P}\left(\overline{U}\left(\tau\right) > B\right) = \int_0^\infty \lambda p e^{-\lambda p t} \mathbf{P}\left(\overline{U}(t) > B\right) dt.$$

Assumption 4: Let $U(t)$ be a spectrally negative Lévy process, i.e., a homogeneous random process with independent increments and without positive jumps.

We believe this assumption to be quite realistic since, from the practical point of view, we are interested in the behavior of the Markov process $U(t)$ near point (upper voltage limit) UL during a finite time interval τ.

The Laplace transform $\mathbf{E}e^{\mu U(1)}$ (reminder that $\mathbf{E}e^{\mu U(t)} = \left(\mathbf{E}e^{\mu U(1)}\right)^t$) for spectrally negative processes is defined for all $\mu \geq 0$ and the equation

$$\log \mathbf{E}e^{\mu U(1)} = s, \; s > 0, \tag{5}$$

has a non-negative solution $\mu = \mu(s) \geq 0$, where $\mu(0) = q \geq 0$ (see [20], [21]). It is then well known that the Laplace transform on the time component of the supremum distribution has a simple explicit form (see [20], [21]):

$$s \int_0^\infty e^{-st} \mathbf{P}\left(\overline{U}(t) > B\right) dt = e^{-\mu(s)B},$$

which can be presented as:

$$\mathbf{P}\left(\overline{U}(\tau) > B\right) = e^{-\mu(\lambda p)B}.$$

Let us first formulate the optimization problem: we have independent exponentially distributed random variables $\tau_1, \tau_2, \ldots, \tau_N$ with parameters $\lambda_i p$ and we need to find $\left(\lambda_1^*, \lambda_2^*, \ldots, \lambda_N^*\right)$ at which, for any $B > 0$

$$\max\left\{\mathbf{P}\left(\overline{U}_1(\tau_1) > B\right), \ldots, \mathbf{P}\left(\overline{U}_N(\tau_N) > B\right)\right\} \to \min.$$

Then under Assumption 4 the optimization problem is formulated as

$$\min\{\mu_1\left(\lambda_1 p\right), \ldots, \mu_N\left(\lambda_N p\right)\} \to \max, \tag{6}$$

where $\mu_i(\cdot)$ is the solution to the equation (5) for the voltage $U_i(t)$ of node i and $p = e^{-2Q\Lambda}$.

IV. SOLUTION TO THE OPTIMIZATION PROBLEM FOR A BROWNIAN MOTION

As an example, let us find a solution of the optimization problem for one particular case of spectrally negative Lévy process.

Assumption 5: Suppose that the voltage of each node is described by a standard Brownian motion $W_i(t)$, $i = 1, 2, \dots, N$, (see [22]):

$$U_i(t) = U_i(0) + \sigma_i W_i(t), \quad \sigma_i > 0. \tag{7}$$

Then the solution to the equation (5) is $\mu_i(s) = \frac{\sqrt{2s}}{\sigma_i}$ (see [21]). So, the problem (6) can be reformulated as:

$$\min \left\{ \frac{\lambda_1}{\sigma_1^2} e^{-2Q\Lambda}, \dots, \frac{\lambda_N}{\sigma_N^2} e^{-2Q\Lambda} \right\} \to max.$$

That is, instead of searching for equal minimal intervals between messages (3) we are looking for equal minimal probabilities of exceeding the upper voltage limit. So, we need to find

$$\frac{\lambda_1}{\sigma_1^2} e^{-2Q\Lambda} \to max, \tag{8}$$

subject to the constraints (1) and

$$\frac{\lambda_i}{\sigma_i^2} = \frac{\lambda_j}{\sigma_j^2}, \; i \neq j.$$

If we denote $h = \frac{\lambda_i}{\sigma_i^2}$ then

$$\Lambda = h \times S,$$

where $S = \sum_{i=1}^{N} \sigma_i^2$, and due to Assumption 3

$$h \leq \frac{1}{2Q} \frac{1}{50} \frac{1}{\max \sigma_i^2}.$$

The optimization problem (8) might be rewritten as

$$he^{-2QhS} \to max.$$

It is easy to note that if

$$S \geq 50 \max_i \sigma_i^2$$

then at $h^* = \frac{1}{2Q} \frac{1}{S}$

$$\lambda_i^* = \frac{1}{2Q} \frac{\sigma_i^2}{S},$$

$$\max_h he^{-2QhS} = \frac{1}{2Q} \frac{1}{S} e^{-1}. \tag{9}$$

If

$$S < 50 \max_i \sigma_i^2$$

then at $h^* = \frac{1}{2Q} \frac{1}{50 \max_i \sigma_i^2}$

$$\lambda_i^* = \sigma_i^2 h^*, \quad \max_h he^{-2QhS} = h^* e^{-2Qh^* S}.$$

Let us summarize this section. According to our model, we have the parameters of the LV electrical grid – the number of nodes N and the value of σ_i for each node – and the parameters of the communication network – the number of nodes N and airtime Q. As a result of the risk minimization problem for the communication network with the optimal intensity of message transmission λ_i^* for each node is obtained.

