An Integral Data Gathering Framework for Supervisory Control and Data Acquisition Systems in Green IoT

Xuemei Xiang, Jinsong Gui, and Neal N. Xiong, *Senior Member, IEEE*

***Abstract*—In Green Internet of Things, energy consumption research is a hot topic. Our research focuses on Supervisory Control And Data Acquisition (SCADA) system, and it is a system consisting of a plurality of self-organized sensing networks. Live data gathering from numerous Sub-connected SCADA (S-SCADA) networks to make unified decisions based on collected data is one of pivot problem, which has not been well studied. In this article, an integral data-gathering framework is proposed to prolong network lifetime by using sink rotating and Unmanned Aerial Vehicles (UAV) path planning. In the proposed Sink Rotating joint UAV Data Gathering (SR-UAV-DG) framework, the data gathering process is divided into two organic components: (a) S-SCADA in-network data collection. The energy consumption of the network is balanced by selecting the node with the least energy consumption as the sink node. (b) Use UAV for data collection between S-SCADA networks. The theoretical analysis results show that the SR-UAV-DG framework proposed in this article reduces the maximum energy consumption of nodes in the network by 99.21% after 1000 rounds of data. The flight time of UAV is reduced by 16.83%. In the case of unreliable communication links, the data reception rate is guaranteed to reach 91.94%.**

***Index Terms*—Green Internet of Things, SCADA system, energy saving, data collection, unmanned aerial vehicles.**

I. Introduction

THE DEVELOPMENT of microprocessor technology has promoted the rapid development of the current Internet of Things (IoT), in order to improve the sustainability, prolong the service life of the network, and reduce the system cost, green IoT has been repeatedly mentioned and studied [1]–[3]. The number of devices currently connected to the Internet far exceeds the number of humans, and will be near 50 billion by 2020 [4], [5]. Different types of sensor nodes can perceive various physical phenomena in the environment [6]–[8]. The storage, computing and communication capabilities of current sensing devices have grown considerably and widely used in various applications [9]–[11]. Among them, the Supervisory Control And Data Acquisition (SCADA) system is an important application, which is a computer network system spread over a wide area in order to remotely control and monitoring a specific process in a working field [12], [13].

The Wireless Sensor Network (WSN) deployed by a number of sensor nodes to a specific area through a variety of deployment methods is a widely used SCADA system [14]–[16]. An important feature of this system is that data from different S-SCADA networks will eventually be transmitted back to the control center [17]–[19].

Therefore, data collection is a pivot problem for SCADA system [20]–[22]. In previous studies, there was a sink node connected to the Internet [16], [18], [23]. In many practical applications, especially in dangerous, disaster, battlefield and other application scenarios, sensor nodes are broadcast by means of aircraft [24], [25], and the sensor network after spreading is self-organized into one or more SCADA systems [25], [26]. Thus, in this type of application, even if sensor nodes send the perceived data to the sink in the network [16], [25], the data still cannot be obtained by the control center [27], and how to collect the data in the SCADA system to the control center is a challenge problem [28].

There are many types of SCADA systems. In the most extreme SCADA system, there is only one sensor device. In the study of Bonola *et al.* [29], a strategy for opportunistic routing data collection for this extreme SCADA system is proposed [29], [30]. In smart city, there is a type of sensing device deployed in a smart garbage bin or running light. Such sensing devices are temporarily deployed on the perceptual object according to the needs of the application [29], so the lack of basic communication facilities or the deployment of basic communication facilities is too expensive. In response to this situation, the method of Bonola *et al.* [29] is to use mobile vehicles as data mules to collect data. In the study of Bonola et al, the sensing device has a certain communication range. When the mobile vehicle enters the communication range of the sensing device, the sensing device will transmit data to the mobile vehicle. When a moving vehicle passes through the control center, the collected data will be dumped to the control center [29]. The network model proposed by Borujeni *et al.* [4]

Manuscript received December 30, 2020; revised February 26, 2021; accepted March 15, 2021. Date of publication March 23, 2021; date of current version May 20, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61772554 and Grant 62072475; in part by the Hunan Provincial Innovation Foundation for Postgraduate under Grant CX20190209; and in part by the Fundamental Research Funds for the Central Universities of Central South University under Grant 2019zzts153. *(Corresponding authors: Jinsong Gui; Neal N. Xiong.)*

Xuemei Xiang and Jinsong Gui are with the School of Computer Science and Engineering, Central South University, Changsha 410083, China (e-mail: xuemeixiang@csu.edu.cn; jsgui2010@mail.csu.edu.cn).

Neal N. Xiong is with the Department of Mathematics and Computer Science, Northeastern State University, Tahlequah, OK 74464 USA (e-mail: xiongnaixue@gmail.com).

Digital Object Identifier 10.1109/TGCN.2021.3068257

is similar to the network model in this article. It considers the data collection in the Edge network [4], [31], [32]. However, Fog nodes in the edge network are connected to the Internet [3], [17], [19]. Therefore, as long as one of the SCADA systems self-organized by multiple sensor nodes is connected to the Fog nodes, the Fog nodes act as sinks. Other nodes only need to route data to Fog nodes in the same way as ordinary WSNs to collect data [4].

Network lifetime is the most critical performance indicator in the sensing-based system [8], [16], [18]. In a network with sink as the data collection center, the amount of data transmission undertaken by sinks and nearby nodes is much greater than that of nodes in other regions [16], [18]. In the network with fixed sink, this phenomenon of uneven energy consumption cannot be avoided. Therefore, researchers have proposed a method for moving sinks for the WSN [33]. In a fixed sink network, this uneven energy consumption cannot be avoided. Therefore, researchers have proposed a method of moving sink [33]. The main problem facing this new method is how to establish a link between the mobile sink and the Internet [33].

UAV technology has been developed rapidly in recent years. Its rapid, low price and super adaptability have made it widely used in life, production and research. Ebrahimi *et al.* [34] proposed the use of UAV for data collection. The main idea of the method they used was to divide the WSN into multiple clusters, and the cluster heads collected the data in the cluster. Then, UAV fly to these cluster heads for data collection. The cluster head rotation can achieve energy consumption balance, which can significantly improve the network life [25], [34]. However, the method proposed by Ebrahimi is to study a WSN. The essence of this strategy is the combination of clustering algorithm and UAV path optimization strategy in WSN. Since the coverage of a WSN is small and the number of cluster heads is not large, UAV can complete data collection in a short period, so it is easier to optimize path optimization and energy consumption [34]. However, the data collection situation in the SCADA system is different from the system studied by Ebrahimi et al. Multiple sub-SCADA networks are geographically dispersed, distant distances from each other, and each SCADA network is large too. At this time, the SCADA system chooses different nodes to act as sinks, which not only has an important impact on the energy consumption of the network (network life), but also has a significant impact on the path optimization of the UAV.

The main contributions in this article are proposed as follows: First, the data collection process of the Sink Rotating joint UAV Data Gathering (SR-UAV-DG) scheme is divided into two organic components. During data collection in the S-SCADA network, the network periodically selects the sink node. Second, during data collection between S-SCADA networks, the UAV is path-planned. After the SR-UAV-DG framework proposed in this article performs 1000 rounds of data transmission, the maximum energy consumption of nodes in the network is reduced by 99.21% compared with the sensor network of fixed sink, the flight time of UAV is reduced by 16.83%. Moreover, the data-receiving rate reach 91.94% when the communication link is unreliable.

The rest of this article is arranged as follows: In Section II, some research related to this article is introduced. In Section III, we introduce some models used in this article. In Section IV, the SR-UAV-DG framework proposed in this article is described in detail, including how the SR-UAV-DG framework works and how to calculate its energy consumption. In Section V, the SR-UAV-DG framework is analyzed in detail, and compared with the no-SR-UAV-DG framework. Finally, Section VI is a summary of the full text.

II. Related Work

A new application formed by the development of microprocessing technology–SCADA system [12], [13]. In the SCADA system, a large number of sensing devices are randomly distributed in a wide area [16], [18], [35]. This brings great difficulties to real-time data acquisition, and it is also a key problem to be solved by the SCADA system [19]–[21]. The sensing devices currently connected to the Internet have reached a very large scale. These sensing devices have good computing, storage, and communication capabilities. These characteristics have caused the current network's center to move from the cloud side to the edge network, resulting in new network architectures such as Fog computing, Edge network [4], [21], [31]. At the same time, along with the development of artificial intelligence technology [37], the depth of processing and the breadth of calculations of various applications have been greatly enhanced, and the hidden information is extracted from these data, thus obtaining huge commercial benefits, and profoundly change the original way of understanding and controlling the world. Data collection becomes not only a SCADA system, but also a key technology for the entire current network. The main researches on data collection are: (1) methods; (2) optimization strategies for energy consumption and delay; (3) other important research related to it. This section is mainly discussed from the above three aspects. Network security and privacy protection [38]–[40] is not the focus of this article, and therefore is not in this section.

