An Optimal Communications Protocol for Maximizing Lifetime of Railway Infrastructure Wireless Monitoring Network

Xiaoping Ma, Honghui Dong, Xiang Liu, Limin Jia, Guo Xie, and Zheyong Bian

***Abstract*—A wireless monitoring network is an effective way to monitor and transmit information about railway infrastructure conditions. Its lifetime is significantly affected by the energy usage among all sensors. This paper proposes a novel cluster-based valid lifetime maximization protocol (CVLMP) to extend the lifetime of the network. In the CVLMP, the cluster heads (CHs) are selected and rotated with the selection probability and energy information. Then, the clusters are determined around the CHs based on the multi-objective optimization model, which minimizes the total energy consumption and balances the consumption among all CHs. Finally, the multi-objective model is solved by an improved nondominated sorting genetic algorithm II. The simulation results show that, compared with two other strategies in the prior literature, our proposed CVLMP can effectively extend the valid lifetime of the network as well as increase the inspected data packets received at the sink node.**

***Index Terms*—Cluster, energy-efficiency, K-means++, lifetime, nondominated sorting genetic algorithm (NSGA) II, railway monitoring network.**

I. Introduction

With the rapid increase of operating speed and mileage, operational safety of railways has attracted wide attention from both academia and industry. A large portion of railway accidents is caused by infrastructure failures. Thus, rail infrastructure inspections are an indispensable solution to obtaining related information regarding infrastructure conditions (e.g., rail defects, track geometry failure) to ensure railway safety [1], [2]. The track inspection vehicle has been a prevalent technology for railway infrastructure inspection. However, track inspection vehicles are usually used to periodically inspect railway infrastructure and can hardly provide real-time information regarding railway conditions [3] by detecting emergency conditions. Moreover, frequent inspection requires intensive human resources and high capital cost. As a supplementary approach for periodic track inspection, industrial wireless sensor networks (IWSNs) emerge as a preferable technology to inspect the operational condition of the railway infrastructure [4]–[6]. In addition, in order to identify infrastructure defects quickly by means of monitoring the system in real time, IWSNs appear to be an efficient method for monitoring railway infrastructures, especially for complex, extensive railway transportation systems, such as the high-speed rail networks.

In IWSNs, a large quantity of sensors is installed on railway infrastructures (e.g., rails, roadbeds, and bridges) to monitor and transmit condition information [7]. The efficiency of IWSNs greatly depends on the availability and reliability of these sensor nodes. However, the processing ability, communication bandwidth, and energy storage of these sensor nodes are often limited. Therefore, an effective communication protocol is needed to utilize the limited sensing resources. Previous literature demonstrated that the cluster-based routing protocol is one of the energy-saving strategies [8]–[20] that minimizes the total energy consumption and/or balances the energy consumption among all sensors.

To optimally utilize IWSNs, this paper proposes a novel cluster-based valid lifetime maximization protocol (CVLMP), which generates the optimal global clusters dynamically in each transmission round. In the CVLMP, the K-means++ algorithm is used to generate initial clusters based on the deployment of the sensors, and then, the initialized clusters are optimized dynamically in the following rounds. During the optimization process, the cluster heads (CHs) are dynamically selected and rotated from all sensors in the monitoring area to balance their energy consumption. To achieve the target, each CH's rotation is performed based on the probability of CH candidacy, the ratio of residual energy, and predicted energy consumption. Then, the genetic algorithm (GA) is used to optimize clusters with the objective of minimizing the total energy consumption and

Manuscript received June 13, 2017; revised September 7, 2017 and October 24, 2017; accepted November 23, 2017. This work was supported in part by the National Science and Technology Major Project under Grant 2016YFB1200100, in part by the State Key Laboratory under Grant RCS2016ZT018, and in part by the National Natural Science Foundation of China under Grant U1534208 and Grant 51207007. Paper no. TII-17-1257. *(Corresponding author: Limin Jia.)*

X. Ma is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, and also with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ 08817 USA (e-mail: xm54@scarletmail.rutgers.edu).

H. Dong and L. Jia are with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: hhdong@bjtu.edu.cn; jialm@vip.sina.com).

X. Liu and Z. Bian are with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ 08817 USA (e-mail: xiang.liu@rutgers.edu; zheyongbian@gmail.com).

G. Xie is with Shaanxi Key Laboratory of Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an 710048, China (e-mail: guoxie@xaut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2017.2785786

balancing the consumption of each sensor. Experimental simulations are designed to test the performance of the proposed CVLMP. The experimental results show that the CVLMP outperforms the typically used approaches, including the energy-balanced routing method based on forward-aware factor (FAF-EBRM) and multi-objective fuzzy clustering algorithm (MOFCA) [10], [11].

The rest of the paper is organized as follows. Section II reviews the related literature and formulates the problem. The overall scheme of the energy-saving protocol is described in Section III. The novel CVLMP is presented in Section IV. Experimental simulations are analyzed in detail in Section V, and Section VI presents the key conclusions.