V. SF ALLOCATION

The LoRa technology uses distributed spectrum modulation with 6 orthogonal spreading factors ($SF = 7, 8, \dots, 12$). Messages with different SF can be transmitted simultaneously. A smaller SF provides a higher data rate. A larger SF increases the receiver's sensitivity and therefore the range of the system [1], [2], [25].

There are many works (see, for example, [25], [26]), where the problem of SF allocation is considered without taking into account the specifics of the transmitted data.

We can reduce our risk by allocating nodes to different SF. We need to split the set of indices of all nodes $G = \{1, 2, \dots, N\}$ into $2 \leq k \leq 6$ subsets to reduce maximal risk among nodes with $SF = 13 - k, 13 - k + 1, \dots, 12$. It means to divide G into k sets $G_{13-k}, G_{13-k+1}, \dots G_{12}$, such that $G = G_{13-k} \cup \dots \cup G_{12}$, $G_i \cap G_j = \emptyset$, $i \neq j$, where G_i – node indices with $SF = i$ and G_j – node indices with $SF = j$. We assume that message transmission in each of the SF occurs independently of the nodes with other SF.

Let Q_j be airtime for $SF = j$, $j = 7, 8, \dots 12$, and

$$S_j = \sum_{i \in G_j} \sigma_i^2.$$

The optimization problem for allocating nodes among k sets might be written as

$$F = \min \{F_{13-k}, F_{13-k+1}, \dots, F_{12}\} \to \max, \tag{10}$$

where

$$F_j = h_j^* e^{-2Q_j h_j^* S_j}, \; j = 13 - k, 13 - k + 1, \dots, 12,$$
$$h_j^* = \frac{1}{2Q_j} \frac{1}{\max\{50 \max_{i \in G_j} \sigma_i^2, S_j\}}.$$

The optimization problem (10) implies an optimal partitioning of the node indices, which is rather difficult to find.

The idea is that nodes with large values of σ_i should send messages more often than others. For convenience, we index the nodes in ascending order such that

$$\sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_N.$$

Since the message airtimes for different spreading factors satisfy the ratio:

$$Q_7 \leq Q_8 \leq \dots \leq Q_{12}$$

let us not consider all possible partitions, but only "sequential" ones:

$$G_{12} = \{1, 2, \ldots, i_{12}\},$$
$$G_{11} = \{i_{12}+1, i_{12}+2, \ldots, i_{11}\},$$
$$\ldots,$$
$$G_{13-k} = \{i_{13-k+1}+1, i_{13-k+1}+2, \ldots, N\},$$

where indices

$$i_{12}, i_{11}, \ldots i_{13-k+1} \in G,$$
$$i_{12} \le i_{11} \le \cdots \le i_{13-k} = N.$$

It is easy to find numerically the solution F_s^* of the optimization problem (10) with sequential partition of G and it is clear that

$$F_s^* \le F^*,$$

where F^* – optimal solution of (10) without restriction on sequential partition of the set G.

Note that without condition (1)

$$F_k = \frac{1}{2Q_{13-k}S_{13-k}e}$$

and the objective function is of the form:

$$\max\{2eQ_{13-k}S_{13-k}, \ldots, 2eQ_{12}S_{12}\} \to \min.$$

If we define

$$\nu = \max_{13-k \le j \le 12} Q_j S_j,$$

then

$$\nu \ge Q_j S_j, \quad j = 13-k, 13-k+1, \ldots, 12,$$

and we can perform our task as a well known linear integer programming problem.

Let x be a vector from $N \times k$ elements

$$x = (x_1, x_2, \ldots, x_{kN}), \; x_i \in \{0, 1\},$$

where

$$x_i + x_{N+i} + \cdots + x_{(k-2)N+i} + x_{(k-1)N+i} = 1, \; i = 1, 2, \ldots, N.$$

Put

$$s_{13-k} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2, 0, \ldots, 0),$$
$$s_{13-k+1} = (0, \ldots, 0, \sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2, 0, \ldots, 0),$$
$$\ldots,$$
$$s_{12} = (0, \ldots, 0, \sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2).$$

For brevity, we can rewrite $S_j = \sum_{i \in G_j} \sigma_i^2 = s_j x^T$, $j = 13-k, 13-k+1, \ldots, 12$, where x^T is a transposed vector x.

It is to formulate the optimal SF allocation problem:

$$\nu \to \min,$$
$$\nu - Q_j s_j x^T \ge 0, \; j = 13-k, 13-k+1, \ldots, 12. \tag{11}$$
$$x_i \in \{0, 1\}, \; i = 1, 2, \ldots, kN,$$
$$x_i + x_{N+i} + \cdots + x_{(k-2)N+i} + x_{(k-1)N+i} = 1, \; i = 1, 2, \ldots, N. \tag{12}$$

Remark that integer programming is quite a consuming procedure. Therefore, if we replace condition (12) with condition

$$0 \le x_i \le 1, \; i = 1, 2, \ldots, kN,$$

it is easy to obtain solution ν^* of linear programming problem (11).