Data collection in clustered networks is a effective data collection strategy in WSN [16], [41]. Several sensor devices at similar distances form a cluster, and several clusters are combined to form a cluster-based WSN. Each cluster has a special node called Cluster Head (CH) node, and other nodes are called member nodes. The clustering network is mainly aimed at data fusion networks [42]. In such a network, there is redundancy between packets perceived by nodes. Therefore, after multiple packets meet, data fusion can be performed to remove redundant data, thereby merging into smaller packets. In a network of data fusion functions, any multi-packet can be fused into one data packet, called convergecast [42].

Bonola *et al.* [29] proposed a data collection strategy for sensing devices deployed in a smart city without communication infrastructure. The network scenario they studied is this: In smart city, sensing devices are deployed in smart bins, streetlights, or roads that need to be monitored. These deployed sensing devices are deployed in areas where there is no basic communication setup, and many are time-critical and need to be constantly transferred. Therefore, these sensing

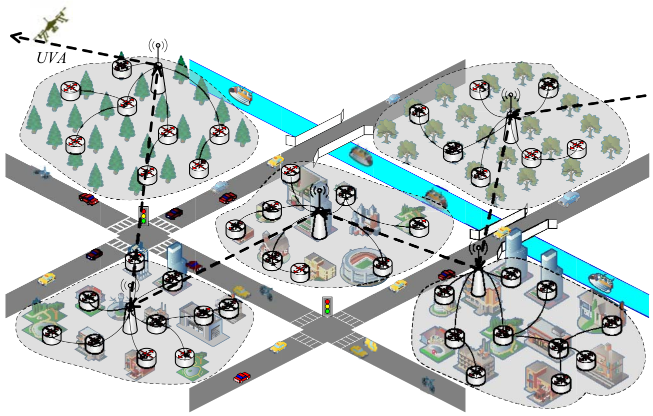


Fig. 1. Network model for our method.

devices can be self-organized into an S-SCADA network, or a single sensing device is a S-SCADA network.

However, such a SCADA system does not have a link to communicate with the Internet. For data collection in such networks, Bonola *et al.* [29] proposed a method for data collection using mobile vehicles in the city. In their strategy, mobile vehicles are used as a medium to "transport" data from sensor devices to the data center. Bonola *et al.* [29] used actual vehicle trajectory data in the city and tested it to prove its effectiveness. They found that the trajectory of mobile vehicles could cover the entire city in a shorter period. However, the above method can only collect data from those transmission devices that the moving vehicle can approach.

The method used by Li *et al.* [25] is to use mobile vehicles for code dissemination on the sensing devices on both sides of the road, and to use the UAV for code dissemination for those who cannot use mobile vehicles.

Ebrahimi *et al.* [34] proposed a method for data collection using UAV for WSN. Their strategy is to divide the network into multiple clusters. Each cluster collects the data of the nodes in the cluster by the CH, and then the UAV flies to each CH for data collection. Ebrahimi *et al.* [34] proposed a corresponding optimization strategy to solve the key technologies in this kind of data collection: How to choose the number and location of CHs to make the network have a long lifetime, and make the UAV flight distance the shortest.

Ebrahimi *et al.* [34] is aimed at data collection of a WSN, and this article is aimed at multiple S-SCADA networks. These S-SCADA networks are complex networks composed of a number of sensing devices, which is rarely studied. Obviously, this kind of complex network is more consistent with the actual network, so its research is more meaningful.

III. System Model and Definitions

A. Network Model

A SCADA system is composed of multiple S-SCADA networks and a central station. Each S-SCADA network independently completes data collection and finally aggregates the data to the central station. In our research, an unmanned aerial vehicle is added as a data transmission medium, which will be responsible for transmitting the data in the S-SCADA network to the central station (as shown in Fig. 1).

TABLE I
Parameter Settings of Symbols

symbol	Meaning	values
d	the distance between the sender and the receiver	
B	the noise bandwidth	$30kHz$
R	the transmission rate	$19.2kbps$
S	the size of a packet	$128\ Byte$
$SNR_{(d)}$	signal-to-noise ratio of two nodes with distance d	
P_s	the power of the node when sending a packet	
$P_{pl(d)}$	the path loss between the sender and the receiver	
P_n	the noise floor	$-115\ dBM$
F	the noise factor	$13\ dB$
T	the ambient temperature	27 °C
B_e	the bandwidth of the system noise	$30\ kHZ$
k	the Boltzmann constant Boltzmann constant	$1.380649 \times 10^{-23} J/K$
d_0	the reference distance	$1m$
$P_{pl(d_0)}$	the path loss at the reference distance	
n	the path loss index	2
x_δ	a shadowing component obeying the Gaussian distribution	
m	hop count	
P	end-to-end reception rate	

B. Reliability Model

The symbols involved in the model and their meanings and values are shown in Table I, refer to [43] and [44]. The probability that the receiver node successfully receives the data packet is called the reception rate. Refer to [44], the reception rate is (1). The value of $SNR_{(d)}$ can be expressed as (2). According to [44], P_n can be expressed as (3), and the relationship between the path loss $P_{pl(d)}$ and the distance between nodes can be expressed as (4).

$$p_{(d)} = \left(1 - \frac{1}{2} \times e^{-\frac{B \times SNR_{(d)}}{2 \times R}}\right)^{8 \times S} \quad (1)$$

$$SNR_{(d)} = P_s - P_{pl(d)} - P_n \quad (2)$$

$$P_n = (F+1)kTB_e \quad (3)$$

$$P_{pl(d)} = P_{pl(d_0)} + 10n{*}\log_{10}\left(\frac{d}{d_0}\right) + x_\delta \quad (4)$$

It is assumed that the data packet has experienced m hops during transmission, and the reception rate of each hop is $p_{(d1)}, p_{(d2)}, \ldots, p_{(dm)}$, respectively. Then, the final receiving rate of the data packet is:

$$P = \prod_{i=1}^{m} p_{(di)}. \quad (5)$$

C. Energy Consumption Model

In our research, energy consumption is mainly related to power and time, and time depends on the amount of data transmitted (the product of the constant of a single data packet and the number of data packets) and the transmission rate. Suppose the node sends/receives Q packets, each packet has a length of L, and the packet transmission rate is R. Thus, the time at which the node transmits the data packet is $L{*}Q/R$, and the

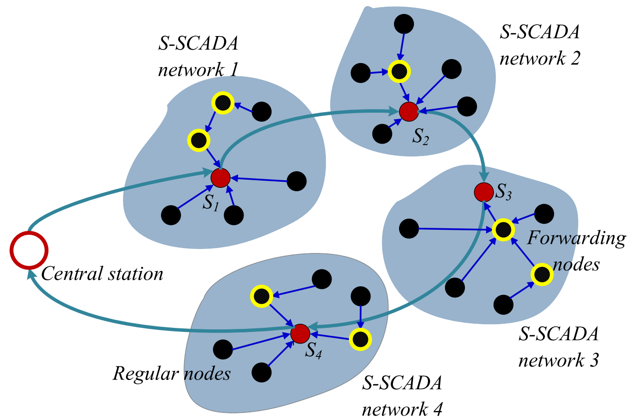


Fig. 2. The process of transmitting packets to a central station.

energy consumption are (6) and (7), where P_{send} ($P_{receive}$) is the power of the node when sending (receiving) a data packet.

$$E_{send} = P_{send} * \frac{L * Q}{R} \tag{6}$$

$$E_{receive} = P_{receive} * \frac{L * Q}{R} \tag{7}$$

where $P_{pl(d)}$ represents the path loss of the node, and:

$$P_{receive} = P_{send} - P_{pl(d)}. \tag{8}$$

D. Problem Statement

Suppose there are M nodes in the network, and E_i is used to represent the energy consumed by node i, where $i = 1, 2, \ldots, M$. Assuming that the maximum energy consumption of nodes in the network is E_{max}, the essential goal of the SR-UAV-DG framework is to make E_{max} as small as possible. Suppose the central station is at point 0, and the sink nodes in each network are at point j ($j = 1, 2, \ldots, N$, N is the number of networks).

Find a path by calculation, and call them as j-th S-SCADA network in the order in which they appear in the path order (as shown in Fig. 2). Therefore, the flight time of the UAV is:

$$t_{fly} = \frac{\sum_{j=0}^{N} \sqrt{\left(x_j - x_{(j+1)}\right)^2 + \left(y_j - y_{(j+1)}\right)^2}}{V} \tag{9}$$

where (x_j, y_j) represents the coordinates of sink node j, V represents the speed of UAV, and (x_{N+1}, y_{N+1}) represents the coordinates of central station. Assume that the UAV stays in network j ($j = 1, 2, \ldots, N$) for t_j. Thus, the time that the UAV flies back to the central station can be expressed as:

$$T_{back} = t_{fly} + \sum_{j=0}^{N} t_j \tag{10}$$

In summary, the main objectives of this article can be expressed as:

$$\begin{cases} min(E_{max}) \\ min(T_{back}). \end{cases} \tag{11}$$

IV. Our Proposed SR-UAV-DG Framework

A. The Selection of Sink Node in the Next Round

After the data transmission in the S-SCADA network is completed, the sink node S first modifies its own routing information to "*Dis* = 0, *Next_Hop* = *Null*", and broadcasts it. Assuming that node A receives this message, it modifies the routing information to "*Dis* = 1, *Next_Hop* = *S*". When node A broadcasts this message, assuming node B has received it, it will modify the routing information to "*Dis* = 2, *Next_Hop* = *A*".

