DOI: 10.58254/viti.4.2023.06.65

UDK 621.396.4

Doctor of Tech. Sciences Romaniuk Valery ORCID: 0000-0002-6218-2327 (MITIT)
PhD Hrymud Andrii ORCID: 0000-0003-4012-5185 (MITIT)

ANALYSIS OF THE RULES FOR CONSTRUCTING A FLIGHT TRAJECTORY OF A COMMUNICATION AERIAL PLATFORM FOR DATA COLLECTION FROM NODES OF A WIRELESS SENSOR NETWORK

Романюк В. А., Гримуд А. Г. Аналіз правил побудови траєкторії польоту для збору даних комунікаційною аероплатформою з вузлів безпроводової сенсорної мережі.

Для збору даних моніторингу з вузлів безпроводових сенсорних мереж із незв’язною топологією пропонується використовувати комунікаційну аероплатформу (далі – КА), яка побудована на базі безпілотного літального апарата. КА в процесі польоту формує тимчасові локальні радіомережі та виконує роль повітряного шлюзу для збору даних з окремих вузлів та головних вузлів у зв’язних фрагментах мережі. Ефективність процесу збору даних залежить від характеристик мережі, способів (правил) побудови траєкторії польоту КА, кількості та локації точок (інтервалів) обміну даними тощо. У статті проводиться оцінка ефективності застосування різних продукційних правил побудови траєкторії польоту КА для збору даних з вузлів мережі для досягнення певних цільових функцій: мінімізації часу збору даних, максимізації часу функціонування мережі.

У роботі сформульована задача пошуку траєкторії польоту та збору даних КА з вузлів як задача пошуку найкоротшої траєкторії переміщення зони покриття КА з початкової в кінцеву точки польоту, яка забезпечує покриття всіх вузлів (кластерів на площі) на мінімальній відстані обміну КА з вузлами. Для оцінки ефективності застосування правил побудови траєкторії польоту та збору даних розроблена відповідна імітаційна модель. Вхідними даними моделювання є характеристики мережі, вузлів та комунікаційної аероплатформи, способи (правила) управління процесом збору даних. Імітаційна модель надає можливість отримувати залежності показників ефективності (час збору даних, витрати енергії батарей, час функціонування мережі) на множині правил побудови траєкторії польоту та збору даних КА при різних вхідних даних.

Результати імітаційного моделювання застосування множини (бази) правил довели можливість зменшення часу збору даних до 20 % або підвищення часу функціонування мережі до 15 % порівняно з існуючими рішеннями.

Ключові слова: *безпроводова сенсорна мережа, комунікаційна аероплатформа, траєкторія польоту, збір даних, база правил, моделювання.*

V. Romaniuk, A. Hrymud Analysis of the rules for constructing a flight trajectory of a communication aerial platform for data collection from nodes of a wireless sensor network.

To collect monitoring data from the nodes of wireless sensor networks with a disconnected topology, it is proposed to use a communication aerial platform (CA), which is built on the basis of an unmanned aerial vehicle. During the flight, the CA forms temporary local radio networks and performs the role of an air gateway for collecting data from individual nodes and main nodes in connected network fragments. The efficiency of the data collection process depends on the characteristics of the network, the methods (rules) of constructing the trajectory of the CA, the number and location of points (intervals) of data exchange, etc. The article assesses the effectiveness of the application of various production rules for building the trajectory of a CA to collect data from network nodes to achieve certain target functions: minimizing the time of data collection, maximizing the time of network operation.

The paper formulates the task of finding a flight path and collecting CA data from nodes as the task of finding the shortest trajectory of the movement of the CA coverage area from the initial to the end point of the flight, which provides coverage of all nodes (clusters on the area) at the minimum distance of the CA exchange with the nodes. In order to evaluate the effectiveness of the application of the flight path construction rules and data collection, a corresponding simulation model was developed. The input data of the simulation are the characteristics of the network, nodes and communication aerial platform, methods (rules) of managing the data collection process. The simulation model provides an opportunity to obtain dependences of efficiency indicators (data collection time, battery energy consumption, network operation time) on a set of flight path construction rules and CA data collection with different input data.

The results of simulated modeling of the application of a set (base) of rules proved the possibility of reducing the time of data collection by up to 20% or increasing the time of network operation by up to 15% compared to existing solutions.

Keywords: *wireless sensor network, communication aerial platform, flight path, data collection, rule base, modeling.*

1. Problem statement. In recent years, the technology of wireless sensor networks (WSNs) has been rapidly developing as the transport basis of the Internet of Things to solve various problems for both civilian and military purposes [1–3]:

monitoring the parameters of objects (territories) for pollution (radiation, chemical, biological, etc.) and compliance with environmental standards;

monitoring the state of state borders, crop fields, forests, rivers, seas, food pipelines, power lines, bridges, etc;

assistance in search and rescue missions;

seismic monitoring, tracking the movement of wild animals, etc;

control of military operations and the battlefield (monitoring the status of their own and enemy troops, obtaining coordinates of movement of military equipment and personnel), etc.

The peculiarity of WSNs [1–3] is the limited resources of sensor nodes in terms of battery power, processor speed, memory capacity, transmitter power, etc. Modern WSNs can have hundreds or thousands of sensor nodes. The classical WSN architecture is a self-organizing radio network that transmits monitoring data along built-in routes from nodes to a radio gateway, which then transmits data to users via the communication infrastructure.

Certain areas of application of WSNs should consider the peculiarities of the nodes' location and the possibility of obtaining an unconnected network topology. For example, in remote areas affected by an emergency, with difficult terrain, occupied by the enemy, there is no possibility of using public communication infrastructure to organize communication with the gateway and it is impossible to build a connect topology of the entire network in conditions of a significant distance between nodes.

In these conditions, it is proposed to use a communication airborne platform (hereinafter referred to as the CA) as a mobile air gateway to collect data from network nodes [1–4]. In contrast to the classical architecture of data collection in sensor networks (an ad-hoc radio network with data collection along transmission routes between nodes to the gateway), this allows first, to organize the collection of monitoring data from unconnected network nodes; second, to obtain line-of-sight radio channels and shorter range, which reduces the level of energy consumption of nodes for data transmission (respectively, increases the network lifetime), etc.

This raises the urgent scientific task of analyzing the methods (rules) for constructing a trajectory for the overflight of the network nodes by the CA to collect data to achieve certain objective functions (hereinafter referred to as the "OF"): minimizing the time of data collection and/or minimizing the energy consumption of nodes (maximizing the network operation time) [3–5]. Solving this problem will allow optimizing the parameters of the data collection management system and improving its efficiency.

2. Analysis of recent publications. The use of CA for data collection in remote high-dimensional WSNs is a well-known solution and has been considered from different angles in many publications [1–16].

The first group of publications [5; 6] considers two ways of solving the problem of constructing a CA flight path:

flying over the entire territory of nodes [5] according to different models – horizontally, spiral, zigzag, Hilbert curve, etc.

most of the territory (number of nodes), considering the limitation on the maximum flight time of the CA (tractor, angular, circular, square flight models of the CA are proposed [6]).

The main disadvantage of the method of collecting data by flying over the territory is the significant time of data collection, but this method will be used during the initial flight of the network of CA to determine the actual parameters of the nodes (position coordinates, battery power level, amount of monitoring data).

The next group of publications [7–10] considers the construction of a CA flight path when collecting data from nodes only as a solution to the classical traveling salesman problem – finding the shortest route between the starting and ending points of the CA flight with a flight over nodes or data collection points (cluster centers). This problem belongs to the class of NP-hard problems. Obtaining an exact solution for a network of significant dimensionality is problematic. Therefore, heuristic algorithms for obtaining an approximate solution are proposed in practice, which have a low computational complexity: Lin-Kernighan traveling salesman heuristic (LKH) [8], Near Neighboring (NN) [9], Spiral Decomposition (SD) [10], Fast Path Planning with Rules (FPPWR) [11], Convex Hull Insertion Heuristic (CHIH) [12], ant algorithm [13], genetic algorithm [14], particle swarm algorithm [15], etc. However, such a problem statement calculates only the shortest flight route but does not consider the state of the nodes' parameters, nor does it optimize the energy consumption of the nodes. Therefore, the application of the shortest path search algorithm can be used for the initial (basic) solution for its further improvement.