II. Related Prior Work and Problem Formulation

As previously mentioned, IWSNs have been widely applied in railway transportation systems, especially in infrastructure condition monitoring, to improve operational safety. IWSNs transmit real-time monitoring information for remote fault diagnosis of critical parts of bridges, tunnels, trains, and rails [4], [5]. However, most existing studies focus on the technological applications of IWSNs in railway systems, whereas the reliability and availability of IWSN information transmission have not received adequate attention. To address this problem, Shafiullah optimized the energy consumption among the sensors installed in the rail cars [6], given a relatively small sensor network. The energy efficiency optimization in a large network with many sensors in railway monitoring systems has rarely been studied, while the number of sensors continues to increase, owing to their wide application in railway inspection.

Some studies focus on improving the energy efficiency of IWSNs using optimization methods. For example, transmission power adjustment technologies help to avoid the interference by controlling the signal coverage area and thus reduce the energy consumption [12]–[14]. Optimizing the deployment of the sensor nodes is useful to reduce the energy consumption as well [15], [17]. The consumption of limited energy is reduced by optimizing the transmitted data packets size [18]. Besides, hierarchization (clustering) is an ideal solution to reducing and balancing the energy consumption. To this end, CH selection/rotation and cluster generation/optimization schemes are two major procedures to establish such a protocol.

For the CH selection and rotation scheme, some existing methods prescribe that those nodes that have not been regarded as CHs in previous rounds might have a higher probability to be selected as CHs [8], [9]. Zhang proposed an FAF-EBRM algorithm for CH selection and rotation scheme using a multihop strategy to balance the energy consumption [10]. In this strategy, the next-hop node is determined by the link weight and forward energy density. This method aims to minimize wasted energy caused by the backward transmission of information. The position of the sensors is considered in the MOFCA [11]. Similarly, the residual energy and distances between sensor nodes are considered in CH selection in predictive energy consumption efficiency, energy-balanced routing protocol (EBRP), and hybrid, energy-efficient, distributed clustering approach [18]–[20]. Kuila applied the particle swarm optimization heuristic algorithm to solve similar problems [21]. However, in these approaches, CH selection and rotation are achieved within each local cluster, which is determined in the initial stage. This may result in a local optimization rather than the global optimization for CH selection and rotation.

For cluster generation, Heinzelman proposed two clustering strategies, LEACH and LEACH-centralized [8], [9], which divide all sensor nodes evenly into k-clusters. Cenedese proposed a distributed clustering strategy to generate the clusters [22]. Khan used the K-means method to classify the sensor nodes according to their relative positions [23]. However, these methods did not consider the impact of cluster size on CH energy consumption. Thus, some improved clustering methods are developed by determining the size of each cluster based on the distances from the sensor nodes to the sink node [24], [25]. However, in those studies, the cluster size remains constant once the clusters are determined, and thus may result in unbalanced sensor node energy consumption in a dynamic operational environment.

To solve the above-mentioned problems, an optimal cluster-based protocol is proposed in this paper for the RIWMS. The lifetime of the system is maximized by minimizing the total energy consumption and balancing the consumption among all sensors simultaneously. The intended contributions of this study can be summarized as follows.

1) The impact factors are comprehensively considered in the modeling of CHs selection and rotation, including the predicted energy consumption of the tentative CHs, the candidacy probability, together with the residual energy of sensors. This addresses the limited consideration of energy consumption after CHs selection in existing studies, which can effectively reduce the selected probability as CHs for incompetent sensors.
2) The cluster scales and the CH-to-members correspondence are dynamically optimized at the beginning of each transmission round, simultaneously considering both the energy consumption minimization and balance, which expands the single focus in previous studies.
3) A hybrid algorithm by blending K-means++ with a nondominated sorting genetic algorithm II (NSGA-II) is developed to achieve the multi-objective optimization model. Using K-means++ can improve the population initialization of NSGA-II so as to accelerate the optimization progress and enhance the quality of the solutions.
4) The proposed CVLMP model enhances the utilization efficiency of the limited energy, which satisfies the real-time and consecutive inspection requirements of the railway wireless monitoring system. This will promote its application in the railway monitoring system and increase the safety of the railway operation.

III. Overall Scheme of the Energy-Efficiency Protocol

A. Overall Structure

In RIWMS, the information transmission network is composed of three layers, as shown in Fig. 1, sensor layer

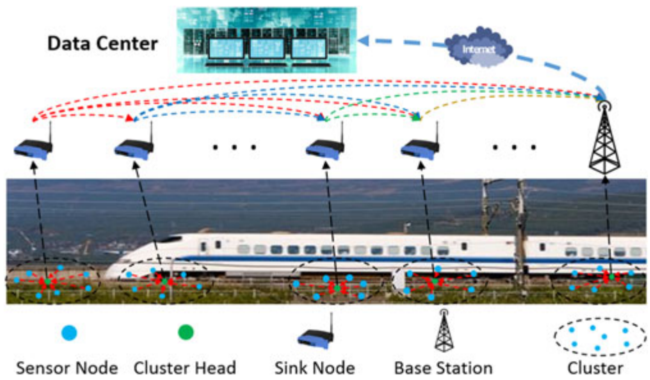


Fig. 1. Schematic structure of the RIWMS.