In this loop, all nodes in the network will know the next hop node. For nodes other than the sink node in the network, such as node *A*. After receiving the message, the information of the next hop node (assumed to be *B*) is known. The sink node *S* selects the node with the smallest energy consumption as the sink node S' in the next round (assumed to be *C*). Since node *S* does not know the path to the node *C*, it uses the above-mentioned repeated broadcast method to inform other nodes.

Assuming that a node *D* is located in the network, after node *D* sends out the data packet(s), it needs to wait for a long period of time before it can receive broadcast message about the node S'. To save energy, we slightly adjust the above steps. In the current round, when the node *S* finds the ID of node S', it does not immediately notify it, but waits until the start of the next round. The specific steps are given in Algorithm 1.

B. Energy Consumed by Transmitting Messages

In the SR-UAV-DG framework, there are mainly two types of data. The first type of data is data packets collected by nodes in the S-SCADA network. This type of data has a relatively large length and is the main data type. We use data packets to call it. The second type of data is some additional auxiliary data generated in order to implement the SR-UAV-DG framework. Their length is shorter than the data packet, and we call it a message.

Theorem 1: Using different sensor nodes as sink node in different round requires an additional part of energy to enable nodes in the network to know the next hop route. This part of the energy is:

$$E_{mes} = \begin{cases} P_{bro} * L_{mess} * \frac{n}{R_{bro}} + P_{rec} * \frac{L_{mess}}{R_{bro}} & node\ S; \\ P_{bro} * L_{mess} * \frac{2n}{R_{bro}} + P_{rec} * \frac{L_{mess}}{R_{bro}} & node\ S''; \\ P_{bro} * L_{mess} * \frac{2n}{R_{bro}} + 2 * P_{rec} * \frac{L_{mess}}{R_{bro}} & others \end{cases} \tag{12}$$

where P_{bro} represents the power when the node broadcasts the message, and P_{rec} represents the power when the node receives the broadcast message. L_{mess} indicates the length of a message, n indicates the number of times a message is repeatedly broadcast, and R_{bro} indicates the data transmission rate at the time of broadcast. Node S'' and node *S* represent sink nodes in previous round and in this round, respectively.

Proof: According to Section IV-A, in the previous round, the sink node S'' has determined the ID of the sink node *S* in this round. Thus, at the beginning of this round, node S'' broadcasts the message with the ID information of node *S*. The node that receives this message (for example, node A) will verify whether the ID information in the message matches its own ID. If not, node A broadcasts the message (*n* times). After that, node S starts to broadcast the message. In summary, in the process of transmitting a message, the number of data packets broadcast by the node is:

$$Q_{bro} = \begin{cases} n & sink\ node\ S \\ 2n & other\ nodes \end{cases} \tag{13}$$

Knowing the number of data packets, according to the energy consumption model in Section III-C, we can get the energy consumption:

$$E_{bro} = \begin{cases} P_{bro} * L_{mess} * n / R_{bro} & sink\ node\ S \\ P_{bro} * L_{mess} * 2n / R_{bro} & other\ nodes \end{cases} \tag{14}$$

When the node S'' ibroadcasts the message, the other nodes need to receive it once. It's the same when node *S* broadcasts a message, therefore, we can get (15) and (16).

$$Q_{rec} = \begin{cases} 1 & node\ S\ and\ S'' \\ 2 & other\ nodes \end{cases} \tag{15}$$

$$E_{rec} = \begin{cases} P_{rec} * L_{mess} / R_{bro} & node\ S\ and\ node\ S' \\ 2 * P_{rec} * L_{mess} / R_{bro} & other\ nodes \end{cases} \tag{16}$$

In summary, the energy consumed by a node to transmit a message can be expressed as (12). ■

C. Energy Consumed by Transmitting Packets

Suppose the S-SCADA network *i* is a network of a circular area with a radius of R_i, and the position of the center of the circle is (X_i, Y_i), denoted by O_i. The sensor nodes are evenly distributed in a certain density, and the position of the sink node S_i is (x_i, y_i), where $i = 1, 2, \dots, N$. Assume that the transmission radius selected by the node in the S-SCADA network *i* is r_i.

Taking the position of the sink node S_i as the center and dividing the entire S-SCADA network *i* into several regions with kr_i as the radius, where *k* starts from 1, and takes an integer upward. In order to ensure the reliability, the opportunistic routing scheme is adopted.

Theorem 2: According to our previous assumption, after dividing the S-SCADA network *i* by the above method, it is divided into M_i sub-regions, and:

$$M_i = \left\lceil \frac{\sqrt{(X_i - x_i)^2 + (Y_i - y_i)^2} + R_i}{r_i} \right\rceil. \tag{17}$$

Proof: In order to assist our calculations, we draw a line segment from the point S_i through the point O_i and intersect the circle (the network edge) at the point A_i. Therefore, the distance d_i is equal to (18). Since the division is performed every r_i when dividing the sub-region, the number of sub-regions can be obtained by dividing the length of the line segment $S_i A_i$ by r_i (In the case of inability to divisible, rounded up).

$$d_i = \sqrt{(X_i - x_i)^2 + (Y_i - y_i)^2}. \tag{18}$$

■

Theorem 3: For an S-SCADA network *i* $(i = 1, 2, \dots, N)$, the area of the sub-region k $(k = 1, 2, \dots, M_i)$ can be expressed as:

$$S_k = \begin{cases} \pi(kr_i)^2 - \pi[(k-1)r_i]^2 & 1 \le k \le (R_i - d_i)/r_i \\ \pi {R_i}^2 - \pi(kr_i)^2 - S_{k+1} & (R_i - d_i)/r_i < k < M_i \\ \pi {R_i}^2 - \pi(M_i r_i)^2 & k = M_i. \end{cases} \tag{19}$$

Proof: There are two types of sub-regions. The first one is a ring shape. The outer radius of this sub-region must not exceed the range of the S-SCADA network *i*, which can be represented by (20). The area of this sub-region can be directly obtained by using the area formula of the ring. Therefore, its area is (21), after simplification, it is (22).

$$kr_i \le R_i - d_i \tag{20}$$

$$S_k = \pi(kr_i)^2 - \pi[(k-1)r_i]^2 \tag{21}$$

$$S_k = (2k-1)\pi {r_i}^2 \tag{22}$$

The second type of sub-area is not a complete circle, but a part of the circle. For the convenience of description, we use O_i to represent the center point of the S-SCADA network *i*, and the area where the S-SCADA network *i* is located is called the circle O_i. We use S_i to represent the sink node, and use the circle S_{ik} to represent a circle with S_i as the circle and kr_i as the radius, where $k > (R_i - d_i)/r_i$. First, the area of the sub-region M_i is numerically equal to the area of the circle O_i minus the area it intersects with the circle S_{iM_i}, and the area is (23). The area of the circle O_i is subtracted from the area of the circle $S_{i(M_i-1)}$, and the area of the sub-regions M_i and $M_i - 1$ is obtained. Then, the area of the sub-region $M_i - 1$ can be obtained as (24). By analogy, we have found a method to calculate the area of the second type of sub-region, denoted as (25). Thus, we can summarize the method of calculating the area of each sub-region.

$$S_{M_i} = \pi {R_i}^2 - \pi(M_i r_i)^2 \tag{23}$$

$$S_{M_i-1} = \pi {R_i}^2 - \pi[(M_i - 1)r_i]^2 - S_{M_i} \tag{24}$$

$$S_k = \pi {R_i}^2 - \pi(kr_i)^2 - S_{k+1} \tag{25}$$

■

Theorem 4: For S-SCADA network *i* $(i = 1, 2, \dots, N)$, the number of packets received and sent by the sensor nodes in sub-region k $(k = 1, 2, \dots, M_i - 1)$ can be expressed as:

$$Q_R(k) = \frac{S_{k+1}}{S_k} * Q_S(k+1) * n \tag{26}$$

$$Q_S(k) = 1 + \frac{S_{k+1}}{S_k} * Q_S(k+1) * [1 - (1-p)^n] \tag{27}$$

especially:

$$Q_S(M_i) = 1 \tag{28}$$

$$Q_R(M_i) = 0. \tag{29}$$

Proof: First, for a node in the sub-region M_i, it generates a data packet and does not need to forward packets. According to Theorem 3, there are ρS_{k+1} nodes in sub-region $k+1$; each node sends $Q_S(k+1)$ packets, and each packet needs to be received *n* times. Then, the sub-region *k* needs to receive a total

Algorithm 1 Determining Routes and the New Sink Node

Determine the routing path to the sink node

For each node in the network such as *A* **Do**
  Clear routing information;
End for
IF a node such as *A* happens to be a sink node *S*
  Broadcast message "*Dis* = 0, *Next_Hop* = *S*" *n* times;
Else
  IF node *A* has no routing information yet
    IF node *A* receives a broadcast message *mes* from node *B* such as "*Dis* = *m*, *Next_Hop* = *B*";
    Store the routing information in node *A*;
      Broadcast message "*Dis* = *m* + 1, *Next_Hop* = *A*";
    End if
  End if
End if