The third group of publications defines [16–21] two main approaches to improve the efficiency of CA data collection from nodes of large-dimensional WSNs:

1. Collecting CA data directly from each node by creating virtual clusters at the actual location of the nodes.
2. Data collection only from the head nodes of real network clusters.

With the first approach (in the absence of connectivity between nodes), the network control center calculates temporary clusters (local radio networks: CA – nodes of the cluster) and forms them during the flight by the CA (air gateway). In the second case, if there is radio communication between certain sensor nodes, the network is self-organized and divided into real clusters with the identification of the main cluster nodes (according to well-known clustering algorithms LEACH, HEED, etc.) [16], which collect data from simple monitoring nodes. The CA flies around and collects data only from the main nodes of the clusters. In the following, we will take both approaches into account.

Publications [17–21] investigate the peculiarities of individual stages of the data collection process and ways to achieve the objective functions: minimum collection time, maximum area coverage (number of nodes), etc.

In [17], the problem of reducing the data collection time by sequentially adding potentially possible hang points is investigated. However, the search for variants of hang points leads to a significant computational complexity, so the proposed solutions can be used in a small-dimensional network.

In [18], several strategies for constructing data collection points from cluster nodes and the CA flight path in a temporary clustered WSN are investigated: through the center of the cluster, data collection on the flight path at the closest node-CA distance, flight through critical nodes in clusters, flight with hovering at one collection point that minimizes the total energy consumption of nodes, etc. However, the authors do not consider the possibility of building multiple data collection points in a cluster, optimizing exchange intervals, optimizing multiple objective functions.

In [19], rules (heuristics) for shortening the CA trajectory are considered by considering the direction of movement of the CA and the location of nodes in the cluster. Further improvement of these rules and evaluation of their effectiveness will be presented later in the article.

In [20], a deep neural network is used to find the 3D flight path of a CA considering the quality of the radio channel, but the set of performance indicators of the collection process is not considered.

In [21], it is proposed to achieve the objective function sequentially by hierarchy levels: network, cluster, CA, group of nodes, and individual node. At the network level, performance indicators are optimized by determining the number of clusters and their sizes, and building the shortest flight path. At the levels of a CA group of nodes and a CA node, the distance is determined, which allows optimizing the time of exchange between them and energy consumption (rules for adjusting the points (intervals) of trajectory data collection are proposed). To reduce the number of options and reduce the time for finding a solution along the trajectory of overflighting nodes and collecting data, an appropriate database of decision-making rules is proposed. However, no evaluation of the effectiveness of the rules and the weighting of their priority is given.

Thus, the unresolved task when considering the rules for constructing a CA flight path is to evaluate the effectiveness of their application and determine the order (priority) of application for constructing a data collection path for a particular WSN to achieve two main optimization criteria: minimizing the time of data acquisition and maximizing the network operation time.

The aim of the article is to analyze the effectiveness of applying the basic rules for building and correcting the flight path and collecting data from the nodes of the BSM by a communication air platform when achieving certain target functions.

Presenting main material. We consider a stationary wireless sensor network of considerable dimensionality (hundreds of sensor nodes) for special purposes. Each sensor node consists of the following basic elements: a battery, a certain set of sensors (e.g., vibration, magnetic, acoustic, etc.), a processor, memory, transceiver, antenna, positioning system, and control system.

During its operation, each sensor node collects and stores environmental parameters (objects of observation) of the monitoring zone assigned to it. The number of collection parameters is determined by the type of sensor modules, and the frequency and methods of data collection (by events, periodically, continuously) are determined by the ground-based network control center (NCC).

The WSN nodes are randomly located in a certain area and are unable to establish a unconnected topology for data transmission to the gateway for various reasons: a significant distance between them, specific terrain, lack of any public communication infrastructure, the need to periodically perform radio silence, economic inexpediency of installing and operating gateways, etc. In other words, under these conditions, the topology of a sensor network will consist of separate unconnected nodes and/or separate unconnected network fragments (clusters). Nodes in connected clusters (if they have an appropriate control system capable of implementing self-organization algorithms) can introduce a control hierarchy: form the head nodes of clusters that will collect and store data from other so-called simple cluster nodes [16].

For WSNs with an unconnected topology, the role of a mobile gateway is played by an unmanned aerial vehicle (UAV) equipped with additional equipment to implement the process of collecting data from sensor nodes: processor, memory, transceiver, antenna, positioning system, and the corresponding control system.

At the planning stage, the NCC calculates the CA flight path and determines the preliminary points (segments) for collecting data from nodes in space. During the CA flight, thanks to a directional antenna at an altitude of $h_k(t)$ forms temporary clusters (coverage and radio communication areas) $C_k(t)$, $k = 1 \ldots K$ nodes with a radius of $R_k(t)$, that is, it creates temporary local radio networks with an air access point (CA). If during the flight of the CA some nodes (head nodes of real clusters of network fragments) fall into the current radio communication zone, then it: establishes radio communication with them (according to the MAC protocol), determines the exchange schedule and determines (or corrects) the point (interval) of the exchange trajectory. When the CA approaches the point (interval) of data collection on the trajectory, the node-CA data exchange process takes place (Fig. 1).

Given:

1. Characteristics of the wireless sensor network:

– the area of the WSN (S) and the type of its geometric shape (for example, rectangular, strip, circle, arbitrary, etc.)

– the number of network nodes (unconnected and/or head nodes of real clusters), coordinates of their location on the ground (x_i, y_i), $i = 1 \ldots N$;

– density of node placement $\alpha = \bigcup_{k=1}^{K} \pi R_k^2 \,/S$ nodes (where R_k – radius k-th coverage area of the CA, $k = 1 \ldots K$) and the type of their placement (homogeneous, grouped, etc.);

– the amount of monitoring data collected by each i-th node – V_{dmi}.

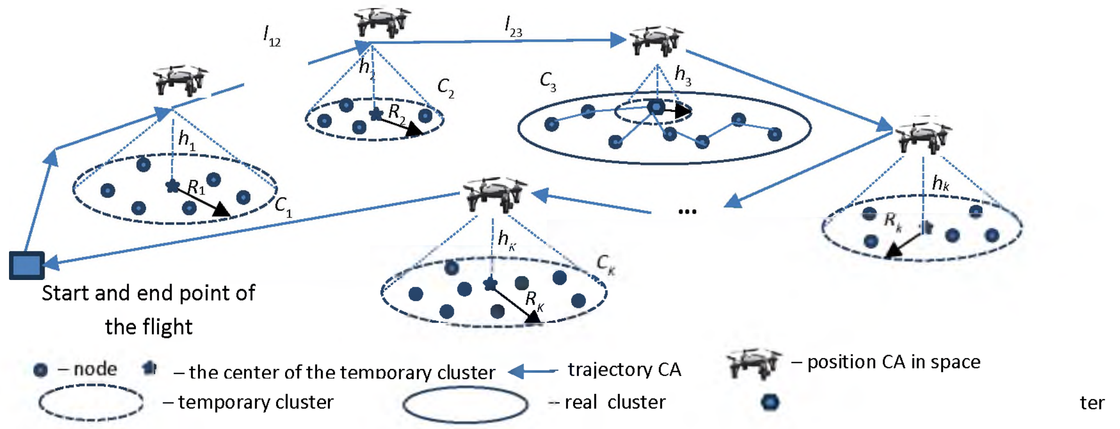


Fig. 1. An example of a CA overflight of clusters and data collection from cluster nodes

2. Characteristics of the node:

– technical characteristics of the node – number and types of sensors, battery power level, energy consumption for monitoring environmental parameters for each type of sensor, etc.;

– communication characteristics of the node – parameters of the antenna, transmitter, energy consumption per bit of reception and transmission for the selected MAC protocol and type of equipment, etc.

3. Characteristics of the communication aeroplatform:

– flight characteristics – speed, altitude, flight time, battery energy, ability to hover and move in space at a constant or variable speed, etc.;

– communication characteristics – MAC protocol, antenna and transceiver parameters, etc.