If we put $F_{lp}^* = 1/(2e\nu^*)$, then we can obtain that

$$F_s^* \le F^* \le F_{lp}^*.$$

This means that if we consider the solution F_s^* of the optimization problem (10) with sequential partitioning, this solution differs from the optimal F^* less than the difference between F_s^* and the solution to the linear programming problem F_{lp}^*, i.e.

$$F^* - F_s^* \le F_{lp}^* - F_s^*.$$

VI. NUMERICAL EXAMPLE

As an example, we consider an LV grid with $N = 150$ nodes. Each node has a LoRaWAN communication network device of Class A. Suppose that due to technical limitations (distance to devices, interference) the transmission can be carried out using spreading factor $SF = 11$, which ensures successful message delivery in the absence of collisions.

Airtime $Q \equiv Q_{11} = 0.8233$ sec. Next, we show how to calculate σ_i parameters for our LV grid and compare the risk values (4) of our model with equal probabilities with respect to the message transmission model with equal message intervals (2).

Consider a grid, each node of which has PV module. We denote the vector of voltages of the grid nodes as

$$U(t) = (U_1(t), U_2(t), \ldots, U_N(t))^T, \; t > 0,$$

given that initial node is balancing, i.e. $U_0(t) = U_0$ is known constant. Define also the current vector

$$I(t) = (I_1(t), I_2(t), \ldots, I_N(t))^T, \; t > 0.$$

The relationship between the vector of voltages and the vector of currents might be expressed as

$$U(t) = U_0 + H \times I(t),$$

where H is the matrix of nodal resistances and $U_0 = (U_0, U_0, \ldots, U_0)^T$.

Suppose we have three-phase LV grid and every phase has linear topology with M nodes, where $3M = N$. For simplicity we assume that phases are independent of each other and

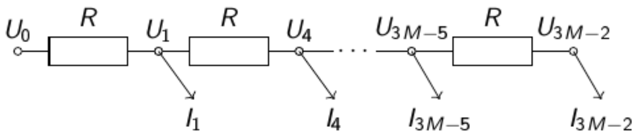

FIGURE 1. The phase-line of the linear LV grid.

active resistances R between the nodes in every phase are the same. We neglect reactive parameters of the LV grid. We number all the nodes so that the phase p, $p = 1, 2, 3$, consists of nodes with indices $3(s-1)+p$, $s = 1, 2, \ldots, M$ (see Fig. (1) with phase $p = 1$).

Then (see, for example, [23])

$$H = -R \times \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ 1 & 2 & 2 & \ldots & 2 & 2 \\ & & & \ldots & & \\ 1 & 2 & 3 & \ldots & M-1 & M \end{pmatrix}$$

or

$$U_{3(s-1)+p}(t) = U_0 - R\left(\sum_{r=1}^{s} rI_{3(r-1)+p}(t) + s\sum_{r=s}^{M-1} I_{3r+p}(t)\right),$$

where U_0 is the known voltage at the beginning of the line.

We suppose that the values of currents $I_i(t)$, $i = 1, 2, \ldots, N$, are the algebraic sum of the current from the PV unit $IS_i(t)$ in the node i and the consumption current $J_i(t)$ from the node to the customer (see [23]):

$$I_i(t) = -IS_i(t) + J_i(t).$$

Let $IS_i(t)$ and $J_i(t)$ be described by means of independent Brownian motion processes, i.e.

$$I_i(t) = I_i(0) - \hat{\sigma}_i \hat{W}(t) + \tilde{\sigma}_i \tilde{W}_i(t),$$

where i is the node index, $\hat{W}(t)$, $\tilde{W}_i(t)$, $i = 1, 2, \ldots, N$, are independent standard Brownian motions, $I_i(0)$ corresponds to the current value at the initial time, $\hat{\sigma}_i$ is the standard deviation for the PV node current and $\tilde{\sigma}_i$ is the standard deviation for the consumption current in the node. For simplicity we put

$$\hat{\sigma}_i = \hat{\sigma}, \quad \tilde{\sigma}_i = \tilde{\sigma}, \; i = 1, 2, \ldots, N.$$

We can find the probability distribution of every node voltage, including last nodes in phases, where the maximal amplitude of voltage fluctuations is achieved. Note that

$$(1+2+\ldots+M)\hat{W}(t) = \frac{M(M+1)}{2}\hat{W}(t),$$

$$\tilde{W}_1(t) + 2\tilde{W}_4(t) + \ldots + M\tilde{W}_{3M-2}(t)) \overset{d}{=} \sqrt{\frac{M(M+1)(2M+1)}{6}}\tilde{W}(t),$$

where $\hat{W}(t)$ and $\tilde{W}(t)$ are independent standard Brownian motions, designation ($\overset{d}{=}$) means the same probability distribution. Thus

$$U_{3(M-1)+p}(t) \overset{d}{=} U_{3(M-1)+p}(0) + \sigma_{3(M-1)+p} W_p(t), \qquad (13)$$

TABLE I. Parameters of the model.