//1: Dis: the number of hops to sink node;
//2: E_{con}^{min}: the minimum energy consumed in all nodes;
//3: E_{con}^{A}: the total energy consumed by node A;

Find the sink node S′ in the next round

// In the current round
When sink node *S* receives packets from all nodes **Do**
  $E_{con}^{min} = \infty$;
  Sink_ID = null;
  For each received packet such as $packet_A$ **Do**
    Get E_{con}^{A} from $packet_A$;
    IF $E_{con}^{A} < E_{con}^{min}$
      Sink_ID = *A*;
      $E_{con}^{min} = E_{con}^{min}$;
    End if
  End for
End
// In the next round
Node *S* Broadcast message "*Sink_ID* = *A*";
For each node such as *B* receives message "*Sink_ID* = *A*" **Do**
  Get the ID information of Node *B* (*B.id*);
  IF *B.id* = *Sink_ID*
    Node *B* is the next sink node S';
  Else
    Broadcast message "*Sink_ID* = *A*";
  End if
End

of $\rho S_{k+1} * Q_S(k+1) * n$ packets, and there are ρS_k nodes as receivers in sub-region *k*, so the average number of packets that node needs to receive can be expressed as (30), simplify to (26).

$$Q_R(k) = \frac{\rho S_{k+1} * Q_S(k+1) * n}{\rho S_k} \tag{30}$$

$Q_S(k)$ is equal to the number of packets it actually receives plus one (generated by itself). There are ρS_{k+1} nodes in sub-region $k+1$, each node sends $Q_S(k+1)$ data packets. The probability that these data packets are finally received is $1-(1-p)^n$, where *p* represents the reception rate. These data packets are finally received by ρS_k nodes. According to above, we can get $Q_S(k)$ as (31), simplify to (27).

$$Q_S(k) = 1 + \frac{\rho S_{k+1} * Q_S(k+1) * [1-(1-p)^n]}{\rho S_k}. \tag{31}$$

■

Theorem 5: The number of packets received and sent by the sink node is related to the number of nodes in the sub-region 1 and the packets sent by each node:

$$Q_R(sink) = S_1 * Q_S(1) * n \tag{32}$$

$$Q_S(sink) = 1 + S_1 * Q_S(1) * [1-(1-p)^n]. \tag{33}$$

Proof: The process of reasoning is similar to that of Theorem 4 (receiver is a sink node). ■

According to the energy consumption model in Section III-C, the energy consumed by the node to transmission data packet can be calculated by the number of data packets.

D. Energy Consumed to Keep Listening

Theorem 6: In the S-SCADA network *i* ($i = 1, 2, \ldots, N$), the energy consumed by the node in the sub-region *k* ($k = 1, 2, \ldots, M_i$) to keep listening during a certain round of data transmission is:

$$E_l = E_l(I) + E_l(II) + E_l(III) \tag{34}$$

and:

$$E_l(I) = \begin{cases} 0 & node\ S'' \\ P_l * (k''-1) * L_{mess} * \frac{n}{R_{bro}} & nodes\ in\ area\ k'' \end{cases} \tag{35}$$

$$E_l(II) = \begin{cases} 0 & node\ S \\ P_l * (k-1) * L_{mess} * \frac{n}{R_{bro}} & nodes\ in\ area\ k \end{cases} \tag{36}$$

$$E_l(III) = \begin{cases} \sum_{i=1}^{M_i} Q_R(i) * L_{pac}/R_{pac} & node\ S \\ \sum_{i=k+1}^{M_i} Q_R(i) * L_{pac}/R_{pac} & nodes\ in\ area\ k \end{cases} \tag{37}$$

Here, the power that the node keeps listening is denoted as P_l. The sub-regions the node belongs to in the previous round and the current round are represented by k'' and *k*. The sink points in the previous round and this round are denoted by S' and *S*. The number of times the message is replayed is *n*. The number of data packets received by the node in sub-region *k* is $Q_R(k)$. The length of the message (data packet) is expressed as L_{mess} (L_{pac}). The broadcast rate of the message (data packet) is expressed as $R_{bro}(R_{pac})$.

Proof: The listening period of a node can be divided into three phases. The first phase begins at the beginning of each round, and all nodes in the S-SCADA network wake up, waiting to receive messages broadcast by the sink node S''. That is, each node broadcasts a message (*n* times) and consumes the time $L_{mess} * n/R_{bro}$, which we call a $slot_{bro}$. Starting from the sink node S'', it first broadcasts a message, requiring 1 $slot_{bro}$. Subsequently, the node S'' enters a sleep state. It works for 1 $slot_{bro}$, and keeps listening for 0 $slot_{bro}$ At the same time, in the sub-region 1, the nodes receives the message and relays it (2 $slot_{bro}$), and the time to keep listening is zero. Thus, in the first phase, the node keeps listening for:

$$t_l(I) = \begin{cases} 0 & node\ S'' \\ (k''-1)*L_{mess}*n/R_{bro} & nodes\ in\ area\ k'' \end{cases} \tag{38}$$

where $k'' = 1, 2, \ldots, M_i''$, M_i'' represents the number of sub-regions in the previous round. Then, in the first phase, the energy that the node keeps listening can be expressed as (35). The second phase is similar to the first phase, which can be expressed as (36). After the above two stages, other nodes in the S-SCADA network begin to send data to the sink node *S*. A packet has a length of L_{pac} and transmute rate is R_{pac}. The unit time required to transmit a packet is L_{pac}/R_{pac}, which we called a $slot_{pac}$. The number of data packets that the node in the sub-region M_i needs to receive is $Q_R(M_i)$, and the total listening time is kept as $Q_R(M_i)$.

The nodes in the sub-region M_i-1 first receive $Q_R(M_i-1)$ packets, and $Q_R(M_i-1)$ slots are required. It then sends $Q_S(M_i-1)$ packets out, requiring $Q_S(M_i-1)$ slots. The nodes in the sub-region M_i-2 remain listening until the start of receiving the data packets. Therefore, the time when the node in the sub-region M_i-1 starts to transmit data is the time when it stops listening. The nodes in the sub-region M_i-1 start receiving packets from the $Q_R(M_i)$ $slot_{pac}$, which needs $Q_R(M_i-1)$ $slot_{pac}$. That is to say, the time at which the node in the sub-region M_i-1starts transmitting data is $[Q_R(M_i)+Q_R(M_i-1)]$ $slot_{pac}$. According to this rule, we find that the time of keeping listening is $[Q_R(M_i)+Q_R(M_i-1)+Q_R(M_i-2)]$ $slot_{pac}$.

Therefore, the time that the node in sub-region *k* ($k = 1, 2, \ldots, M_i$) keeps listening is (39), the time that the sink node keeps listening is (40), and the time that the node remains listening during the third phase is (41). Then, the energy consumed by keeping listening during the third phase can be expressed as (37).

In summary, the energy consumed by the node to listen can be expressed as (34).

$$\sum_{i=k+1}^{M_i} Q_R(i)*L_{pac}/R_{pac} \tag{39}$$

$$\sum_{i=1}^{M_i} Q_R(i)*L_{pac}/R_{pac} \tag{40}$$

$$t_l(III) = \begin{cases} \sum_{i=1}^{M_i} Q_R(i)*L_{pac}/R_{pac} & node\ S \\ \sum_{i=k+1}^{M_i} Q_R(i)*L_{pac}/R_{pac} & nodes\ belong \\ & to\ area\ k. \end{cases} \tag{41}$$

■

E. Delay of the in-S-SCADA Network Data Collection

Theorem 7: In the S-SCADA network *i* ($i = 1, 2, \ldots, N$), during the round of data transmission, the time at which the sink node completes the data collection is:

$$T_i^{tot} = 2*\left\lceil \frac{2R_i}{r_i} \right\rceil *L_{mess}*\frac{n}{R_{bro}} + \left(Q_S(1)*m + \sum_{k=2}^{M_i} Q_S(k) \right) *\frac{L_{pac}}{R_{pac}} \tag{42}$$

where R_i is the radius of the S-SCADA network *i* and r_i is the transmission radius of the node. L_{mess} and L_{pac} represent the length of a message and a packet, respectively, R_{bro} and R_{pac} represent the rate at which messages and packets are transmitted, respectively. *n* indicates the number of times the message is replayed, and *m* indicates the number of retransmissions when the node sends a packet to the sink node. $Q_S(k)$ represents the number of packets sent by the node in the sub-region *k* ($k = 1, 2, \ldots, M_i$).