4. Objective control functions (1)–(4) of data collection implemented by the NCC [21]:

– minimize data collection time T_{col}

$$T_{col} = \frac{L}{v} = \sum_{m=1}^{M} \frac{l_m}{v_m} \to min \tag{1}$$

ensuring a given network operation time $T_{fun} \geq T_{fungiv}$,

$l_m, v_m, m = 1 \dots M$ – trajectory intervals between data collection points and flight speed;

– maximize network operation time T_{fun} by reducing (redistributing) the energy consumption of nodes e_{coni}

$$T_{fun} \to max \quad \left(E_{con} = \sum_{i=1}^{N} e_{coni} \to min\right) \tag{2}$$

ensuring the specified data collection time$T_{col} \leq T_{co\,lg\,iv}$;

– optimization of both criteria
$$\begin{cases} T_{col} \to min \\ T_{fun} \to max \end{cases}; \tag{3}$$

– obtaining an acceptable solution

$$T_{col} \leq T_{colgiv} \text{ and/or } T_{fun} \geq T_{fungiv}, \tag{4}$$

with restrictions Ω:

– type of CA (rotary); speed $v=[v_{\min}, v_{\max}]$, altitude $h=[h_{\min}, h_{\max}]$, time $t_{fly} \leq t_{fly\max}$ and the range of its flight $L \leq L_{\max}$;

– number of clusters in the network – $1 \leq k \leq K$;

–initial energy of the node batteries $e_i \leq e_{\max}$ and KA $e_{KA} \leq e_{KA\max}$;

– the amount of monitoring data for each i-th node – $V_{dmi} \leq V_{dm\max}$;

– radio communication range between node and CA – $d_{i\text{-TA}} \leq d_{\max}$ – the propagation of radio waves in line-of-sight conditions is considered;

– radius of the CA coverage area (cluster) – $R_{\min} \leq R \leq R_{\max}$.

Network operation time T_{fun} can be determined by the following indicators:

a) The period of stable operation of the network T_{pso} (5) – network operation time (monitoring and data transmission in each round of the CA flight) until the first node fails due to battery depletion

$$T_{pso} = \min_{i\in N} t_{funi}(Nround), \tag{5}$$

where t_{funi} – is the time of operation of the *i*-th node before its failure, which is determined by the number of rounds of circling (*Nround*).

б) Percentage of nodes that failed (6) relative to the number of CA rounds

$$T_{fun} = \frac{N_{fail}(Nround)}{N} \text{as a percentage,} \tag{6}$$

where N_{fail} $(Nround)$ – the number of nodes where the battery energy is less than the permissible level on the *Nround* round.

5. A set of methods (rules) for constructing a flight path and collecting CA data from the WSN nodes.

Restrictions and requirements:

– the CA flight area has no prohibited zones, its flight trajectory is formed in the form of certain coordinates of points in space, modeling of the CA flight process is not considered;

– information on node status parameters (location coordinates, battery energy level, amount of monitoring data) is collected during the initial overflight of the CA network, and then the information on the node status is updated during each round of overflight;

– the CA has the ability to collect data both during hovering and during flight;

– The CA and sensor nodes have radio facilities with the same MAC protocol, which allows to change the data transmission rate depending on the state of the radio channel (signal-to-noise ratio) and to regulate the transmission power (energy consumption for transmission), for example, IEEE 802.11;

– memory capacity of sensor nodes and CA is sufficient to store monitoring data;

– the energy value of the CA battery is sufficient for a round of network overflight;

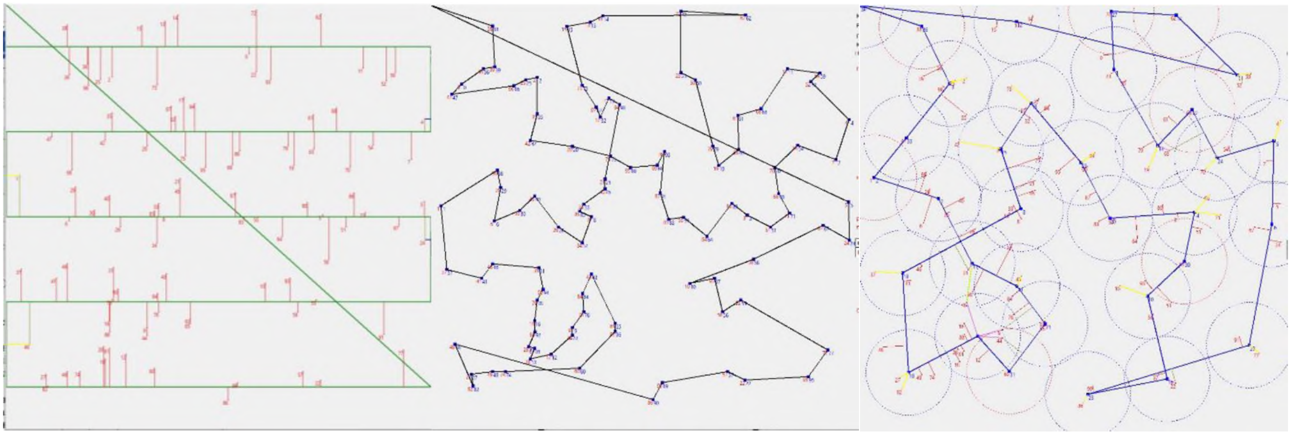
– algorithms for controlling the data collection process, which are implemented by the control systems of nodes and CA, should have a low computational complexity due to the need to implement autonomous flight and ensure the data collection process in real time.

It is necessary to: to analyze the effectiveness of applying various rules for constructing and correcting the CA flight path to collect data from the WSN nodes while achieving certain objective functions.

Solution.

The set of solutions for achieving the defined objective functions (1)–(4) lies between two limiting methods: flying over the entire area of the WSN and flying over each node of the network.

1. CA overflight of the entire territory (area) occupied by the WSN nodes, with simultaneous collection of monitoring data from network nodes. Thus, studies [5; 6] analyzed various options for flying over the entire area and collecting data from the main nodes of real clusters: by lanes (Fig. 2, *a*), by angle, by square, by circle. The aim of the study is to find overflight options that allow to reduce the length of the overflight route or maximize the number of serviced (covered) nodes for a limited time of the CA flight. It is shown that there is no single optimal overflight option: the "strip" option is effective for maximizing the coverage area of the CA, the "circle" option is more efficient in terms of overflight time. However, the length of the route and the flying time over the entire area of the CA remains very long. For example, according to the results of the authors' simulation modeling for a network of 100 nodes, the length of the CA trajectory when flying over the entire area horizontally at $R = 100$ is $L = 7500$ at the total energy consumption of the nodes $E_{con} = 4477$. When the coverage radius is reduced by half $R = 50$ the length of the trajectory is already $L = 11878$ with a halving of energy consumption to the value of $E_{con} = 2115$ (Fig. 2, *a*), which imposes additional requirements on the flight characteristics of the CA. The overflight of the entire network will usually be used during the initial overflight of the network to collect initial information about the parameters of the network nodes (position coordinates, amount of monitoring data, battery energy level, etc.).



$L = 7232 \quad E_{con} = 4477\ (R = 100)$ $L = 8215\ E_{con} = 2000\ (R = R_{\min})$ $L = 5987\ E_{con} = 3068\ (R = 100)$

$L = 11878 \quad E_{con} = 2582\ (R = 50)$ $L = 7152\ E_{con} = 2114\ (R = 50)$

a) ***b)*** ***c)***

Fig. 2. Ways to fly CA around the WSN nodes:
a – the entire area of the WSN ("by lanes"); *b* – of each node; *c* – temporary clustering

2. Flying around each node at the minimum flight altitude of the CA. As a result of the simulation of this method, we obtain the minimum energy consumption of nodes for data exchange $E_{con} = 2000$ (due to $R \leq R_{min}$) and a significant length of the route $L = 8215$ (Fig. 2, *b*). It is advisable to use this method to minimize energy consumption – objective function (2), but it is not guaranteed that the flight time (data collection) limit will be met $L \leq L_{max}$. Different types of algorithms can be used to solve the problem of finding the shortest route to fly around each of the nodes: full search (for WSNs of small dimensionality), heuristic (NN – Fig. 2, *b*, FPPWR, CHIH, etc.), genetic, etc. Each of them shows different results depending on the network parameters and implementation features (see below in Section 3).

The results of modeling the performance evaluation indicators of the first two methods are borderline and are used for comparison with the results of applying other methods.

3. Between the above two limiting ways of flying over the CA (the entire area and each node) is a hybrid variant of flying over (**temporary clustering of the network**) – **flying over only certain zones (clusters) of the WSN area** (for example, through the centers of the clusters, Fig. 2, *c*), **in which groups of nodes or individual nodes are actually located**.