Parameter	Value
Number of nodes (N)	150 nodes
Number of phases	3
Number of nodes in each phase (M)	50 nodes
Payload	10 bytes
Spreading Factor (SF)	11, 12
Airtime (Q_{11})[1]	0.8233 s.
Airtime (Q_{12})[1]	1.4828 s.
Active resistance (R)	0.01 Ω
Standard deviation (SD) of PV current ($\tilde{\sigma}$)	0.01 A
SD of consumption current ($\hat{\sigma}$)	0.018 A
Upper voltage limit (UL)	253 V
Initial value ($U_{3(M-1)+p}(0), j = 1, 2, 3$)	248 V
B	5 V

where $W_p(t)$, $p = 1, 2, 3$, are standard Brownian motions, $U_{3(M-1)+p}(0)$ – node voltage at some initial time and

$$\sigma^2_{3(M-1)+p} = R^2\left(\hat{\sigma}^2\frac{M^2(M+1)^2}{4} + \tilde{\sigma}^2\frac{M(M+1)(2M+1)}{6}\right).$$

That is, we defined σ_i (see (7)) from the Assumption 5 for our example with simplified three-phase linear LV grid.

Firstly, based on our assumptions, we can find and compare the maximal probabilities of exceeding the upper voltage limit among all nodes (maximal risk) for the case with equal intensities (2) and the case when the specifics of the transmitted data are taken into account for one $SF = 11$.

Secondly, the communication system, which at $SF = 11$ ensures successful message delivery in the absence of collisions, also ensures successful delivery at $SF = 12$. We define airtimes for every SF as Q_{11} and Q_{12}. In considering several SF, we assume that all assumptions remain valid except for one spreading factor limitation. We need to divide indices of all nodes $G = \{1, 2, \ldots, N\}$ into two sets G_{11} and G_{12}, such that $G = G_{11} \cup G_{12}$, $G_{11} \cap G_{12} = \emptyset$, where G_{11} – node indices with $SF = 11$ and G_{12} – node indices with $SF = 12$. We assume that message transmission in each of the SF occurs independently of the nodes with other SF.

This allows us to find and compare maximal risks between cases with equal and various intensities of transmission.

Let us first consider the case where all nodes belong to the same spreading factor $SF = 11$.

We take $UL = 253\ V$ since the voltages in a LV grid must not exceed the level $UL = U_{nom} + 10\%$, where $U_{nom} = 230V$ is the nominal voltage, over 5% time (see [24]). The other parameters of the example are given below (see Table I).

For the case where all nodes belong to the same $SF = 11$ and send messages with equal intensities (2), it is easy to obtain that

$$\max\left\{\mathbf{P}\left(\overline{U}_1(\tau_1) > 5\right), \ldots, \mathbf{P}\left(\overline{U}_N(\tau_N) > 5\right)\right\} = \exp(-0.4109 \times 5) = 0.1281.$$

TABLE II. Maximal risk.

	Maximal risk with SF 11	Maximal risk with SF 11 and 12
Equal λ for every SF	0.1281	0.0998
Various λ_i for nodes and every SF	0.0605	0.031

To study the maximal risk among all nodes, which have various intensities, we consider individual probabilities of exceeding the upper voltage limit. In addition, as we have already mentioned, the last in the phases nodes will have the maximum voltage variance, therefore

$$\max\left\{\mathbf{P}\left(\overline{U}_1(\tau_1) > 5\right), \ldots, \mathbf{P}\left(\overline{U}_N(\tau_N) > 5\right)\right\}$$
$$= \max\left\{\mathbf{P}\left(\overline{U}_1(\tau_1) > 5\right), \ldots, \mathbf{P}\left(\overline{U}_{3M-2}(\tau_{3M-2}) > 5\right)\right\}$$
$$= \mathbf{P}\left(\overline{U}_{3M-2}\left(\tau_{3M-2}\right) > 5\right) = \exp(-0.5611 \times 5) = 0.0605.$$

Let us consider our problem for the case $SF = 11, 12$. Instead of finding the optimal node partitioning (see (10)), consider sequential partitioning with $k = 2$ and

$$G_{12} = \{1, 2, \ldots, i_{12}\},\ G_{11} = \{i_{12}+1, i_{12}+2, \ldots, N\}. \tag{14}$$

It is easy to obtain that $i_{12} = 92$,

$$F_s^* = 0.2412,$$

and find the solution of the linear programming problem (11)

$$F_{lp}^* = 0.2448.$$

Thus for two spreading factors and various message sending intensities we obtain that maximal risk with sequential partitioning

$$\exp\left(-\sqrt{2F_s^*} \times 5\right) = \exp(-0.6947 \times 5) = 0.031.$$

Let us find the maximum risk for partitioning (14) and equal message sending intensities within each spreading factor:

$$\exp(-\min(0.6607, 0.4608) \times 5) = 0.0998.$$

The table below (see Table II) shows the values of maximal risk among all nodes for one spreading factor (first column) and two spreading factors with partitioning (14) and $i_{12} = 92$ (second column), for equal (first row) and various message intensities (second row).