Proof: The process of going through can be divided into three phases like Theorem 6. In the first phase, sufficient time must be given. In the first phase, sufficient time must be given to ensure that no matter where the sink node is, all messages in the network can be received. When the transmission radius is r_i, the node farthest from the node S'' can only receive the broadcast message by $\lceil 2R_i/r_i \rceil$ hopping. Similarly, in the second phase, $\lceil 2R_i/r_i \rceil$ $slot_{bro}$ is also required.

$$T_I = T_{II} = \left\lceil \frac{2R_i}{r_i} \right\rceil *slot_{bro} = \left\lceil \frac{2R_i}{r_i} \right\rceil *L_{mess}*\frac{n}{R_{bro}} \tag{43}$$

In the third phase, the number of data packets that each node in sub-region *k* needs to send is $Q_S(k)$, and the time consumed is $Q_S(k)*slot_{pac}$, where $k = 2, 3, \ldots, M_i$. In the last stage, the nodes in subarea *k* need $Q_S(k)*slot_{pac}$ to send data packets, where $k = 2, 3, \ldots, M_i$. In particular, in order to ensure the acceptance rate, nodes in sub-area 1 send data packets to the sink node through repeated broadcasting, so one data packet requires $m*Q_S(k)*slot_{pac}$:

$$T_{III} = Q_S(1)*m*slot_{pac} + \sum_{k=2}^{M_i} Q_S(k)*slot_{pac} = \left(Q_S(1)*m + \sum_{k=1}^{M_i} Q_S(k) \right) *\frac{L_{pac}}{R_{pac}} \tag{44}$$

In summary, the delay of the data collection in-S-SCADA network *i* can be expressed as (42). ■

F. UAV Path Planning

The central station uses genetic algorithms to plan the path. As long as the UAV enters the sink communication range, it can complete data reception, as shown in Fig. 3. The vertical distance between the UAV and the plane is *h*, and the transmission is reliable when the linear distance of the UAV from the sink node is *d*. Therefore, when the horizontal distance of the UAV from the sink node is $l = \sqrt{d^2 - h^2}$, the transmission is reliable. A path is obtained according to the genetic algorithm, and each network is referred to as the *j*-

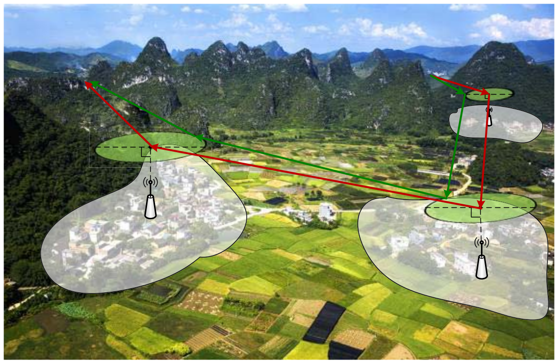


Fig. 3. Optimization of the flight path of the UAV.

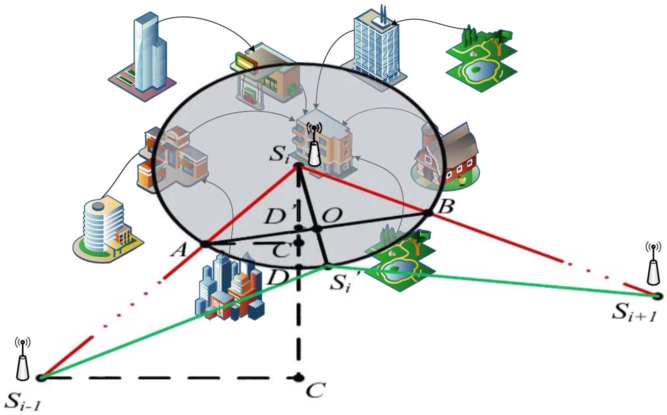


Fig. 4. Analysis chart of the optimization process.

th S-SCADA network in the order in which it appears on this path. For a SCADA network group with N S-SCADA networks, the numbers are from 0 to $N+1$. Among them, the numbers 0 and $N+1$ both represent the central station. For the j-th S-SCADA network ($j = 1, 2, \ldots N$), the coordinates of its sink node S_j are (x_j, y_j). And the coordinates of the sink nodes S_{j-1} and S_{j+1} are (x_{j-1}, y_{j-1}) and (x_{j+1}, y_{j+1}) respectively, as shown in Fig. 4.

Theorem 8: By optimization, the UAV does not need to fly to the sky above each sink node but fly to a specific point $S_j{}'$ to complete the data collection. The coordinates of $S_j{}'$ can be expressed as (X'_{S_j}, Y'_{S_j}):

$$\begin{aligned} X'_{S_j} &= x_j + \sqrt{\left(x_j - X_O\right)^2 + \left(y_j - Y_O\right)^2}/l*\left(X_O - x_j\right) \\ Y'_{S_j} &= y_j + \sqrt{\left(x_j - X_O\right)^2 + \left(y_j - Y_O\right)^2}/l*\left(Y_O - y_j\right) \end{aligned} \tag{45}$$

where point O is an auxiliary point, and its coordinates (X_O, Y_O) can be expressed as $((X_A + X_B)/2, (Y_A + Y_B)/2)$. Point A and point B are the other two auxiliary points whose coordinates are expressed as (47), where l represents the horizontal distance of the UAV from the sink node and can be expressed as (48), d is the distance of reliable transmission, and h is the flying height of the UAV.

$$\begin{cases} X_A = x_j + l/\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}*\left(x_{j-1} - x_j\right) \\ Y_A = y_j + l/\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}*\left(y_{j-1} - y_j\right) \\ X_B = x_j + l/\sqrt{\left(x_j - x_{j+1}\right)^2 + \left(y_j - y_{j+1}\right)^2}*\left(x_{j+1} - x_j\right) \\ Y_B = y_j + l/\sqrt{\left(x_j - x_{j+1}\right)^2 + \left(y_j - y_{j+1}\right)^2}*\left(y_{j+1} - y_j\right) \end{cases} \tag{46}$$

$$l = \sqrt{d^2 - h^2}. \tag{47}$$

Proof: Take S_j as the center and l as the radius to make a circle. The line segments $S_{j-1}S_j$ and S_jS_{j+1} intersect the circle at point A and point B, respectively.

Obviously, the route $S_{j-1} \to A \to B \to S_{j+1}$ is shorter than the line $S_{j-1} \to S_j \to S_{j+1}$.

Through the center S_j, make a line perpendicular to AB, which intersects the circle at point $S_j{}'$ and intersects line AB at point O. We make a horizontal auxiliary line through the node S_{j-1}, make a vertical auxiliary line through the node S_j, and the two auxiliary lines intersect at the point C.

Make a vertical line through point A perpendicular to the line segment S_jC, and the vertical point is C'. Since $\triangle S_{j-1}S_jC$ and $\triangle AS_jC$ are similar, there are:

$$\frac{\left|S_jA\right|}{\left|S_jS_{j-1}\right|} = \frac{\left|C'A\right|_x}{\left|CS_{j-1}\right|_x} = \frac{\left|S_jC'\right|_y}{\left|S_jC\right|_y} \tag{48}$$

where $|AB|$ represents the length of the line segment AB, $|AB|_x$ and $|AB\|_y$ represents the distance between node A and node B (horizontally and vertically). Therefore, we can first calculate the change in the abscissa, it can be expressed as (49). Similarly, the change in the ordinate can be expressed as (50).

Therefore, the abscissa and ordinate of node A and node B can be expressed as (46). The coordinates of the point O can be expressed as $((X_A + X_B)/2, (Y_A + Y_B)/2)$, for convenience, we use (X_O, Y_O) to represent. In order to calculate the coordinates of $S_j{}'$,vertical lines are made to the ordinate by points O and $S_j{}'$, and the vertical points are D' and D, respectively. Through the same idea, the abscissa of $S_j{}'$ can be expressed as (45).

$$\begin{aligned} \left|C'A\right|_x &= \frac{\left|AS_j\right|}{\left|S_{j-1}S_j\right|}*\left|CS_{j-1}\right|_x \\ &= \frac{l}{\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}}*\left(x_{j-1} - x_j\right) \end{aligned} \tag{49}$$

$$\left|S_jC'\right|_y = \frac{l}{\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}}*\left(y_{j-1} - y_j\right). \tag{50}$$

■

G. Delay of Return to the Central Station (UAV)

Theorem 9: The UAV needs to fly from (x_c, y_c) to the following points: $(x_1{}', y_1{}')$, $(x_2{}', y_2{}'), \ldots, (x_i{}', y_i{}'), \ldots, (x_{N-1}{}',$

$y_{N-1}{}'$), $(x_N{}', y_N{}')$, collect the data packets sent by the corresponding sink node, and finally fly back to (x_c, y_c). The time required for this process can be expressed as (51).

$$t_c^{arr} = max\left(t_N^{arr}, T_N^{tot}\right) + n{*}\frac{Q_S(sink_N)}{R_{pac}} + \frac{\sqrt{(x_c - x_N{}')^2 + (y_c - y_N{}')^2}}{V_{max}} \tag{51}$$

and:

$$t_j^{arr} = max\left(t_{j-1}^{arr}, T_{j-1}^{tot}\right) + n{*}\frac{Q_S\left(sink_{j-1}\right)}{R_{pac}} + \frac{\sqrt{\left(x_j{}' - x_{j-1}{}'\right)^2 + \left(y_j{}' - y_{j-1}{}'\right)^2}}{V_{max}} \tag{52}$$

$$t_1^{arr} = \frac{\sqrt{(x_c - x_1{}')^2 + (y_c - y_1{}')^2}}{V_{max}} \tag{53}$$

where t_j^{arr} indicates the time when the UAV reaches *j*-th S-SCADA network ($j = 2, 3, \ldots, N$), and T_j^{tot} indicates the time when the *j*-th S-SCADA network completes data collection. t_c^{arr} indicates the time when the UAV returns to the central station. n indicates the number of times the packet is retransmitted, $Q_S(sink_j)$ indicates the number of packets sent by the sink node in the *j*-th S-SCADA network, and R_{pac} indicates the rate at which the packet is transmitted. (x_c, y_c) represents the coordinates of the central station, and $(x_j', y_j{}')$ represents the coordinates of the point at which the UAV hovered in the *j*-th S-SCADA network ($j = 1, 2, \ldots, N$). V_{max} indicates the speed at which the UAV is flying at full speed.