That is, the classical salesman problem is transformed into **the problem of finding the shortest route (or rather trajectory) to move the CA coverage area from the starting to the end point of the flight, which ensures coverage of all nodes (points on the area) at the minimum distance of exchange between the CA and the nodes**.

Achievement of the objective functions (1)–(2) has the opposite dependence. Reducing the length of the CA trajectory leads to an increase in the distance between the CA and the nodes and, accordingly, an increase in the energy consumption of the nodes for data transmission and vice versa. Therefore, it is proposed to use the lexicographic method to solve the two-criteria optimization problem (3). To do this, before the CA flight, the ground-based NCC determines the priority of the objective functions in accordance with the current operational situation. It is proposed to search for the optimal solution according to the hierarchy according to the priority of the objective functions in the following steps [21]:

virtual clustering of the network is carried out according to a certain cluster analysis algorithm - it is proposed to use the iterative cluster analysis algorithm FOREL (FORmal Element), which has certain advantages over others – low computational complexity, correspondence of the cluster concept to the physical content of the CA coverage area, etc;

the shortest (or permissible length) route for flying around the data collection points is built according to a certain algorithm for solving the salesman problem;

the data collection points (intervals) of CA monitoring in clusters are determined according to the adopted cluster overflight strategy and the priority of the objective functions (in the simplest case, the center of the cluster is determined as the data collection point).

It should be noted that the effectiveness of the final optimization solution and the speed of its obtaining (computational complexity) is largely determined by the initial clustering solution (depending on the choice of the starting point of the area and radius of coverage of the satellite), the algorithm for finding the shortest flight path, the strategy of flying around the nodes in the cluster, etc.

So data collection time T_{col} of CA from the network nodes depends on the following parameters (5):

$$T_{col} = f(N, K, TR(St_k), v, V_{dmi}, Q_k, INT_i, s_{i-\text{КА}}(d_{i-\text{ТА}}, MAC), h_k, \Omega) \qquad (5)$$

– number of nodes i =1… N and the coordinates of their location on the ground (x_i, y_i);

– quantities k = 1…K clusters, their area, the number of nodes in a cluster n_k, and their relative position;

– trajectories TR of the CAs flight in the network defined by the strategy St_k overflight and data collection from the nodes of each k-th cluster (data collection during the flight and/or when the CA is hovering, one or more hovering points in the cluster, etc;)

– speeds $v = \{v_m\}$ CA flight on each of the trajectory segments $l_m, m = 1 \dots M$;

– the amount of monitoring data V_{dmi} in network nodes;

– number of data collection points Q_k with coordinates in space $(x,y,h)_k$ in each k-th cluster when the CA hangs up;

– locations in the space and time of intervals $INT_i = \{(x, y, h)_{begin}, (x, y, h)_{finish}, t_{begin}, t_{finish})\}_i$ CA flight trajectories, which are defined for data collection (exchange) in motion with the i-th node;

– transmission speed of the MAC protocol $s_{i-CA}(d_{i-CA}, MAC)$, which depends on the distance d_{i-CA} and parameters of the radio channel (signal-to-noise ratio), transmitter, receiver, antennas, etc;

– flight altitude h_k, limitations Ω resources of nodes and CA, etc.

Increase in network operation time T_{fun} can be achieved:

reducing the energy consumption of nodes for data reception and transmission (reducing the transmission power of the node) by reducing the distance between the CA and the node

$$d_{i-TA} = g(K(R_k), n_k, TR_k, Q_k, INT_i),$$

which is achieved by optimizing the number of clusters K (the size of the coverage area R_k), is the number of nodes in the k-th cluster n_k, trajectories TR_k, position of the points Q_k (intervals INT_i) exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster exchange in the cluster;

redistribution of energy consumption between nodes competing for transmission (if a node has a higher battery energy level, then it must consume more energy).

In addition, when determining the trajectory of the cluster nodes and data exchange, it is necessary to take into account:

– relative location of nodes relative to the trajectory (it is desirable to exchange data in the nearest intervals of the CA flight trajectory from the node);

– critical energy level of the node's battery (plan to fly over "exhausted" nodes at a minimum distance);

– the amount of node monitoring data – selecting the point (interval) of trajectory collection that is closer to this node.

A simulation model was developed to evaluate the effectiveness of the proposed different methods (algorithms, rules) of collecting data from WSN nodes by CA. It is written in Python, has a user-friendly interactive interface, and allows visual tracking of all stages of the process of collecting CA data from network nodes.

The input data for the modeling are as follows.

1. Characteristics of the network, nodes, CA (according to the problem statement): S=1000×1000 conventional units; N = 50…200; node placement (homogeneous, with grouping), the initial energy of the nodes and the amount of monitoring data e_i; energy consumption of the node for transmission (calculated by a simplified formula $e_{con} = c * d_{i-\text{KA}}^2$, where c =const, d – distance between the node and the CA); coverage radius R; radio communication range between node and CA – $d_{i\text{-KA}}$; number of overflight rounds *Nround* etc.

2. Methods (from the whole area, from each node, by clusters), methods (directly from each node, from the head nodes of clusters) of data collection.

3. Algorithms for temporary network clustering (FOREL, *k*-averages, etc.).

4. Algorithms for finding the shortest route: heuristic (nearest neighbor, FPPWR, CHIH, etc.).

5. Clustering rules, determination of data collection points (intervals) on the CA flight path rules, cluster overflight rules, rules for building the CA path, etc. (Table 1) [21].

Table 1

Decision-making hierarchy for collecting CA data from the WSN nodes

The stage decision	Managed parameter	Action (rule)	
		Obective function minT_{col}	Objective function min E_{con}
1. Virtual clustering of the WSN (network level)	Number and size of homogeneous clusters R^*, starting point of clustering, number of nodes in clusters, size of each cluster R_k	Reduce the number of clusters: increase R, determine the starting point of clustering in the places of grouping nodes, redistribution of nodes between clusters	Increase the number of clusters: reduce R while maintaining connectivity between cluster nodes and CA
2. Search for the initial flight trajectory of the CA (network level)	An algorithm for finding the shortest route (trajectory) to fly around the centers of clusters (collection points) of the WSN	Selection of the best heuristic algorithm from the set, optimization of the algorithms' own parameters	–
3. Adjustment of data collection points in each cluster (cluster level)	Position of the data collection point (interval) relative to the CA trajectory and location of the cluster nodes	Rules for trajectory shortening in a cluster when ensuring radio communication CA - the farthest node	–
4. Determination of the data collection strategy in the cluster (cluster level)	Location of the trajectory relative to the position of the nodes in the cluster, number and position of data collection points	Rules for reducing the distance of a CA node that has a significant amount of data (reducing the exchange time)	Rules for reducing the node-CA distance for battery power consumption
5. Correction of collection points (intervals), calculation of the CA-node exchange schedule (CA node level)	Number of points (length of intervals) of data collection on the CA trajectory segments, CA flight speed	Redistribution of collection points along the CA trajectory segments, selection of the maximum flight speed while meeting the exchange time	Rules for the priority of exchange of node-CA with a lower (low) battery energy level

The simulation model provides an opportunity:

to obtain dependencies of performance indicators (data collection time (trajectory length), battery power consumption, network operation time) on a set of control parameters (rules) for building a flight path and collecting CA data under different input data (network dimension, type of node location, number of overflight rounds, etc;)

explore the optimization parameters – the number and size of clusters, the number and location of data collection points (intervals) on the flight path, cluster flying strategies; algorithms for finding the shortest route, etc.

Let's model and evaluate the effectiveness of each stage of solving the problem of collecting CA data from the WSN nodes and the corresponding rules for their implementation

1. Finding a solution for network clustering – optimization of CA coverage areas.

Reducing the number of clusters k in the network potentially reduces the length of the CA flight route (and thus the data collection time), but leads to an increase in sensor node energy consumption (due to the increase in the node-CA distance) and an increase in node-CA communication time (decreases the MAC protocol transmission rate). And vice versa: an increase in the number of clusters leads to an increase in the length of the CA flight route, but reduces the node-CA distance and, accordingly, reduces their exchange time and node energy consumption. In other words, there is a certain optimum of the number of clusters k^*, their size R^*, and the location of points in them Q_k and intervals INT_i CA data collection, which determines a compromise decision regarding the TF. Therefore, the following basic rules apply [21].