(infrastructure condition monitoring), communication layer (information transmission), and data processing layer (fault diagnosis and prediction). Since the base station (BS) is located at one end of the monitoring region, it is hard for the sensors to send information to the BS directly. To address this problem, the monitoring region is divided into several small regions, and the sink node is equipped to collect the information from each small region locally. The information transmission progress is shown as follows. First, the information inspected by the sensors is sent to CHs and then forwarded and collected at the local sink nodes. Then, the collected information is transmitted among the sink nodes and forwarded to the BS. Finally, the infrastructure information arrives at the data center from BS via internet. The sink nodes, whose energy storage, computing, and communication ability are all superior to the sensors, act as the local BS in RIWMS to collect the information from the nearby sensors.

B. Overall Framework

We adopt and optimize a cluster-based protocol to reduce the energy consumption in the information transmission between the sensors and sink nodes in the small monitoring region. In this protocol, all the sensors in the small monitoring region are divided into several clusters. One of the sensors in each cluster is selected as the CH to receive the information from other non-CH sensors. Due to the small scale of the monitoring region and high real-time demands, the single-hop protocol is adopted in the communication between the sensors and the corresponding sink node.

As shown in Fig. 2, the scheme is composed of three layers: the input information layer, cluster optimization strategy layer, and output information layer.

In the input layer, the input information, including the energy and location information, is centered around the sensors deployed in the monitoring region.

In the strategy layer, the CHs and clusters are generated with the objective of minimizing the total energy consumption and balancing the consumption among all sensors, and the approach will ultimately extend the RIWMS lifetime. First, the number of the clusters, which is essential to the K-means++, is calculated in advance. Then, the K-means++ algorithm is adopted to initialize the clusters and CHs. The basic idea of this approach

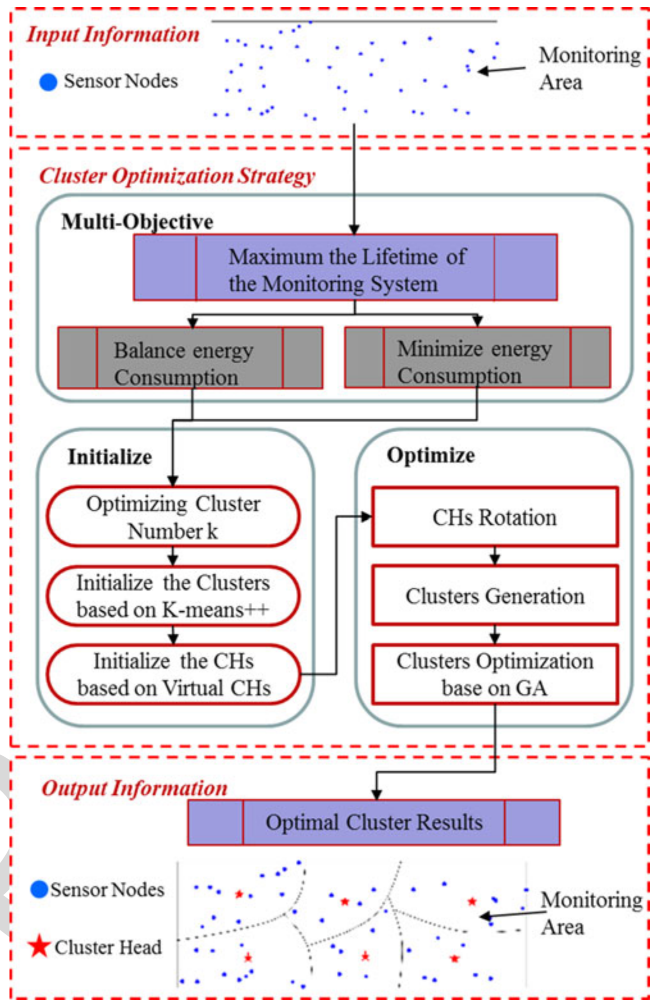


Fig. 2. Overall scheme of RIWMS sensor cluster optimization.

is to cluster the sensors that are close to each other in order to reduce the energy consumption. However, for the complex RIWMS, the CHs and clusters generated by the K-means++ are not necessarily the optimal solutions, and thus, we develop a more effective method in which clusters and CHs are optimized in the following steps. In each optimization round, the CHs are rotated and selected at first, and then, clusters are generated around them. The candidacy probability, residual energy, and predicted energy consumption of all sensors are all considered to determine the CHs. Furthermore, the clusters are obtained by solving a multi-objective optimization model using the GA.

In the output layer, the sensors are allocated to the clusters according to the optimal solutions. The transmission of the inspected information based on the optimal solutions is expected to extend the lifetime of the RIWMS significantly.