Proof: The flight time of the UAV can be divided into three parts: the UAV is flying at full speed, and the UAV waits for the sink node to complete the data collection work and the UAV receives the data packet. The time required for the UAV to wait for the sink node to complete the data collection work is related to the time when the UAV reaches the agreed receiving location and the time at which the sink node completes the data collection. The time T_j^{tot} at which the nodes in the *j*-th S-SCADA network ($j = 1, 2, \ldots, N$) complete data collection can be obtained from Theorem 7. The time when the UAV reaches the *j*-th S-SCADA network can be calculated by the UAV reaching the (*j* − 1)-th S-SCADA network(assumed to be t_{j-1}^{arr}). The time for the (*j* − 1)-th S-SCADA network to complete data collection is T_{j-1}^{tot}. If $t_{j-1}^{arr} > T_{j-1}^{tot}$, it means that the UAV can receive the data packet as soon as it arrives, and the time when the data starts to be received is t_{j-1}^{arr}. Otherwise, the time at which data begins to be received is T_{j-1}^{tot}. The UAV starts receiving packets at the time of $max(t_{j-1}^{arr}, T_{j-1}^{tot})$. A total of $Q_S(sink_{j-1})$ packets are sent in the (*j* − 1)-th S-SCADA network, each data packet is repeatedly transmitted n times, and the rate of data packet transmission is R_{pac}. That is, the time required to transmit a packet is $n{*}Q_S(sink_{j-1})/R_{pac}$. After the data-transfer is completed, the time required for the UAV to fly from the (*j* − 1)-th S-SCADA network to the *j*-th S-SCADA network is:

$$\frac{\sqrt{\left(x_j{}' - x_{j-1}{}'\right)^2 + \left(y_j{}' - y_{j-1}{}'\right)^2}}{V_{max}} \tag{54}$$

TABLE II
INFORMATION ABOUT S-SCADA NETWORKS

S-SCADA network i	Center coordinates(m)	Radius (m)	Number of sensor nodes
1	(-2459,964)	534	895
2	(1213,2215)	495	769
3	(-2054,-778)	571	1024
4	(2164,-2362)	511	820
5	(748,-633)	547	939
6	(2240,24)	578	1049
7	(-186,535)	590	1093
8	(-658,-2212)	506	804
9	(-2153,-2497)	501	788
10	(-962,2333)	515	833
11	(2429,1475)	506	804

In summary, the time when the UAV reaches the *j*-th S-SCADA network can be expressed as (52) and (53). The time when the UAV reaches the *N*-th S-SCADA network is t_N^{arr}. It also needs to start receiving data packets after the data collection work in the sink node in the S-SCADA network, and finally fly back to the central station. Thus, the time when the UAV returns to the central station can be expressed as (51). ■

V. Experimental Analysis

On a horizontal plane, there are N S-SCADA networks, which are called S-SCADA network *i* ($i = 1, 2, \ldots N$). The coordinates of the center of the S-SCADA network *i* ((X_i, Y_i)) satisfy $|X_i| \leq 2500$, $|Y_i| \leq 2500$ (one unit length represents one meter). The radius of the S-SCADA network *i* (R_i) satisfies $400 \leq R_i \leq 600$. Sensor nodes are randomly distributed with a density of $\rho = 0.001$. These nodes will generate one data packet in each round, and the length of each data packet is 128 bytes.

The length of the message broadcast by the node is smaller than the packet size, which is 8 bytes. Whether it is transmitting data packets or messages, the reliability model in Section III-B is adopted. The data transfer rate in this model is 19.2 kbps. We placed a central station on the point with coordinates $(-5000, 0)$. The UAVs flying out of the central station pass through the S-SCADA network and collect data. These S-SCADA networks are called the *j*-th ($j = 1, 2, \ldots, N$)S-SCADA network in turn. We randomly generated several S-SCADA networks. Table II gives the information of these S-SCADA networks, such as center coordinates and radius.

A. Analysis and Comparison of Acceptance Rates

In Fig. 5, we show the number of data packets sent by the nodes, received by the UAV and received by the UAV in the S-SCADA network with no-SR-UAV-DG framework (see Theorem 4). In this case, the success rate of the UAV receiving the data packet is only 45.92% (average of the 2000 rounds).

With SR-UAV-DG framework, we selected different transmission powers for the nodes according to the reliability model in Section III-B according to different transmission methods and transmission distances (see Table III for details). Under such an arrangement, the reception rate is up to 99.7%, as can be seen from Fig. 6.

TABLE III
OPTIMIZED RECEIVE RATE

	power	distance	Receive rate	optimization	Optimized reception rate
Node broadcast message	0dBm	60m	0.970752	Repeatedly broadcast 2 times	0.999145
Transfer packets between regular nodes	-2dBm	60m	0.867948	Opportunistic routing (4 candidate nodes)	0.999696
The regular node sends packets to the sink.	0dBm	60m	0.970752	Repeatedly sent 2 times	0.999145
Sink sends packets to UAV	3dBm	100m	0.912816	Repeatedly sent 3 times	0.999337

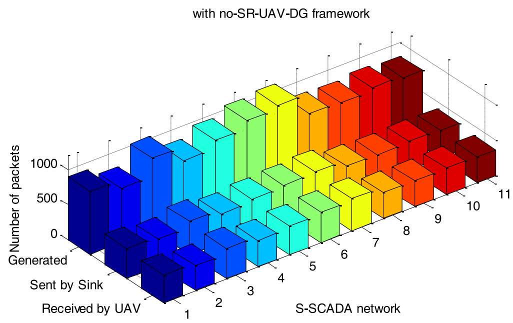


Fig. 5. Number of packets with no-SR-UAV-DG framework.

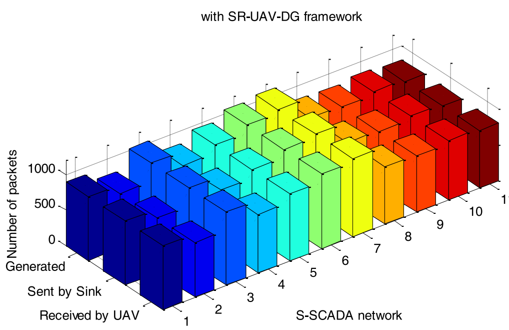


Fig. 6. Number of packets with SR-UAV-DG framework.

TABLE IV
UNIT ENERGY CONSUMED WHEN TRANSMITTING DATA

Type	Sender	Receiver	Length (byte)	Energy consumption（j）	
				Sender	Receiver
Message	node	node	8	3.33333e-006	2.10319e-016
Packet	node	node	128	3.36511e-005	2.12324e-015
	node	Sink	128	5.33333e-005	3.36511e-015
	Sink	UAV	128	0.000106414	2.673e-015

TABLE V
UNIT ENERGY CONSUMED WHEN KEEP LISTENING

Type	Length	Time	Power of LPL	Energy consumption
Message	8 *byte*	0.00333333 s	-15 *dBm*	1.05409e-007 *j*
Packet	128 *byte*	0.0533333 s	-15 *dBm*	1.68655e-006 *j*

B. Analysis and Comparison of Energy Consumption

According to some researches in Sections III and IV, we have given the unit energy consumption of nodes to transmit messages/data packets and keep listening. For details, see Table IV and Table V.

According to Table IV and Table V, we can calculate the energy consumption of each node in the S-SCADA network. In Fig. 7, we show the energy consumption in S-SCADA network 1 (shown every 1000 rounds). In the case of the SR-UAV-DG framework, the maximum energy consumed by the node in the S-SCADA network 1 is 1.9456 Joules (after 2000 rounds), which is 99.295% less than the no-SR-UAV-DG framework (276.074 Joules) (by (55)).

$$\begin{cases} Q_S(M_i) = 1; \\ Q_R(M_i) = 0; \\ Q_R(k) = \frac{S_{k+1}}{S_k} * Q_S(k+1); \\ Q_S(k) = 1 + \frac{S_{k+1}}{S_k} * Q_S(k+1) * p; \\ Q_R(sink) = S_1 * Q_S(1); \\ Q_S(sink) = 1 + S_1 * Q_S(1) * p. \end{cases} \tag{55}$$

In Fig. 8, the ratio of energy savings after using SR-UAV-DG framework is given.