The rule for determining the number of clusters in the network: IF the priority of the objective function T_{col}→min (T_{fun}→max), THEN increase (decrease) the size and number of clusters.

The results of the study of the proposed rule when applying various heuristic algorithms for finding the shortest path (nearest neighbor, FPPWR, CHIH) through the centers of clusters are shown in Fig. 3 and Fig. 4 (N = 100, nodes have the coordinates shown in Fig. 2). Cluster size is optimized R^* within the values of R = 50…100 for each of the algorithms. According to the FOREL clustering algorithm, the number of clusters is determined k = 69…33. We observe that for each shortest path search algorithm there is an optimal value of R^* ($R^*_{NN} = 94, R^*_{FPPWR}$=100, $R^*_{CHIH} = 100$), which allows to significantly (more than 20%) reduce the primary length of the CA flight path.

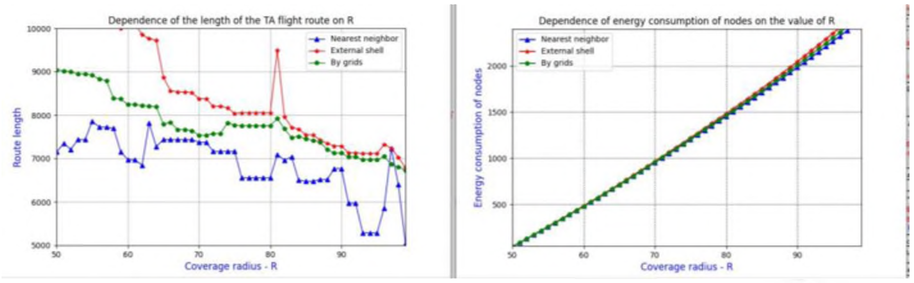


Fig. 3. Estimation of the trajectory length and energy consumption of nodes as a function of the coverage area size R with homogeneous node placement under different shortest route search algorithms

For a network with grouped nodes, the simulation results are shown in Fig. 4.

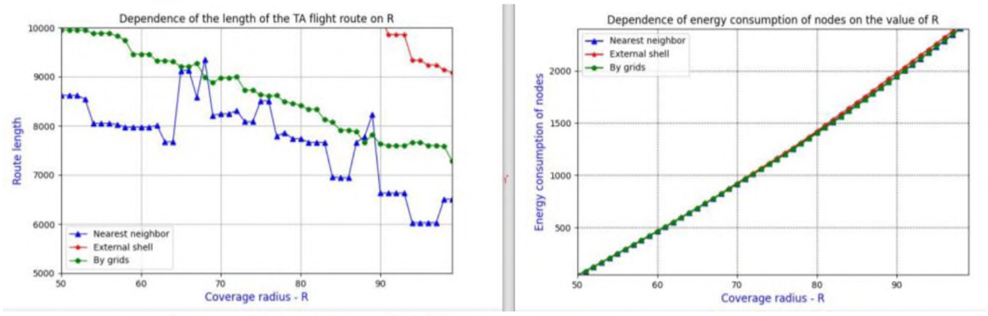


Fig. 4. Estimation of the trajectory length and energy consumption of nodes as a function of the coverage area size R when nodes are placed in groups

The application of other network rules (Table 1) affects the number of clusters and, accordingly, the length of the route, the energy consumption for the exchange of CA data with nodes.

The rule for selecting the initial clustering point: to select the initial clustering point of the network, select the point with the maximum number of nodes in the cluster (the goal is to reduce the number of clusters).

The rule for redistributing nodes between clusters: IF there is a small number of nodes in a cluster, then redistribute the nodes of this cluster to other clusters if possible (i.e. reduce the number of clusters).

The rule for adapting the size of each cluster: IF objective function (2) THEN reduce R by reducing the flight altitude of the CA while maintaining the connectivity of the CA nodes in the cluster.

The benefit of applying these rules can reach 5–7 %.

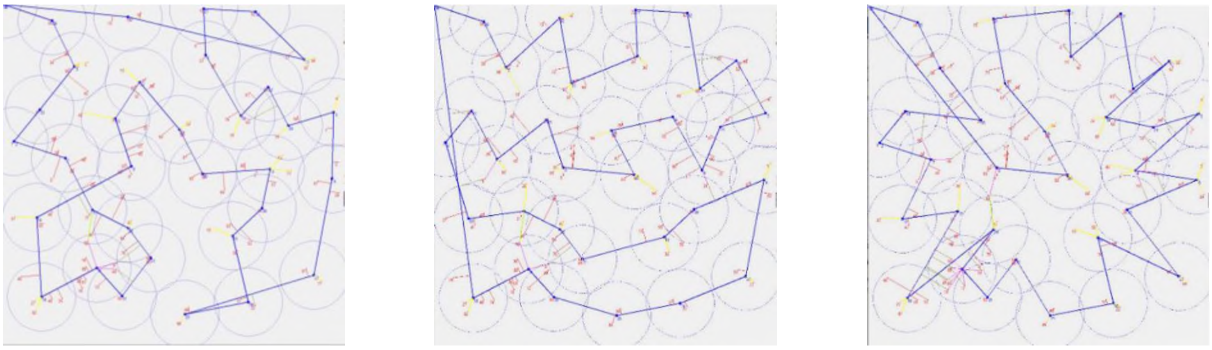
That is, optimization of the coverage area radius, application of the rules (selection of the initial clustering point at the planning stage, redistribution of nodes between clusters, adaptation of cluster sizes) has a significant impact on the efficiency of the data collection process during network operation..

2. Rules for finding the shortest route (trajectory) for flying over CA clusters.

To solve the problem of finding the shortest route to fly around each cluster, various known algorithms can also be used: full search (for a small number of clusters), heuristic, genetic, etc. Each of them shows different results depending on the parameters of the network, nodes, and CA. Figures 5–8 show the performance indicators (L, E_{con}, T_{fun}) when using three heuristic algorithms for finding the shortest path (NN, FPPWR, CHIH) with homogeneous and group placement of nodes, the size of the coverage area (R = 50, 100), network dimensions (N = 100).

The modeling results demonstrated a significant dependence of the performance indicators on the nature of the location of universities in the area (homogeneous, grouped) and the adopted algorithm for finding the shortest route to fly around the clusters. In our case, in most cases, the nearest neighbor algorithm has the advantage of the three defined algorithms with the best performance indicators.

In addition, for each of the algorithms, there is an additional possibility of optimization by the internal parameters of the shortest path search algorithms themselves. For example, the nearest neighbor algorithm can be used to select the number of steps to the next cluster (one, two, three, etc.), by squares (optimizing the size of a lattice square), convex hull (optimizing the size of each hull). At the same time, for specific network parameters (area, placement, node parameters, CA, etc.), each of them may have an advantage.