Fig. (2) displays on the ordinate axis the values of maximal risk at $B = 5$ for the cases of equal (blue line) and various (purple line) message sending intensities, where the abscissa axis displays the values of i_{12} when nodes are sequentially partitioned (see (14)) into $SF = 11$ and $SF = 12$.

Thus, by accounting for the specifics of the data being transmitted, it is possible to choose various parameters for sending messages from different nodes in order to reduce the probability of an unwanted event.

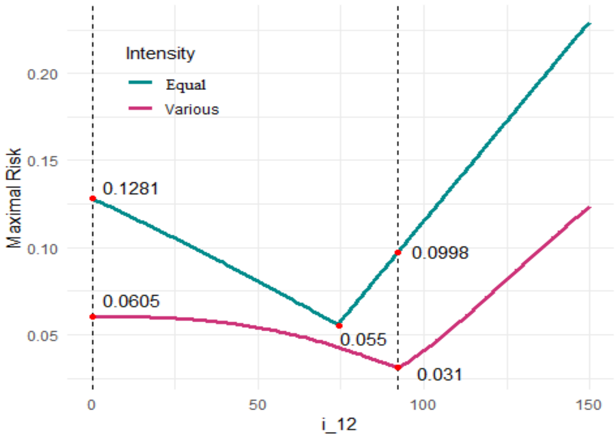

FIGURE 2. Values of maximum risk at $B = 5$ when sequentially partitioned by $k = 2$ SF for cases of equal and various message sending intensities .

VII. ACCURACY ASSESSMENT

In partial example conditions, we provide a number of simulation examples to compare them with theoretical derivations.

Assume that in the example conditions all nodes belong to the same spreading factor $SF = 11$.

If all nodes send messages with equal intensity, then the average time between two successfully delivered messages to the operator from every node is equal to (see (3))

$$2Q_{11}Ne = 671.3 \text{ s}.$$

The simulation results for the lengths of intervals between successfully delivered messages when nodes send messages with various intensities (see (9)) of exponential distribution are listed in Table III. The first column "Index" shows the node indices, the second column "Model" shows our theoretical values

$$\frac{1}{\lambda_i^* p} = \frac{2Q_{11}Se}{\sigma_i^2}$$

under the assumption of exponential distribution of time between messages (Assumption 1). The third column "Sim. (Exp)" shows the simulation results for the lengths of the intervals between successfully delivered messages. The results in the third column are obtained by simulating the intervals between messages with a shifted exponential distribution $Q_{11} + \theta$, where θ has an exponential distribution with parameter λ_i^*. Assumption 1 regarding the random exponential distribution of message intervals is necessary to derive rigorous mathematical conclusions. In practice, some equipment manufacturers for LoRaWAN networks allow configuring the transmission of messages not through fixed, but through random, uniformly distributed intervals of time. Thus, the fourth column shows the simulation results when each node sends messages in independent random intervals with a shifted uniform distribution of $Q_{11} + \alpha$, where α has a uniform distribution on $(0, 2/(\lambda_i^* p))$.

TABLE III. Comparison of the average interval length between successfully delivered messages according to the model and simulation results .

Index	Average interval (s.)		
	Model	Sim. (Exp)	Sim. (Unif)
25	3469	3483	3508
50	1162	1174	1165
75	654.1	651.1	646
100	453.0	444	448
110	418.8	416.7	409
120	394.0	389.8	394
130	372.4	370.7	367
140	363.5	360.3	356
150	359.9	357.9	359

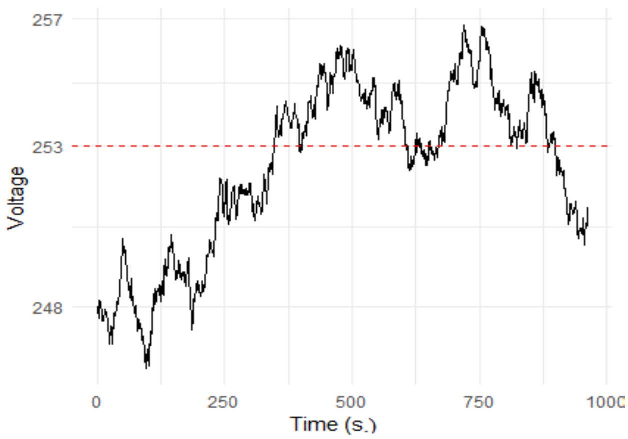

FIGURE 3. Voltage trajectory between successfully delivered messages.

TABLE IV. Simulation results for risk with equal, exponentially and uniformly distributed interval lengths between messages .

Index	Risk		
	Equal	Var. (Exp)	Var. (Unif)
25	0.0011	0.0623	0.0569
50	0.0249	0.0602	0.0559
75	0.0597	0.0607	0.0562
100	0.0973	0.0602	0.0564
110	0.1086	0.0565	0.0536
120	0.1167	0.0595	0.0587
130	0.1177	0.0637	0.0553
140	0.1297	0.0596	0.0564
150	0.1249	0.0599	0.0564

The results in Table III show that rigorous theoretical conclusions derived from the mathematical model are well supported by the simulation results.