It can be seen that in the network using the SR-UAV-DG framework, the energy consumed by the node is greatly reduced and as the number of transmission rounds increases, the proportion of energy saving is greater.

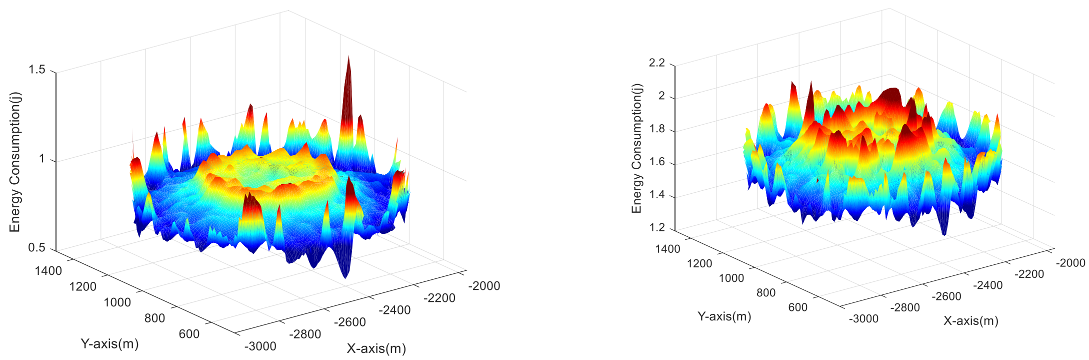
In general, the reason why the SR-UAV-DG framework can maintain node energy balance is that it dynamically selects the sink node according to the energy consumption of the node. The node with the lowest energy consumption will be selected as the sink node.

In the case of the same transmission rounds, in the SCADA system using the SR-UAV-DG framework, the maximum energy consumption value of the node is greatly reduced, which greatly extends the life time of the SCADA system.

C. Analysis and Comparison of Delay

In Sections V-A and V-B, we analyzed the reception rate and energy consumption of nodes in the S-SCADA network using SR-UAV-DG framework, and compared them with the case of using no-SR-UAV-DG framework. In this section, we briefly analyze the delay of S-SCADA network using SR-UAV-DG framework. Here, we define two different concepts of delay respectively.

Beginning with the time when UAV departs from the central station (message and data in the S-SCADA network start to transmit at the same time), the first delay is the time when the sink node in the S-SCADA network completes data collection. It is easy to conclude that in the S-SCADA network using SR-UAV-DG framework, it takes extra time to determine the new sink node and determine the candidate node in the opportunistic route. Therefore, when SR-UAV-DG framework is used, the time for sink nodes to complete data collection increases. According to our research, the increase rate is about 150% to 220% (compared to using no-SR-UAV-DG framework, according to Table VI). Although the time has increased, it does not have much impact on the return time of UAV. This is because: if the central station is far from the sub-networks, the sink nodes have completed data collection before the UAV arrives; if the central station is relatively close, UAV only needs to hover and wait for a while at the first sub-network that passes



(a) After 1000 rounds (b) After 2000 rounds

Fig. 7. Energy consumption in the S-SCADA network 1 using the SR-UAV-DG framework.

TABLE VI
UAV FLIGHT RECORD (WITH AND WITHOUT SR-UAV-DG FRAMEWORK)

Fly				Collect packets				Σ Time (s)	
Distance (m)		Time (s)		Packets Number		Time (s)			
Without	With	Without	With	Without	With	Without	With	Without	With
32071.04	23660.34	400.89	295.75	4706.18	9803.42	31.37	196.06	432.26	491.82

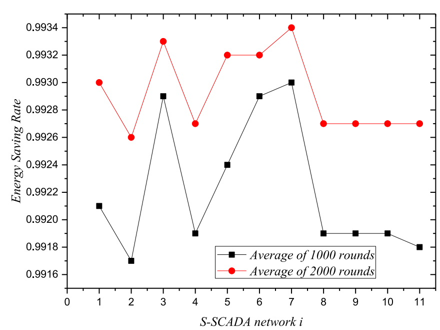


Fig. 8. Average energy saving rate.

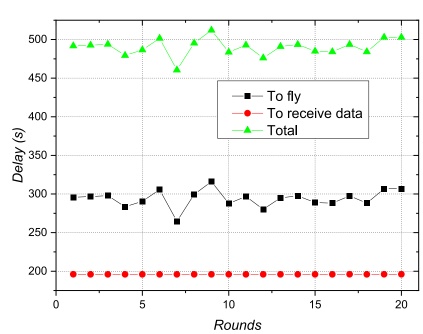


Fig. 9. Delay of SCADA network using SR-UAV-DG framework.

through (we think all S-SCADA networks are of the same scale).

VI. CONCLUSION AND FUTURE WORK

In today's social environment that advocates energy conservation, the green Internet of Things has been proposed and has been rapidly developed. In response to the call for energy saving in the Green Internet of Things, we propose a SR-UAV-DG framework in this article. In the framework we propose, the node with the lowest energy consumption is always selected as the sink node in the S-SCADA network, and the central station uses an optimized genetic algorithm to plan the path for the UAV. Research shows that when the transmission proceeds to the 1000th round, the SR-UAV-DG framework can reduce the maximum energy consumption of nodes in the network by 99.21% (by the 2000th round, it is 99.295%). In the no-SR-UAV-DG framework, the central station is finally able to receive less than 50% of the packets. However, it can reach 99.7% or more with the SR-UAV-DG framework.

In general, the SR-UAV-DG framework provides a new way of thinking about data collection across multiple networks. That is to say, the energy consumption of the node is reduced within the network, the network can run for a longer period, the time for the central station to collect data packets in all networks is reduced, and at the same time, the data reception rate is ensured. In the future, we will try to put SR-UAV-DG framework into practical application scenarios to find more features and application possibilities.

ACKNOWLEDGMENT

The authors thank to the person who provided meticulous and suggestions for improving this article.

REFERENCES

[1] M. Yu, A. Liu, N. Xiong, and T. Wang, "An intelligent game based offloading scheme for maximizing benefits of IoT-edge-cloud ecosystems," *IEEE Internet of Things Journal*, early access, Nov. 23, 2020, doi: 10.1109/JIOT.2020.3039828.

[2] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, and T. Melodia, "Machine learning for wireless communications in the Internet of things: A comprehensive survey," *Ad Hoc Netw.*, vol. 93, pp. 1–46, Jul. 2019.

[3] S. Huang, Z. Zeng, K. Ota, M. Dong, T. Wang, and N. Xiong, "An intelligent collaboration trust interconnections system for mobile information control in ubiquitous 5G networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 1, pp. 347–365, Jan.–Mar. 2021, doi: 10.1109/TNSE.2020.3038454.

[4] E. Borujeni, D. Rahbari, and M. Nickray, "Fog-based energy-efficient routing protocol for wireless sensor networks," *J. Supercomput.*, vol. 74, no. 12, pp. 6831–6858, Aug. 2018.

[5] Q. Liu, G. Wang, X. Liu, T. Peng, and J. Wu, "Achieving reliable and secure services in cloud computing environments," *Comput. Elect. Eng.*, vol. 59, pp. 153–164, Apr. 2017.

[6] Y. Ouyang, A. Liu, N. Xiong, and T. Wang, "An effective early message ahead join adaptive data aggregation scheme for sustainable IoT," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 1, pp. 201–219, Jan.–Mar. 2021, doi: 10.1109/TNSE.2020.3033938.

[7] K. Xie, L.Wang, X. Wang, G. Xie, and J. Wen, "Low cost and high accuracy data gathering in WSNs with matrix completion," *IEEE Trans. Mobile Comput.*, vol. 17, no. 7, pp. 1595–1608, Jul. 2018.

[8] S. Huang, J. Gui, T. Wang, and X. Li, "Joint mobile vehicle–UAV scheme for secure data collection in a smart city," *Ann. Telecommun.*, to be published, doi: 10.1007/s12243-020-00798-9.

[9] K. Xie *et al.*, "An efficient privacy-preserving compressive data gathering scheme in WSNs," *Inf. Sci.*, vol. 390, pp. 702–715, Dec. 2015.

[10] W. Huang, K. Ota, M. Dong, T. Wang, S. Zhang, and J. Zhang, "Result return aware offloading scheme in vehicular edge networks for 6G driving application," *Comput. Commun.*, vol. 164, pp. 201–214, Dec. 2020.

[11] X. Zhu, Y. Luo, A. Liu, W. Tang, and M. Z. A. Bhuiyan, "A deep learning-based mobile crowdsensing scheme by predicting vehicle mobility," *IEEE Trans. Intell. Transp. Syst.*, early access, Oct. 7, 2020, doi: 10.1109/TITS.2020.3023446.

[12] D. Gaushell and H. Darlington, "Supervisory control and data acquisition," *Proc. IEEE*, vol. 75, no. 12, pp. 1645–1658, Jan. 1988.