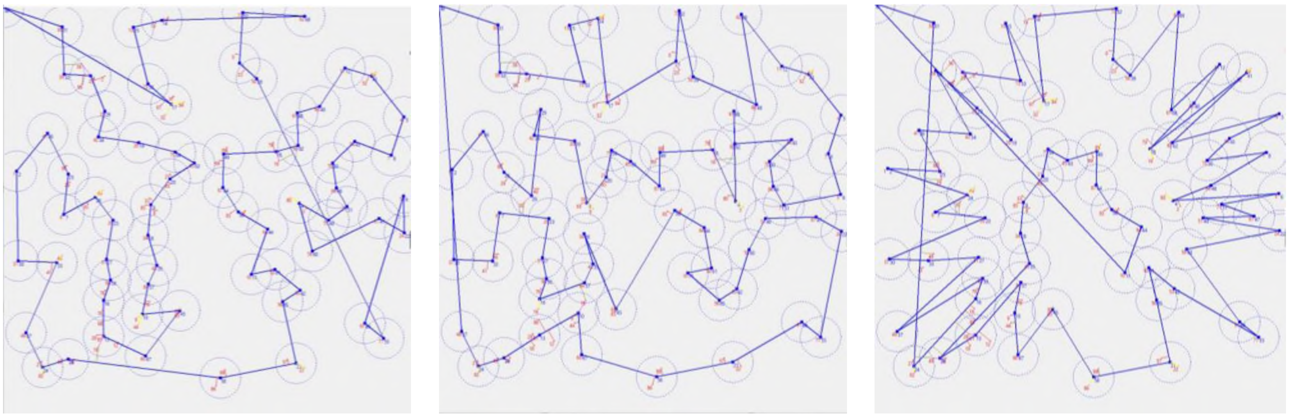


$L = 5987$ $E_{con} = 3068$ $(R = 100)$ *a)*

$L = 6596$ $E_{con} = 3157$ $(R = 100)$ *b)*

$L = 6899$ $E_{con} = 3290$ $(R = 100)$ *c)*

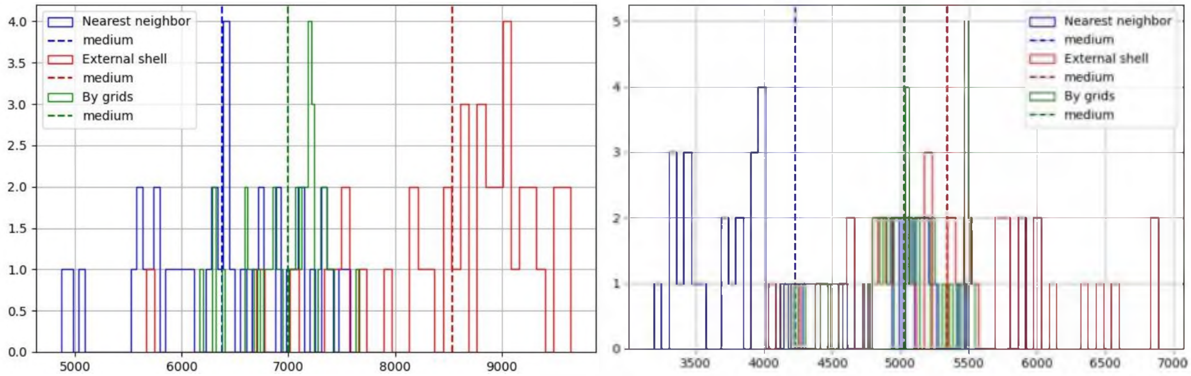
Fig. 5. Results of modeling the trajectory of a CA flight in a clustered network with uniformly distributed nodes in the WSN using different shortest route search algorithms (N=100, R=100): *a* – the nearest neighbor; *b* – FPPWR; *c* – CHIH with 2-layer shell



$L = 7153\ E_{con} = 2115\ (R = 50)$ $L = 9042\ E_{con} = 2128\ (R = 50)$ $L = 11844\ E_{con} = 2143\ (R = 50)$

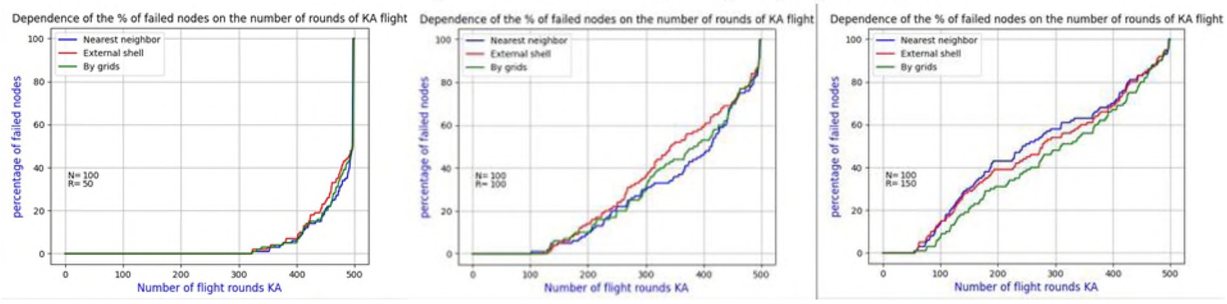
a) *b)* *c)*

Fig. 6. The result of modeling the trajectory of a CA over a clustered WSN by grouping nodes using different shortest route search algorithms (*N*=100, *R*=50):
a – the nearest neighbor; *b* – FPPWR; *c* – CHIH 2-layer shell



a) *b)*

Fig. 7. The result of modeling and estimating the length of the CA flight route of uniformly distributed and grouped nodes in the WSN using different algorithms for finding the shortest route
(*R*=100, a sample of 100 random locations of nodes on the area):
a – uniform placement of nodes; *b* – grouping of nodes

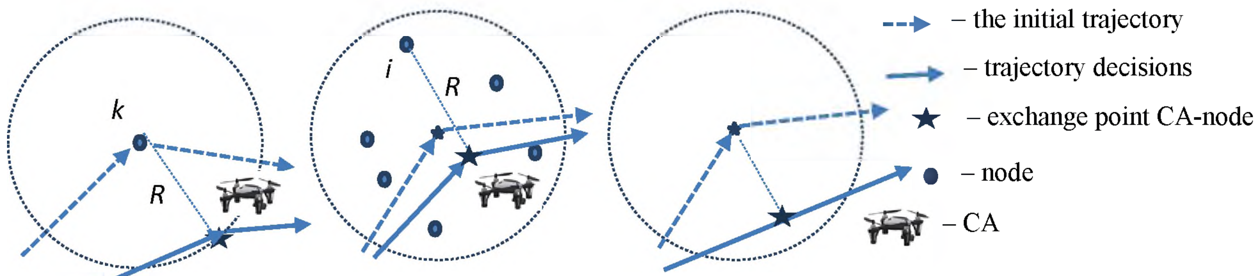


a) *R* = 50 ***b)*** *R* = 100 ***c)*** *R* = 150

Fig. 8. Dependence of network operation time on the number of CA rounds and coverage radius
(*N* = 100 at *R* = 50, 100, 150)

3. Rules for calculating (adjusting) data collection points in a cluster with a reduction in route length to improve the basic solution (through cluster centers).

The task of finding the shortest route for a CA flight differs from the classical traveling salesman problem. In our formulation of the problem, it is enough for a node to fall within the coverage area of the CA. Therefore, let us consider heuristic rules for shortening the route relative to the initial solution (see sections 1, 2) (Fig. 9) [19; 21].



a) node k is the only one in the cluster, the center of the next cluster does not belong to the CA movement sector

b) node i is located at a distance R from the exchange point

c) node k is the only one in the cluster, the center of the next cluster belongs to the sector of the CA movement direction

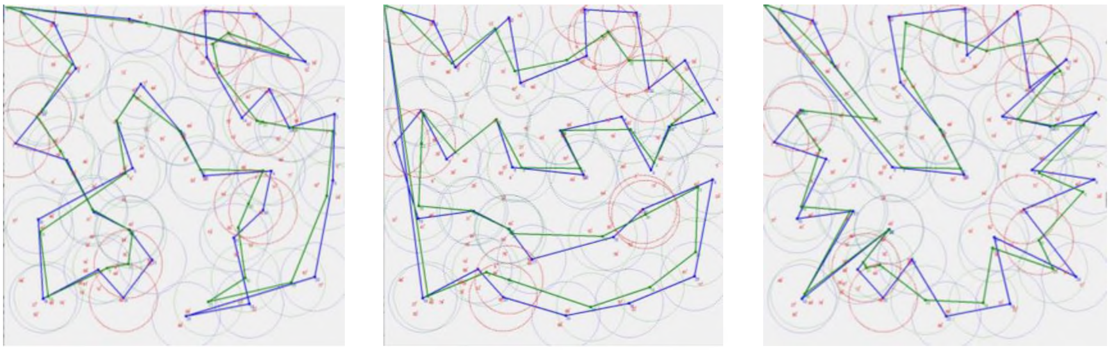
Fig. 9. Illustration of rules for reducing flight path length and CA data collection

Rules for reducing the length of the CA flight path in a cluster (Fig. 9).

IF the priority objective function T_{col}→min, the initial cluster flight route passes through the center of the cluster with one node k, and the next cluster center does not fall in the sector of the CA movement direction, THEN determine the point on the tangent line at a distance R from the new route in the CA movement direction as the data collection point (Fig. 9, *a*).

IF the priority objective function T_{col}→min, the initial route of the cluster overflight passes through the center of the cluster, the i-th node is at the greatest distance from the CA flight path, THEN determine the point that is at a distance R from the new route in the direction of the CA movement as the data collection point (Fig. 9, *b*).