The work is devoted to the problem of the optimal choice of interval lengths to minimize the maximum risk among all nodes, that is, the probability of exceeding the upper voltage limit, provided the nodes have the same initial value. This means that if the operator receives a message from the ith node with a voltage of 248 V, the operator is interested in the risk of the voltage exceeding the upper voltage limit of 253 V before the next successfully delivered message. The simulation results for one such event are presented in Fig. 3.

We reduced this risk by choosing various interval lengths between the messages. If the average interval is the same for all nodes, the maximum risk is 0.1281; if it varies according to (9), the maximum risk is 0.0605. The Table IV shows the simulation results of the risk value for varoius nodes ("Index") for the cases of equal average interval ("Equal") and various intervals ("Var. (Exp)"). The fourth column ("Var. (Unif)") shows the simulation results for the risk when each node sends messages at shifted uniformly distributed intervals.

TABLE V. List of abbreviations.

Abbreviation	Explanation
A	Ampere
DER	Distributed Energy Resources
Equal	the simulation results for the risk when each node sends messages with equal average intervals
LoRaWAN	Long Range Wide-Area Networks
Index	Index of the node
LV	Low Voltage
M	number of phases
Model	conclusions derived from the mathematical model
N	Number of nodes
PV	Photovoltaic
R	active Resistance
s.	seconds
SF	Spreading Factor
SD	Standard Deviation
Sim.(Exp)	the simulation results for the time between successfully delivered messages and the intervals between them, which have a shifted exponential distribution.
Sim.(Unif)	the simulation results for the time between successfully delivered messages and the intervals between them, which have a shifted uniform distribution
UL	Upper Voltage Limit
U_{nom}	Nominal Voltage
V	Volt
Var.(Exp)	the simulation results for the risk with shifted exponentially distributed intervals
Var.(Unif)	the simulation results for the risk with shifted uniformly distributed intervals
Ω	Ohm

VIII. CONCLUSION

We studied the communication system for monitoring the voltages of the electrical grid, based on LoRaWAN technology, which allows the creation of low-cost and energy-efficient data collection and transmission systems. We considered two related objects – the LV electrical grid with DER and the communication network, which used LoRaWAN technology to transmit node voltage monitoring data. The model, reflecting the main specific features of these objects – the random character of voltage changes in the electrical grid and the possible loss of data due to conflicts in the communication network – was proposed. The notion of risk was introduced – the probability of loss of information about voltage exceeding the upper limit. The optimization problem of minimizing the maximal risk by selecting the parameters of the communication network, based on the properties of a random voltage changes in the grid, was formulated. The solution of this problem was found and illustrated by the example for the voltages described by Brownian motions.

For the convenience of readers, we present Table V with the abbreviations used in the article.

According to our model, each node sends messages to the grid operator at exponentially distributed with the parameter λ_i random time intervals for each node i. Due to possible collisions, sent messages can be lost. Thus, the grid operator estimates the voltages of the grid nodes only on the basis of successfully delivered messages.

Therefore, each time a successfully delivered message is received from a certain node, the operator must estimate the risk that the voltage at that node will exceed the threshold level UL until the next successfully delivered message:

$$\mathbf{P}\left(\overline{U}_i(\tau_i) > B\right), \ B = UL - U_i(0)$$

where i – index of the node, $U_i(0)$ – voltage value at the moment of receiving the last successfully delivered message, $\overline{U}_i(t) = \sup_{0 \leq s \leq t} U_i(s)$, τ_i – exponentially distributed random variable with parameter $\lambda_i p$, UL – upper voltage limit.

If this probability (or risk) is "high", the operator takes costly measures to avoid this event, for example turning off solar generation at the node. Otherwise, it is reasonable to wait until the next successfully delivered message. According to our model, one must choose a higher message sending intensity λ_i in nodes where the risk of exceeding the threshold level UL is higher (provided the initial values of $U_i(0)$ are the same) than in less risky nodes in order to equalize risks. In LV grid the voltage fluctuations of the nodes distant from the transformer are larger than those close to it, so it is natural to send messages from distant nodes more often.

As an extension of our problem, it is natural to consider a model with several spreading factors. We considered the problem of optimal partitioning of nodes by spreading factors and studied the maximal risks of exceeding the threshold among all nodes.

According to requirements [24], the voltage should be between certain upper and lower limits. In our work, we focused only on the risk of "over voltage" and did not study "under voltage" case, which is similar to the "over voltage," but is not the concern of our study. The problem that considers both over and under voltage risks is more difficult. If we consider the voltage as a random process under general requirements, then only the asymptotic results for the problem with two boundaries are known [27], [28]. They are more complicated for analysis than the probability of exceeding the upper voltage limit and can be used for further research.

We proposed an approach to the selection of communication network parameters, which took into account the properties of the randomly changing parameters of the object of monitoring. This approach might be applied for various monitoring systems, in which LoRaWAN or any other network with random multiple access channel was used as communication network.