[13] O. Paladino, F. Fissore, and M. Neviani, "A low-cost monitoring system and operating database for quality control in small food processing industry," *J. Sensor Actuator Netw.*, vol. 8, no. 4, p. 52, Oct. 2019.

[14] Q. Liu, Y. Guo, J. Wu, and G. Wang, "Effective query grouping strategy in clouds," *J. Comput. Sci. Technol.*, vol. 32, no. 6, pp. 1231–1249, 2017.

[15] K. Xie *et al.*, "Recover corrupted data in sensor networks: A matrix completion solution," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1434–1448, Jan. 2017.

[16] X. Liu, H. Song, and A. Liu, "Intelligent UAVs trajectory optimization from space-time for data collection in social networks," *IEEE Trans. Netw. Sci. Eng.*, early access, Aug. 19, 2020, doi: 10.1109/TNSE.2020.3017556.

[17] S. Huang, A. Liu, S. Zhang, T. Wang, and N. Xiong, "BD-VTE: A novel baseline data based verifiable trust evaluation scheme for smart network systems," *IEEE Trans. Netw. Sci. Eng.*, early access, Aug. 7, 2020, doi: 10.1109/TNSE.2020.3014455.

[18] Y. Ren, T. Wang, S. Zhang, and J. Zhang, "An intelligent big data collection technology based on micro mobile data centers for crowdsensing vehicular sensor network," *Pers. Ubiquitous Comput.*, to be published, doi: 10.1007/s00779-020-01440-0.

[19] T. Li, W. Liu, T. Wang, M. Zhao, X. Li, and M. Ma, "Trust data collections via vehicles joint with unmanned aerial vehicles in the smart Internet of Things," *Trans. Emerg. Telecommun. Technol.*, to be published, doi: 10.1002/ett.3956.

[20] B. Jiang, G. Huang, T. Wang, J. Gui, and X. Zhu, "Trust based energy efficient data collection with unmanned aerial vehicle in edge network," *Trans. Emerg. Telecommun. Technol.*, to be published, doi: 10.1002/ett.3942.

[21] J. Tan, W. Liu, T. Wang, M. Zhao, A. Liu, and S. Zhang, "A high-accurate content popularity prediction computational modelling for mobile edge computing by using matrix completion technology," *Trans. Emerg. Telecommun. Technol.*, to be published, doi: 10.1002/ett.3871.

[22] J. Ge, B. Liu, T. Wang, Q. Yang, A. Liu, and A. Li, "Q-learning based flexible task scheduling in a global view for Internet-of-Things," *Trans. Emerg. Telecommun. Technol.*, to be published, doi: 10.1002/ETT.4111.

[23] Y. Ren, Z. Zeng, T. Wang, S. Zhang, and G. Zhi, "A trust-based minimum cost and quality aware data collection scheme in P2P network," *Peer-to-Peer Netw. Appl.*, vol. 13, pp. 2300–2323, Mar. 2020, doi: 10.1007/s12083-020-00898-2.

[24] S. Hajiheidari, K. Wakil, M. Badri, and N. Navimipour, "Intrusion detection systems in the Internet of Things: A comprehensive investigation," *Comput. Netw.*, vol. 160, pp. 165–191, Sep. 2019.

[25] T. Li, A. Liu, S. Zhang, T. Wang, and N. Xiong, "A trustworthiness-based vehicular recruitment scheme for information collections in distributed networked systems," *Inf. Sci.*, vol. 545, pp. 65–81, Feb. 2021.

[26] W. Mo, T. Wang, S. Zhang, and J. Zhang, "An active and verifiable trust evaluation approach for edge computing," *J. Cloud Comput.*, vol. 9, Art. no. 51, Sep. 2020, doi: 10.1186/s13677-020-00202-w.

[27] G. Li, F. Li, T. Wang, J. Gui, and S. Zhang, "Bi-adjusting duty cycle for green communications in wireless sensor networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2020, Art. 166, Aug. 2020, doi: 10.1186/s13638-020-01767-5.

[28] N. Zhang, P. Yang, J. Ren, D. Chen, Y. Li, and X. Shen, "Synergy of big data and 5G wireless networks: Opportunities, approaches, and challenges," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 12–18, Feb. 2018.

[29] M. Bonola, L. Bracciale, P. Loreti , R. Amici, A. Rabuffi, and G. Bianchi, "Opportunistic communication in smart city: Experimental insight with small-scale taxi fleets as data carriers," *Ad Hoc Netw.*, vol. 43, pp. 43–55, Jun. 2016.

[30] M. Chen, Y. Hao, C.-F. Lai, D. Wu, Y. Li, and K. Hwang, "Opportunistic task scheduling over co-located clouds in mobile environment," *IEEE Trans. Service Comput.*, vol. 11, no. 3, pp. 549–561, May/Jun. 2018, doi: 10.1109/TSC.2016.2589247.

[31] G. Zhang, T. Wang, G. Wang, A. Liu, and W. Jia, "Detection of hidden data attacks combined fog computing and trust evaluation method in sensor-cloud system," *Concurrency Comput. Pract. Exp.*, to be published, doi: 10.1002/cpe.5109.

[32] X. Chen, Q. Shi, L. Yang, and J. Xu, "ThriftyEdge: Resource-efficient edge computing for intelligent IoT applications," *IEEE Netw.*, vol. 32, no. 1, pp. 61–65, Jan. 2018.

[33] S. Maurya, V. Jain, and D. Chowdhury, "Delay aware energy efficient reliable routing for data transmission in heterogeneous mobile sink wireless sensor network," *J. Netw. Comput. Appl.*, vol. 144, pp. 118–137, Oct. 2019.

[34] D. Ebrahimi, S. Sharafeddine, P. Ho, and C. Assi, "UAV-aided projection-based compressive data gathering in wireless sensor networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1893–1905, Oct. 2018.

[35] S. Liu, G. Huang, J. Gui, T. Wang, and X. Li, "Energy-aware MAC protocol for data differentiated services in sensor-cloud computing," *J. Cloud Comput.*, vol. 9, Art. no. 48, Aug. 2020, doi: 10.1186/s13677-020-00196-5.

[36] Y. Liu, M. Ma, X. Liu, N. Xiong, A. Liu, and Y. Zhu, "Design and analysis of probing route to defense sink-hole attacks for Internet of Things security," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 356–372, Jan–Mar. 2020, doi: 10.1109/TNSE.2018.2881152.

[37] X. Liu *et al.*, "Adaptive data and verified message disjoint security routing for gathering big data in energy harvesting networks," *J. Parallel Distrib. Comput.*, vol. 135, pp. 140–155, Jan. 2020.

[38] H. Teng *et al.*, "A novel code data dissemination scheme for Internet of Things through mobile vehicle of smart cities," *Future Gener. Comput. Syst.*, vol. 94, pp. 351–367, May 2019.

[39] J. S. Gui, L. H. Hui, N. N. Xiong, and J. Wu, "Improving spectrum efficiency of cell-edge devices by incentive architecture applications with dynamic charging," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 795–808, Feb. 2021, doi: 10.1109/TII.2020.2987089.

[40] Y. Liu, T. Wang, S. Zhang, X. Liu, and X. Liu, "Artificial intelligence aware and security-enhanced trace-back technique in mobile edge computing," *Comput. Commun.*, vol. 161, pp. 375–386, Sep. 2020, doi: 10.1016/j.comcom.2020.08.006.

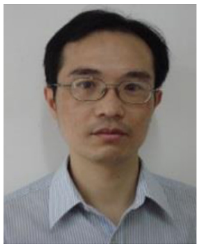
[41] M. Mezghani, "An efficient multi-hops clustering and data routing for WSNs based on Khalimsky shortest paths," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 4, pp. 1275–1288, Jul. 2018.

[42] Z. Li, Y. Liu, A. Liu, S. Wang, and H. Liu, "Minimizing convergecast time and energy consumption in green Internet of Things," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 3, pp. 797–813, Jun. 2020.

[43] A. Laouid, A. Dahmani, A. Bounceur, R. Euler, F. Lalem, and A. Tari, "A distributed multi-path routing algorithm to balance energy consumption in wireless sensor networks," *Ad Hoc Netw.*, vol. 64, pp. 53–64, Sep. 2017.

[44] C. Joo and N. Shroff, "On the delay performance of in-network aggregation in lossy wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 662–673, Apr. 2014.

Xuemei Xiang is currently pursuing the degree with the School of Information Science and Engineering, Central South University, China. Her research interests include device-to-device communications, crowdsensing networks, and wireless sensor networks.



Jinsong Gui received the M.S. and Ph.D. degrees from Central South University, China, in 2004 and 2008, respectively.

He is currently a Professor with the School of Information Science and Engineering, Central South University. His research interests cover the general area of distributed systems, as well as related fields, such as wireless network topology control, performance evaluation, and network security.

Neal N. Xiong (Senior Member, IEEE) received the both Ph.D. degrees from Wuhan University and the Japan Advanced Institute of Science and Technology, respectively. He is currently an Associate Professor (5th year) with the Department of Mathematics and Computer Science, Northeastern State University, Tahlequah, OK, USA.