IV. Proposed Cluster-Based Valid Lifetime Maximization Protocol

The CVLMP is proposed and introduced in detail in this section, including cluster initialization, CH rotation and selection, and cluster optimization by solving a multi-objective model

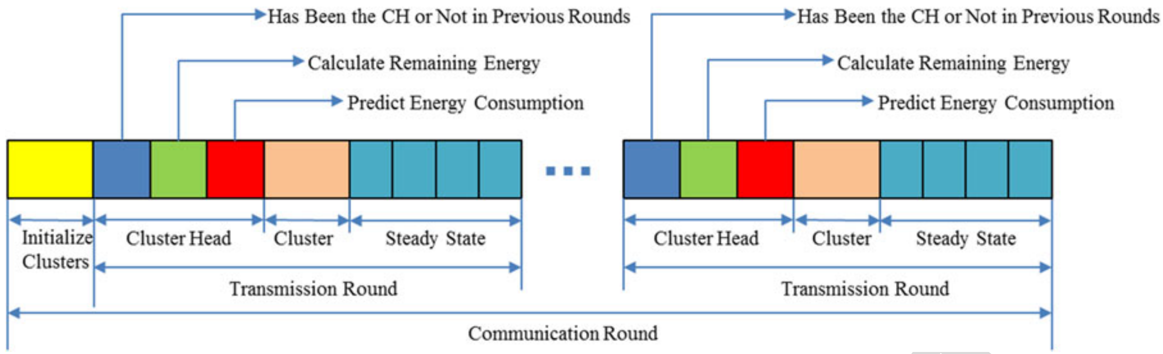


Fig. 3. Structure of the communication round of the proposed CVLMP.

using the NSGA-II. The operational process of the proposed CVLMP is shown in Fig. 3.

The protocol is executed by several communication rounds, including one initialization clustering round and several transmission rounds. In the initialization round, all sensor nodes are divided into several clusters using K-means++. In the transmission rounds, three phases (CH selection and rotation, cluster generation and optimization, and steady transmission) are executed repeatedly.

A. Clusters Initialization by the K-Means++ Algorithm

In the cluster-based protocol, the energy consumption of non-CHs is determined by the distance to the CH and the size of data packets. In this section, the K-means++ algorithm is adopted to initialize the clusters, and they will accelerate the cluster optimization in the following stages.

1) Optimal Number of Clusters: For the K-means++ algorithm, the number of the clusters k should be determined in advance, as it will affect the total energy consumption of all the N sensors in the small monitoring region. Take a short section of the rail and the corresponding infrastructure as monitoring objects; the monitoring field could be considered as a rectangular region. Then, the rectangular region is divided into several small square regions, as mentioned above. The coordinates of the small region are ($x \in [-M/2, M/2]$,$y \in [-(1+\alpha)M, -\alpha M]$). α is the distance coefficient, which is used to adjust the vertical distance between the square region to the sink node. The sink node is located at the origin SK$(0,0)$. The joint probability density of the sensor nodes coordinates is $\rho(x,y) = 1/M^2$.

The total energy consumption of all the sensor nodes in one transmission round is calculated as

$$E_{\text{total}} = \sum_{i=1}^{k} \left(E^{i}_{\text{CH}} + \sum_{j=1}^{n-1} E^{ji}_{\text{Non-CH}} \right) \tag{1}$$

where E^{i}_{CH} is the energy consumed by the ith CHs, $E^{ji}_{\text{Non-CH}}$ represents the energy consumption of the jth sensor node in ith cluster, and n is the number of the sensors in each cluster. We assume that all the N sensors are divided into k clusters equally. Hence, there are $n = N/k$ sensors in each cluster, including one CH sensor and $n-1$ non-CH sensors.

The energy of the CH sensors is consumed in three phases: data packet reception ($E^{i}_{\text{CH}-Rx}$), aggregation ($E^{i}_{\text{CH}-Dx}$), and transmission ($E^{i}_{\text{CH}-Sx}$). The energy consumption of the CHs can be written as

$$\begin{aligned} E^{i}_{\text{CH}} &= E^{i}_{\text{CH}-Rx} + E^{i}_{\text{CH}-Dx} + E^{i}_{\text{CH}-Sx} \\ &= (n-1) \times l \times E_{\text{ele}} + n \times l \times E_{\text{DA}} \\ &\quad + (l \times E_{\text{ele}} + l \times \xi_{\text{mp}} \times d^{4}_{\text{toSK}i}) \\ &= l \times (n \times E_{\text{ele}} + n \times E_{\text{DA}} + \xi_{\text{mp}} \times d^{4}_{\text{toSK}i}) \end{aligned} \tag{2}$$

where E_{ele} denotes the electronics energy coefficient, ξ_{mp} is the amplifier energy coefficients for the multipath fading model, l is the size of data packet, and $d_{\text{toSK}i}$ is the distance from the ith CH sensor to the sink node. It is calculated as

$$d_{\text{toSK}i} = \sqrt{x(i)^2 + y(i)^2}. \tag{3}$$

The non-CHs just transmit their data packets to the corresponding CHs, and the energy consumption is calculated as