The authors would like to thank Dr. Dmitri Moltchanov for help and sound advice.

REFERENCES

[1] F. S. D. Silva, E. P. Neto, H. Oliveira, D. Rosário, E. Cerqueira, C. Both, S. Zeadally, and A. V. Neto, "A survey on long-range wide-area network technology optimizations," *IEEE Access*, vol. 9, pp. 106079–106106, 2021, doi: 10.1109/ACCESS.2021.3079095.

[2] M. Jouhari, N. Saeed, M.-S. Alouini, and E. M. Amhoud, "A survey on scalable LoRaWAN for massive IoT: Recent advances, potentials, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 3, pp. 1841–1876, 3rd Quart., 2023, doi: 10.1109/COMST.2023.3274934.

[3] A. P. Kenneth and K. Folly, "Voltage rise issue with high penetration of grid connected PV," *IFAC Proc. Volumes*, vol. 47, no. 3, pp. 4959–4966, 2014, doi: 10.3182/20140824-6-za-1003.01989.

[4] H. Rajab, T. Cinkler, and T. Bouguera, "IoT scheduling for higher throughput and lower transmission power," *Wireless Netw.*, vol. 27, no. 3, pp. 1701–1714, Apr. 2021, doi: 10.1007/s11276-020-02307-1.

[5] O. Gusev, A. Turlikov, S. Kuzmichev, and N. Stepanov, "Data delivery efficient spreading factor allocation in dense LoRaWAN deployments," in *Proc. XVI Int. Symp. 'Problems Redundancy Inf. Control Syst.' (REDUNDANCY)*, Moscow, Russia, Oct. 2019, pp. 199–204, doi: 10.1109/REDUNDANCY48165.2019.9003343.

[6] A. Burkov, "Signal power and energy-per-bit optimization problems in systems mMTC," *Inf. Control Syst.*, vol. 10, no. 5, pp. 51–58, Oct. 2021, doi: 10.31799/1684-8853-2021-5-51-58.

[7] M. Heusse, C. Caillouet, and A. Duda, "Performance of unslotted ALOHA with capture and multiple collisions in LoRaWAN," *IEEE Internet Things J.*, vol. 10, no. 20, pp. 17824–17838, Oct. 2023.

[8] O. T. Sanchez, D. Raposo, A. Rodrigues, F. Boavida, and J. S. Silva, "Private LoRaWAN network gateways: Assessment and monitoring in the context of IIoT-based management," *Eng. Proc.*, vol. 47, no. 4, pp. 1–7, 2023, doi: 10.3390/engproc2023047004.

[9] L.-C. Hwang, C.-S. Chen, T.-T. Ku, and W.-C. Shyu, "A bridge between the smart grid and the Internet of Things: Theoretical and practical roles of LoRa," *Int. J. Electr. Power Energy Syst.*, vol. 113, pp. 971–981, Dec. 2019, doi: 10.1016/j.ijepes.2019.06.001.

[10] G. Y. Odongo, R. Musabe, D. Hanyurwimfura, and A. D. Bakari, "An efficient LoRa-enabled smart fault detection and using self-powered IoT devices," *IEEE Access*, vol. 10, pp. 73403–73420, 2022.

[11] G. Premsankar, B. Ghaddar, M. Slabicki, and M. D. Francesco, "Optimal configuration of LoRa networks in smart cities," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7243–7254, Dec. 2020.

[12] S. H. Gupta, A. Singh, A. Tyagi, and J. S. Jadon, "Optimizing LPWAN using particle swarm for the smart city," in *Computational Intelligence in Urban Infrastructure*. Boca Raton, FL, USA: CRC Press, 2024, pp. 56–76.

[13] M. Hao, N. Kou, and C. Zeng, "Intelligent sensing and monitoring system for high-voltage transmission line status of smart grid based on IoT technology," *World Electr. Vehicle J.*, vol. 14, no. 8, p. 224, Aug. 2023.

[14] A. Farhad and J.-Y. Pyun, "LoRaWAN meets ML: A survey on enhancing performance with machine learning," *Sensors*, vol. 23, no. 15, p. 6851, Aug. 2023.

[15] S. Hosseinzadeh, M. Ashawa, N. Owoh, H. Larijani, and K. Curtis, "Explainable machine learning for LoRaWAN link budget analysis and modeling," *Sensors*, vol. 24, no. 3, p. 860, Jan. 2024.

[16] M. M. Mijwil, "Artificial neural networks advantages and disadvantages," *Mesopotamian J. Big Data*, vol. 2021, pp. 29–31, Aug. 2021.

[17] M. A. Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and apache spark," *IEEE Netw.*, vol. 30, no. 3, pp. 22–29, May 2016.