IF the priority objective function T_{col}→min, the initial route of the cluster flight passes through the center of the cluster with one node k, and the next center of the cluster falls in the sector of the direction of movement of the CA, THEN the trajectory of movement is built directly to the next center of the cluster. determine the point on the trajectory closest to the center of the cluster as the data collection point (Fig. 9, *c*).



$L = 5987$ $E_{con} = 3068$
$L_{rule} = 4999$ $E_{conrule} = 3212$
a)

$L = 6596$ $E_{con} = 3157$
$L_{rule} = 5117$ $E_{conrule} = 3244$
b)

$L = 6899$ $E_{con} = 3290$
$L_{rule} = 5337$ $E_{conrule} = 3302$
c)

Fig. 10. Results of modeling the application of route shortening rules (N=100, R=100):
a – the nearest neighbor; *b* – FPPWR; *c* – CHIH 2-layer shell

Based on the results of the simulation of the rules for shortening the route length, we can conclude that the trajectory length is reduced by up to 20 %, while energy consumption increases by up to 10 %. According to the simulation results (Fig. 10), the heuristic with the nearest neighbor algorithm wins again, but with other network parameters, any algorithm for finding the shortest route can show the best performance.

4. Rules for selecting points (intervals) of CA data collection in clusters (overflight strategies).

At the cluster level, the number and coordinates of the points (intervals) of data collection of the CA (cluster overflight strategy) are determined. Possible variants of point selection rules are shown in Fig. 11 [18; 21]:

a – selection of exchange points closer to the CA trajectory (through the center or center of mass of the cluster);

b – data collection only when the CA hovers in the center (center of mass) of the cluster;

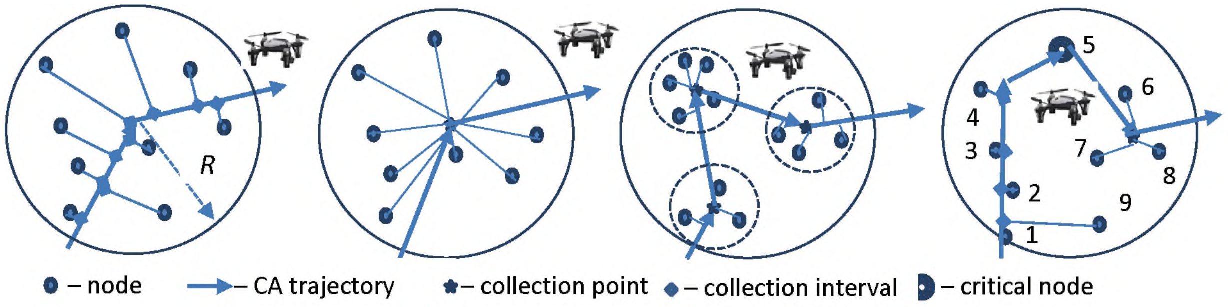
c – in-flight data collection with additional clustering if there are groups of nodes;

d – in-flight data collection taking into account the low battery energy level of individual nodes.

The result of each strategy in the k-th cluster is evaluated by a set of parameters:

– is the energy consumption of each cluster node for data collection (transmission and reception) e_{coni}, the total energy consumption of cluster nodes $E_{con}^{k} = \sum_{i\in k} e_{coni}$;

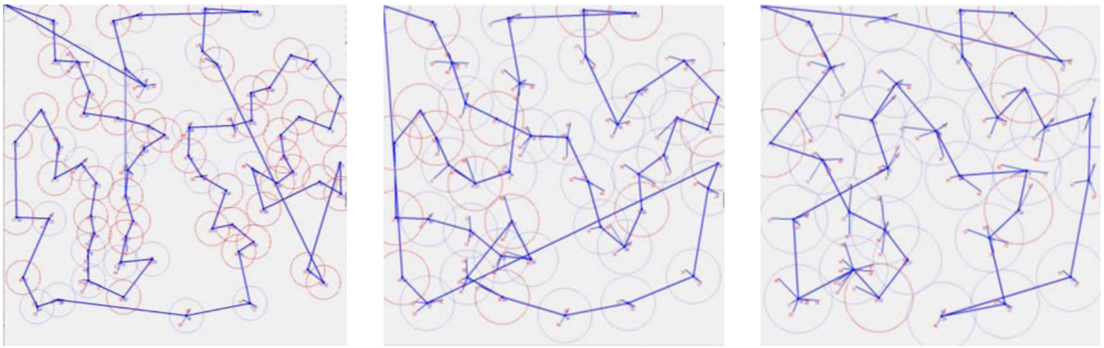
– time of data collection in the cluster t_{con}^{k}, which is determined by the flight time and hover time of the CA.



a) *b)* *c)* *d)*

Fig. 11. Basic rules for determining points (intervals) of CA data collection in a cluster:
a – in the flight; *b* – only in the center of the cluster; *c* – with additional clustering; *d* – through critical nodes

The simulation results showed a reduction in node energy consumption when using the in-flight data collection strategy (Fig. 11, *a*) E_{contr} compared to collecting data only in the center of the cluster E_{conz} (Fig. 11, *b*) – up to 20% reduction in node energy consumption for reception and transmission (Fig. 12).



$E_{contr} = 2115\ (R = 50)$ $E_{contr} = 2512\ (R = 80)$ $E_{contr} = 3841\ (R = 100)$

$E_{conz} = 2210\ (R = 50)$ $E_{conz} = 2890\ (R = 80)$ $E_{conz} = 3069\ (R = 100)$

Fig. 12. Results of modeling two strategies for flying around clusters
(along the trajectory and only in the center)

If there is a grouping of nodes in the cluster, it is advisable to carry out their additional clustering in order to reduce energy consumption and reduce the exchange time (with a significant amount of monitoring data in the nodes) – Fig. 11, *c*.

The rule for determining the hover points for data collection: IF there is a cluster of nodes in the cluster (loaded or with low battery energy), THEN determine the hover point for collecting data from these nodes in the cluster that minimizes the communication time or energy consumption of these nodes.

Rule for correcting the CA flight path: IF it is necessary to reduce the transmission time in the node-CA radio channel and/or reduce the energy consumption of the node, THEN it is necessary to place (move) the exchange points on the satellite path closer to the node.

Rule for reducing and equalizing transmission energy costs: IF multiple nodes compete for exchange intervals with the CA, THEN determine the closest exchange interval *INT* on the CA flight path of a node with a lower battery energy.

Rules for determining the number and location of hover points, exchange intervals, overflight and exchange strategies in the cluster:

IF objective function T_{col}→min, the number of nodes in the cluster is small (average) and the amount of data is insignificant, THEN determine the basic strategy (flight of the CA through the center of the cluster with the determination of exchange intervals on the flight path) (Fig. 11, *a*);

IF objective function TF T_{fun}→max, a large number of nodes in the cluster, their data volume is significant, THEN cluster the cluster with the definition of additional hang points (Fig. 11, *d*).

5. *Rule with redistribution of nodes in different intervals of the constructed CA trajectory.*

Each constructed CA flight trajectory consists of segments between collection points, which can be estimated by the length of the segment l_m and the number of exchange points. If there are trajectory intervals with a significant (very small) number of collection points, then we try to add (remove) nodes to (from) it and re-cluster without taking into account the added (removed) nodes. The goal is to shorten the length of the trajectory or reduce the energy consumption of the nodes.

Rule for redistribution of data collection points: IF a trajectory segment has a significant (small) number of collection points, THEN fix this trajectory segment, delete (redistribute) nodes according to the objective function and re-cluster the network to achieve a certain objective function.

The results of modeling the application of the proposed rule demonstrated the possibility of obtaining a gain in efficiency of up to 8 %.

The interval of exchange between a node and a CA is determined by the following considerations: the node's flight time should be at least as long as the exchange time [21].

Since each rule is focused on achieving a certain objective function, it has a different result of its achievement, therefore, their hierarchy in the form of meta-rules is proposed. For example [21].