$$E^{ji}_{\text{Non-CH}} = l * E_{\text{ele}} + l * \xi_{\text{fs}} * d^{2}_{\text{toCH}ji} \tag{4}$$

where ξ_{fs} is the amplifier energy coefficients for the free space model, and $d_{\text{toCH}ji}$ represents the distance from the jth sensor to the CH in the ith cluster, which is calculated as

$$d_{\text{toCH}ji} = \sqrt{(x(j)-x(i))^2 + (y(j)-y(i))^2}. \tag{5}$$

Statistically, the opportunity of each sensor to be the CH is equal. The expected total energy consumption, as shown in (1), can be rewritten as

$$\begin{aligned} E[E_{\text{total}}] &\approx k * E(E^{i}_{\text{CH}}) + N * E[E^{ji}_{\text{Non-CH}}] \\ &= l * (2N * E_{\text{ele}} + N * E_{\text{DA}} + k * \xi_{\text{mp}} * E[d^{4}_{\text{toSK}i}] \\ &\quad + N * \xi_{\text{fs}} * E[d^{2}_{\text{toCH}ji}]). \end{aligned} \tag{6}$$

The expected fourth (4th) power of $d^4_{\mathrm{toSK}i}$ is calculated as

$$\begin{aligned} E[d^4_{\mathrm{toSK}i}] &= \int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\left(\sqrt{x^2+y^2}\right)^4\rho(x,y)dxdy \\ &= \int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\left(x^2+y^2\right)^2\rho(x,y)dxdy \\ &= \left(0.0125+\frac{(\alpha+1)^3-\alpha^3}{18}+\frac{(\alpha+1)^5-\alpha^5}{5}\right)*M^4. \end{aligned} \tag{7}$$

The expected of $d^2_{\mathrm{toCH}ji}$ is expressed as

$$\begin{aligned} E[d^2_{\mathrm{toCH}ji}] &= \frac{1}{2}\int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2} \\ &\quad\times\left(\sqrt{(x_j-x_i)^2+(y_j-y_i)^2}\right)^2 \\ &\quad\times\rho(x_j,x_i,y_j,y_i)dx_jdx_idy_jdy_i \\ &= \frac{M^2}{6k}. \end{aligned} \tag{8}$$

The initial optimal number of clusters is calculated by setting the derivative of $E[E_{\mathrm{total}}]$ in (6), with respect to k, to zero, as

$$k_{\mathrm{opt}} = \sqrt{\frac{\xi_{\mathrm{fs}}*N*M^2}{6\xi_{\mathrm{mp}}*d^4_{\mathrm{toSK}i}}}. \tag{9}$$

The optimal number of clusters k_{opt} is calculated by (9) when the energy of sensors is sufficient to act as CHs. However, the residual energy of the sensors will decrease as the communication travels. Once the residual energy is not enough to transmit all the information to the sink node as a CH in the current cluster scale, the number of clusters will increase to reduce the sensors in one cluster and decrease the load on the CHs. In this manner, the energy consumption of each CH will be reduced, and the information transmission will be resumed. The strategy is helpful to extend the system lifetime.

2) Initialize Clusters Based on K-Means++: After the number of clusters k is determined, the K-means++ is performed to initialize the clusters. In this algorithm, the selected initial cluster centers are spread out as much as possible [27], ensuring the dispersion of the clusters. The algorithmic details of the K-means++ are shown in Table I.

In step 4, the sensors far from the selected cluster centers will be selected as the new cluster center with high probability. In step 5, all the sensors will be divided into the k_{opt} clusters based on the selected cluster centers.

3) Initialize CHs Based on K-Means++: As shown in Fig. 4, the initial CHs are selected from the k clusters. The virtual CHs (the center points of clusters) are calculated by (10). Then, the sensors nearest to each virtual CH are selected as the initial CHs as follows:

$$\text{Virtual CHs } (X(k),Y(k)) = \left(\frac{\sum_{i=1}^{n_k}x(i)}{n_k},\frac{\sum_{i=1}^{n_k}y(i)}{n_k}\right) \tag{10}$$