[18] (2017). *LoRaWAN 1.1 Regional Parameters, LoRa Alliance*. [Online]. Available: https://lora-alliance.org/wp-content/uploads/2020/11/lorawan-regional-parameters-v1.1ra.pdf

[19] D. K. Kim, G. D. Georgiev, and N. V. Markovskaya, "A model of random multiple access in unlicensed spectrum systems," in *Proc. Wave Electron. Appl. Inf. Telecommun. Syst. (WECONF)*, May 2022, pp. 1–4, doi: 10.1109/WECONF55058.2022.9803810.

[20] B. A. Rogozin, "Distribution of the maximum of a process with independent increments," *Siberian Math. J.*, vol. 10, no. 6, pp. 989–1010, 1970, doi: 10.1007/bf00990775.

[21] A. E. Kyprianou, *Introductory Lectures on Fluctuations of Lévy Processes With Applications*. New York, NY, USA: Springer, 2006.

[22] F. Milan and R. Zárate-Miñano, "A systematic method to model power systems as stochastic differential algebraic equations," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4537–4544, Nov. 2013, doi: 10.1109/TPWRS.2013.2266441.

[23] G. Georgiev, D. Kim, and M. Negnevitsky, ''Probabilistic approximation of voltage values for power distribution network with large number of nodes and solar generation,'' in *Proc. 10th Int. Sci. Symp. Elect. Power Eng.*, 2019, pp. 53–57.

[24] *EN 50160: Voltage Characteristics of Electricity Supplied By Public Distribution Systems*, Eur. Committee for Electrotechnical Standardization, Brussels, Belgium, 2011.

[25] D. Bankov, E. Khorov, and A. Lyakhov, ''LoRaWAN modeling and MCS allocation to satisfy heterogeneous QoS requirements,'' *Sensors*, vol. 19, no. 19, p. 4204, Sep. 2019, doi: 10.3390/s19194204.

[26] D. Garlisi, I. Tinnirello, G. Bianchi, and F. Cuomo, ''Capture aware sequential waterfilling for LoRaWAN adaptive data rate,'' *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 2019–2033, Mar. 2021, doi: 10.1109/TWC.2020.3038638.

[27] V. I. Lotov, ''Asymptotic analysis of distributions in problems with two boundaries. I,'' *Theory Probab. Appl.*, vol. 24, no. 3, pp. 480–491, Jan. 1980, doi: 10.1137/1124059.

[28] V. I. Lotov, ''Asymptotic analysis of distributions in problems with two boundaries. II,'' *Theory Probab. Appl.*, vol. 24, no. 4, pp. 869–876, Jan. 1980, doi: 10.1137/1124103.

DMITRIY KIM received the bachelor's degree in mathematics from Al-Farabi Kazakh National University, Kazakhstan, in 2000, and the master's and Ph.D. degrees in probability theory and mathematical statistics from Novosibirsk State University, Russia, in 2002 and 2005, respectively. He has been a Researcher at a number of Kazakhstani and European universities. Since 2017, he has been with the School of Digital Technologies, Narxoz University, Almaty, Kazakhstan. He specializes in the boundary problems for random processes and in application of probabilistic methods for modeling complex systems.

ANDREY TURLIKOV is currently the Head of the Internet of Things and Cyber-Physical Systems Laboratory, Higher School of Economics (HSE), HSE Campus in St. Petersburg, Russia. He has more than 35 years of experience in the area of telecommunications. During the last ten years, he participated in and supervised over ten successful long-term research projects, focusing on next-generation telecommunications technologies (3GPP LTE-Advanced, IEEE 802.16m, prominent IEEE 802.11-based solutions). His research interests include energy efficiency improvement, collaborative techniques, heterogeneous networking, and support for advanced services toward standardization in IEEE.

GEORGI GEORGIEV received the master's and Ph.D. degrees in electrical engineering from Moscow Power Engineering Institute, in 1978 and 1985, respectively. His professional career was entirely related to various responsible positions in Bulgarian electricity company, namely the Head of the Territorial Dispatch Center, from 1986 to 1995, and the Department for Steady State and Transient Analysis, from 1996 to 2003, and the Chief Specialist in the design, programming, and implementation of a pilot project for remote measurement and control of electricity meters using power line communication, from 2004 to 2006. From 2007 to 2011, he was the Head of investments in a large thermal power plant, and from 2012 to 2022, he was dedicated to the design, analysis and management of PV power stations, and solar parks, as the Head of the Center for Energy Solutions. His research interests include the problems of distributed green generations, mainly on power and energy balance, voltage management, and their electrical stability with widespread application of communication technologies. He passed away in 2022.

TATYANA DEDOVA received the B.S. degree in applied mathematics from Kazakh National University, Kazakhstan, in 1996, the M.S. degree from Kazakh State Academy of Management, in 1998, and the Ph.D. degree in geoinformation Modeling from the Space Research Institute, Kazakhstan, in 2002. She was with the Institute of Space Research, from 1996 to 2004, and KazGeoCosmos JSC, from 2004 to 2018. Since 2015, she has been with EcoRisk LLP, as the Scientific Project Leader. She has been involved in projects related to environmental monitoring, risk mapping, geoinformation of weather prediction modeling, and GIS development, including aerial photography and the creation of digital topographic plans.