Meta rule 1: IF objective function T_{col}→min, THEN (single-criteria optimization) to find:

– the maximum (specified) number of network clusters;

– set the basic trajectory of the CA through the center of the clusters, find the shortest path (use one algorithm from the set to find the shortest path) to fly around the centers;

– apply the rules for reducing the route length, determine possible additional collection points (hovering) of the CA in accordance with the position and amount of data in the cluster nodes (grouping of nodes with a significant amount of data);

– determine the strategy of flying around the cluster nodes;

– set the maximum speed of CA movement in the cluster that meets the requirements for data exchange between CA and cluster nodes;

– calculate the intervals and schedule of node transmissions during the CA flight, taking into account the state of the nodes, using rules that are focused on increasing the transmission rate in the radio channel.

Meta rule 2: IF objective function T_{col}→min and T_{fun}→max, the first objective function takes precedence over the second, THEN (lexicographic optimization method):

– find the maximum (defined) number of network clusters;

– determine the collection points (hovering) of the CA in accordance with the priority of the objective function;

– determine the strategy of flying around the cluster nodes;

– calculate the trajectory of the CA through the collection points;

– calculate the intervals and schedule of node transmissions during a flight at a minimum distance using rules that take into account the available energy of the nodes.

Evaluation of the effectiveness of the application of meta-rules showed the possibility of reducing up to 20% of the data collection time and increasing up to 15% of the network operation time compared to previously proposed solutions. It is important to note that the proposed situational management model [21] can be used in special software of the data collection management system of the NCC and CA. The low computational complexity of the model allows the use of CA in an autonomous mode and making (correcting) decisions in real time.

Conclusions. The only solution for collecting data from WSN nodes with an unconnected topology is the use of communication aerial platforms. In this case, the task of building a CA flight path with the determination of data collection points (intervals) to ensure the main objective functions: minimum data collection time and/or maximum network operation time.

A specific WSN will differ in many parameters: area size and shape; number, coordinates and density of nodes, their relative positioning; battery power level, amount of monitoring data; characteristics and limitations of nodes and CA parameters, etc. The developed simulation model was used to study the dependence of data collection efficiency indicators on the application of various rules for constructing a CA flight path with certain network, node, and CA parameters. It is proved that there is no single method (set of rules) for finding a CA flight path that provides an optimal solution for all variants of the WSN and possible situations on the network.

To optimize the decision on the flight trajectory and data collection, a rule base is proposed that implements a hierarchy of rules for achieving the objective functions. The research made it possible to determine the priority and order of application of the rules in the database, i.e., the development of meta-rules. The results of simulation modeling have shown that the use of the rule base can reduce the data collection time by up to 20 % or increase the network operation time by up to 15 % compared to existing solutions.

The direction of further research is to improve the rule base for building a flight path and collecting CA data from the WSN nodes for other objective control functions.

REFERENCES

1. Amodu O.A., Nordin R., Jarra C., Bukar U.A., Raja Mahmood R.A., Othma, M.A. (2023). Survey on the Design Aspects and Opportunities in Age-Aware UAV-Aided Data Collection for Sensor Networks and Internet of Things Applications. Drones, 7, 260. DOI: https://doi.org/10.3390/ drones7040260.
2. Minh T. Nguyen, Cuong V. Nguyen, Hai T. Do, Hoang T. Hua, Thang A. Tran, An D. Nguyen, Guido Ala, and Fabio Viola. (2021). UAV-Assisted Data Collection in Wireless Sensor Networks: A Comprehensive Survey. Electronics. 10, 2603. DOI: 10.3390/electronics10212603.
3. Imad Jawhar, Nader Mohamed, Jameela Al-Jarood (2015). UAV-based data communication in wireless sensor networks: Models and Strategies. International Conference on Unmanned Aircraft Systems (ICUAS). DOI: 10.1109/ICUAS.2015.7152351.
4. V. Romaniuk, O. Lysenko, A. Romaniuk, O. Zhuk (2020). Increasing the efficiency of data gathering in clustered wireless sensor networks using UAV. Information and Telecommunication Sciences,11 (1), 102–107. DOI: 10.20535/2411-2976.12020.102-107.

5. Zhiqing Wei, Mingyue Zhu, Ning Zhang, Lin Wang (2022). Zhiyong Feng UAV Assisted Data Collection for Internet of Things: A Survey. IEEE Internet of Things Journal 9(17):1-1, DOI:10.1109/JIOT.2022.3176903.

6. Sarmad Rashedand Mujdat Soyturk (2017). Analyzing the Effects of UAV Mobility Patterns on Data Collection in Wireless Sensor Networks Sensors. 17, 413. DOI: 10.3390/s17020413.

7. Weihuang Huang, Jeffrey Xu Yu. (2017). Investigating TSP Heuristics for Location-Based Services Data Sci. Eng. 2: 71–93. DOI: 10.1007/s41019-016-0030-0.

8. Helsgaun K. (2017). An Extension of the Lin-Kernighan-Helsgaun TSP Solver for Constrained Traveling Salesman and Vehicle Routing Problems; Roskilde University: Roskilde, Denmark. DOI:10.13140/RG.2.2.25569.40807.

9. Hahsler M., Hornik, K. (2007). TSP – Infrastructure for the traveling salesperson problem. J. Stat. Softw., 23, 1–21.

10. Wu Yue, Zhu Jiang (2018). Path Planning for UAV to Collect Sensors Data Based on Spiral Decomposition. Procedia Computer Science 131, 873–879. DOI: 10.1016/j.procs.2018.04.29.

11. Chengliang W, Jun-hui Y (2015). Path Planning for UAV to Collect Sensor Data in Large-Scale WSNs. Transaction of Beijing Institute of Technology; 35: 1044–1049. DOI: 10.1016/j.procs.2018.04.291.

12. Kumar Nitesh, Prasanta K. Jana (2019). Convex hull based trajectory design for mobile sink in wireless sensor networks/International Journal of Ad Hoc and Ubiquitous Computing 30(1):26 DOI: 10.1504/IJAHUC.2019.097092.

13. Q F., Yu W., Xiao K., Liu C., Liu W. (2022). Trajectory generation and optimization using the mutual learning and adaptive colony algorithm in uneven environments. Appl. Sci., 12, 4629. DOI: https://doi.org/10.3390/app12094629.

14. Katoch S., Chauhan S.S., Kumar V. (2021). A review on genetic algorithm: past, present, and future. Multimed. Tools Appl., 80, 8091–8126. DOI: 10.1007/s11042-020-10139-6.

15. Emambocus B.A.S., Jasser M.B., Hamzah M., Mustapha A., Amphawan A. (2021). An enhanced swap sequence-based particle swarm optimization algorithm to Solve TSP. IEEE Access, 9, 164820–164836. DOI: 10.1109/ACCESS.2021.3133493.

16. Haider S.K., Jiang A., Almogren A., Rehman A.U., Ahmed A., Khan W.U., Hamam H. (2021). Energy Efficient UAV Flight Path Model for Cluster Head Selection in Next-Generation Wireless Sensor Networks. Sensors., 21, 8445. DOI: https://doi.org/10.3390/s21248445.

17. Josiane da Costa Vieira Rezende, RoneIlídio da Silva, Marcone Jamilson Freitas Souza. (2020). Gathering Big Data in Wireless Sensor Networks by Drone. Sensors, 20, 6954. DOI: 10.3390/s20236954.

18. Dac-Tu Ho, EstenIngar Grotli, and Tor Arne Johansen (2013). Heuristic Algorithm and Cooperative Relay for Energy Efficient Data Collection with a UAV and WSN. International Conference Computing, Management and Telecommunications (ComManTel). DOI: 10.1109/ComManTel.2013.6482418.

19. Cariou, C., Moiroux-Arvis, L., Pinet, F., Chanet, J.-P. (2023). Evolutionary Algorithm with Geometrical Heuristics for Solving the Close Enough Traveling Salesman Problem: Application to the Trajectory Planning of an Unmanned Aerial Vehicle. Algorithms, 16, 44. DOI: https://doi.org/10.3390/a16010044.

20. Nguyen, K. K., Duong, T. Q., Do-Duy, T., Claussen, H., & Hanzo, L. (2022). 3D UAV Trajectory and Data Collection Optimization via Deep Reinforcement Learning. IEEE Transactions on Communications. DOI: 10.1109/TCOMM.2022.3148364

21. Hrymud A., Romaniuk V. (2023). A model of situational control of the telecommunication aerial platform flight trajectory to collect data from nodes of a wireless sensor network. Communication, informatization and cyber-security systems and technologies, № 3. p. 88–100. DOI: 10.58254/viti.3.2023.12.101.