TABLE I
CLUSTER INITIALIZATION BASED ON K-MEANS++

K-means++ Algorithm Cluster Initialization
Input: The N sensors deployed in a square monitoring region;
Output: The N sensors are divided into k_{opt} clusters;
Step 1: determine the number of the clusters k_{opt} based on (13).
Step 2: select one sensor as the first cluster center (CC_1).
Step 3: calculate the distance between sensors and selected cluster centers;
repeat
for $i=1,2,\ldots,k$; cluster centers
for $j=1,2,\ldots,N-k$; cluster members
$d_{\mathrm{toCC}ji}=\sqrt{(x(j)-x(i))^2+(y(j)-y(i))^2}$; calculate distance
until all the distances are calculated.
Step 4: select the kth cluster center (CC_k, $1\le k\le k_{\mathrm{opt}}$)
repeat
for $j=1,2,\ldots,N-k$
for $i=1,2,\ldots,k$
$d_{\mathrm{toCC}j}=\min(d_{\mathrm{toCC}ji})$; select the minimum distance;
$\mathrm{sum}(d(x_{\mathrm{toCH}j}))=\sum_{j=1}^{N-k}d_{\mathrm{toCC}_j}$
$p_j=d(x_{\mathrm{toCH}j})/\mathrm{sum}(d(x_{\mathrm{toCH}j}))$; define the selection
probability of the cluster centers
until all the k_{opt} cluster centers are selected.
Step 5: Generate the initial clusters based on the k_{opt} CHs.
repeat
for $i=1,2,\ldots,k_{\mathrm{opt}}$
for $j=1,2,\ldots,N-k_{\mathrm{opt}}$
$d_{\mathrm{toCC}ji}=\sqrt{(x(j)-x(i))^2+(y(j)-y(i))^2}$
for $j=1,2,\ldots,N-k_{\mathrm{opt}}$
for $i=1,2,\ldots,k_{\mathrm{opt}}$
$d_{\mathrm{toCC}j}=\min(d_{\mathrm{toCC}ji})$
dividing the jth sensor into the ith cluster;
until all the sensors are grouped into the corresponding clusters.

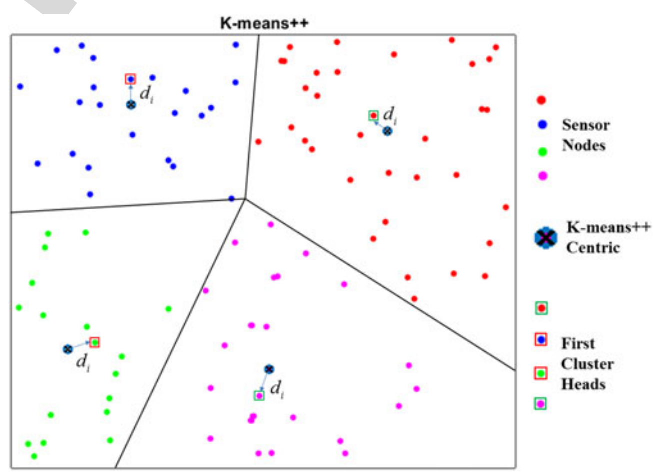


Fig. 4. Cluster initialization based on K-means++.

where n_k is the number of sensors in the kth cluster.

The initial clusters and CHs obtained by K-means guarantee that the sensors close to each other are divided into the same cluster, which can reduce the energy consumption of non-CHs. However, the size of the clusters is unpredictable, which is related to the balance of CHs energy consumption. To prolong the lifetime of the RIWMS, the initial CHs and clusters should be further optimized. The K-means++ method obtaining the initial CHs and clusters is not necessarily able to fully utilize the limited energy. They will be used as the initial clusters in the following cluster optimization.

B. CH Selection and Rotation Probability Model

In the cluster-based protocol, the energy consumption of CH is much more than that of non-CH sensors. Selecting and rotating the CHs before each transmission round starts will help to balance the energy among all sensors. In the CH selection and rotation, in addition to the candidacy probability and the sensors residual energy, the predicted energy consumption after the election is considered. This model could reduce the CH's selected probability of incompetent sensors.

1) Probability of CHs Candidacy: As described above, to balance the energy consumption among all sensors, each sensor should have a completely equal opportunity to operate as CH. This means that the N/k sensors in each cluster will have one chance to act as the CH in the following N/k rounds [8], [9].

The probability of CH candidacy considers whether the sensor is selected as CH in the last $r-1$ rounds. The probability is defined as [8]

$$p_{1i}(r) = \begin{cases} \frac{k}{N-k*r}, & C_i(r)=1 \\ 0, & C_i(r)=0 \end{cases} \tag{11}$$

where $r \in [1, \lfloor N/k \rfloor]$ means that when all sensor nodes have rotated as CHs in previous rounds, r will be reset to 1 and increase to $\lfloor N/k \rfloor$ in each subsequent round. $\lfloor N/k \rfloor$ means that r will always round down to the nearest integer. $C_i(r)$ represents the condition of the node i. If the node has been the CH in the previous r rounds, we set $C_i(r) = 0$ and vice versa.

2) Ratio of Residual Energy: The residual energy of the sensors will vary after several transmission rounds, since

1) the energy consumptions of CHs and non-CHs are different;
2) the energy consumption of different CHs varies due to the different cluster scales and transmission distances; and
3) the different distances from non-CHs to the CHs will also lead to the different energy consumptions of non-CHs.

Therefore, in the CH selection and rotation phases, a higher ratio of residual energy yields a greater probability of CH selection. The ratio of the residual energy is defined as

$$p_{2i}(r) = \frac{\left|E^i_{\text{Re}}(r) - E^{\min}_{\text{Re}}(r)\right|}{\sum_{i=1}^{n}\left|E^i_{\text{Re}}(r) - E^{\min}_{\text{Re}}(r)\right|} \tag{12}$$

where $E^i_{\text{Re}}(r)$ represents the residual energy of the ith sensor before the rth round, and $E^{\min}_{\text{Re}}(r)$ is the minimum residual energy of all sensors in the cluster. Furthermore, the residual energy of the ith sensor before the rth round is computed as

$$E^i_{\text{Re}}(r) = E^i_{\text{Re}}(r-1) - E^i_{\text{Co}}(r-1) \tag{13}$$

where $E^i_{\text{Co}}(r-1)$ is the energy consumed by the ith sensor in the $r-1$th round. If the sensor is CH in the $r-1$th round, it is computed according to

$$E_{\text{Co}-\text{CH}}(r-1) = E_{Rx}(r-1) + E_{Dx}(r-1) + E_{Tx}(r-1), \tag{14}$$

$$E_{Rx}(r-1) = (n_k - 1) * l * E_{\text{ele}}, \tag{15}$$

$$E_{Dx}(r-1) = n_k * l * E_{\text{DA}}, \tag{16}$$

$$E_{Tx}(r-1) = l * E_{\text{ele}} + l * \xi_{\text{mp}} * d^4_{\text{toSK}}(r-1). \tag{17}$$

If the sensor is non-CH in the $r-1$th round, it is computed by

$$E_{\text{Co}-\text{NCH}}(r-1) = l * E_{\text{ele}} + l * \xi_{\text{fs}} * d^2_{\text{toCH}}(r). \tag{18}$$

3) Ratio of the Predicted Energy Consumption: The predicted energy consumption of the CHs after the election is another essential factor. The sensors with less energy consumption as CHs will be selected in the following transmission round with high probability. The ratio of the predicted energy consumption in the following $r+1$th round is defined as

$$p_{3i}(r) = \frac{\left|E^i_{\text{Co}}(r+1) - E^{\max}_{\text{Co}}(r+1)\right|}{\sum_{i=1}^{n_k}\left|E^i_{\text{Co}}(r+1) - E^{\max}_{\text{Co}}(r+1)\right|} \tag{19}$$

where $E^i_{\text{Co}}(r+1)$ represents the predicted energy consumption of the ith sensor in the following $r+1$th round while it is selected as CH. $E^{\max}_{\text{Co}}(r+1)$ is the maximum predicted consumption of the sensors in the cluster. The predicted energy consumption can be calculated by (14)–(17).

The above-mentioned three factors are all considered in the CH selection and rotation probability model. The probability is defined as

$$\begin{aligned} &p_i(r) = \omega_1 p_{1i}(r) + \omega_2 p_{2i}(r) + \omega_3 p_{3i}(r) \\ &\text{S.T.} \quad \sum_{i=1}^{3}\omega_i = 1 \\ &\qquad 0 \le p_{1i}(r), p_{2i}(r), p_{3i}(r) \le 1 \end{aligned} \tag{20}$$

where $\omega_i, i = 1, 2, 3$ are the weighting coefficients, and they are used to adjust the importance of each factor to the model.

The most appropriate CHs are selected and rotated based on this model before each transmission round starts. They can be used to support the cluster generation and optimization in the following phases.

C. Cluster Generation and Optimization Probability Model

After the CHs are determined, the clusters will be updated accordingly to optimize the communication protocol and extend the system lifetime.

For the RIWMS, the death of any sensor may potentially lead to system instability or inspection failures. In this paper, the valid lifetime of the system is defined as the time when more than 90% of the sensors are alive. Hence, the optimal clusters should guarantee all sensors remain alive as long as possible.

The optimal clusters are generated based on the following two schemes. 1) Adjust the clusters scales to balance the energy consumption among CHs. 2) Optimize the correspondence among the CHs and non-CHs to minimize the total energy consumption.

1) Scale of Clusters: As mentioned before, the energy consumption of the CHs is related to d_{toSK} and the scale of clusters (number of non-CHs). d_{toSK} changes as the CHs rotate, and hence, the scales of clusters should vary correspondingly to balance the energy consumption among the CHs. The optimization model is defined as

$$f_1(E_{\mathrm{CH}i}(n_i)) = \min \frac{\sum_{i=1}^{k}\left(E_{\mathrm{CH}i}(n_i) - \overline{E}_{\mathrm{CH}}\right)^2}{k} \tag{21}$$

where $f_1(E_{\mathrm{CH}i}(n_i))$ aims to minimize the variance of the CHs energy consumption, which represents the energy consumption balance degree among all CHs. $E_{\mathrm{CH}i}$ represents the energy consumption of the CH in the ith cluster; $\overline{E}_{\mathrm{CH}}$ represents the average energy consumption for all CHs; and k is the number of clusters.

The energy consumed by the CH in the ith cluster is calculated based on (1), as

$$E_{\mathrm{CH}i}(n_i) = l * (n_i * E_{\mathrm{ele}} + n_i * E_{\mathrm{DA}} + \xi_{\mathrm{mp}} * d_{\mathrm{toSK}i}^4)$$
$$\text{S.T.} \quad \sum_{i=1}^{k} n_i \le N \tag{22}$$

where $n_i (i = 1, 2, \ldots, k)$ is the number of sensors in the ith cluster. The constraint implies that only the living sensors participate in the clusters generation and optimization.

The average energy consumption of all the CHs is calculated as

$$\bar{E}_{\mathrm{CH}} = \frac{\sum_{i=1}^{k} E_{\mathrm{CH}i}(n_i)}{k}. \tag{23}$$

2) CHs to Non-CHs Correspondence: Based on the above-mentioned model, the CHs have almost identical energy consumptions, while the energy consumption of each non-CHs sensor would vary while divided into different clusters. In the following model, the correspondences between the CHs and non-CHs will be optimized to minimize the energy consumption of non-CHs. The total energy consumption optimization model is given by

$$f_2(E_{\mathrm{NCH}i}, E_{\mathrm{CH}j}) = \min\left(\sum_{i=1}^{N-k} E_{\mathrm{NCH}i} + \sum_{j=1}^{k} E_{\mathrm{CH}i}\right), \tag{24}$$

$$E_{\mathrm{NCH}i} = l * E_{\mathrm{ele}} + l * \xi_{\mathrm{fs}} * \sum_{j} r_{ij} * d_{\mathrm{toCH}j}^2, \tag{25}$$

$$r_{ij} = \begin{cases} 1, & \text{sensor } i \text{ is allocated to cluster } j \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

where $f_2(E_{\mathrm{NCH}i}, E_{\mathrm{CH}j})$ aims to minimize the total energy consumption of all sensors. $E_{\mathrm{NCH}i}$ is the energy consumption of the ith non-CH sensor and $E_{\mathrm{CH}j}$ is the energy consumption of the CH in the jth cluster. r_{ij} is the decision variable.

The scales of the clusters are optimized to balance the energy consumption among the CHs, while the total energy consumption is minimized by optimizing the correspondence between the CHs and non-CHs members. The optimization of the correspondence between the CHs and non-CHs will reversely affect the cluster scales. It is obvious that the cluster generation is a multi-objective optimization problem. It can be formulated as

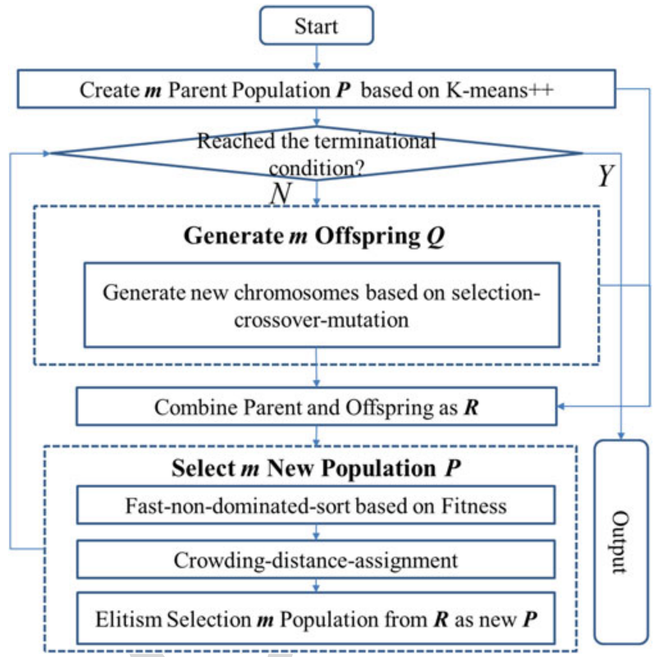


Fig. 5. Flowchart for cluster optimization using NSGA-II.

$$\min f(E) = (f_1(E_{\mathrm{CH}i}), f_2(E_{\mathrm{CH}i}, E_{\mathrm{NCH}i})). \tag{27}$$

The following two constraints should be considered: 1) there are at least two nodes in each cluster, i.e., $2 \le n_i \le N$; and 2) the energy-exhausted, "dead" sensors should be excluded from cluster generation, i.e., $\sum_{i=1}^{k} n_i \le N$.

The cluster generation and optimization is an NP-hard problem. Exact analytical methods face difficulty in obtaining the optimal solutions when the scale of the problem is large. Heuristic methods such as the GA are effective to solve such multi-objective optimization problems in practice. Q1

D. Cluster Optimization Based on the NSGA-II

The multi-objective optimization algorithm is adopted to optimize the clusters to minimize total energy consumption and balance the consumption among all CHs. The NSGA-II is one of the effective GA methods to solve multiple objective optimization problems [26]. The optimization process is described in Fig. 5. The optimal clusters can be obtained by the NSGA-II.

1) Population Initialization: The optimization process based on the NSGA-II is intended to obtain the best solution (chromosomes) based on the initial population. The population consists of m chromosomes. Each of them is composed of M genes, and their positions and contents denote the sensor index and cluster number [see Fig. 6(a)]. The sensor index $[1, 2, \ldots, N]$ will be assigned to sensors when they are deployed in the monitoring region. The cluster number $[1, 2, \ldots, k]$ represents which cluster the sensors belong to.

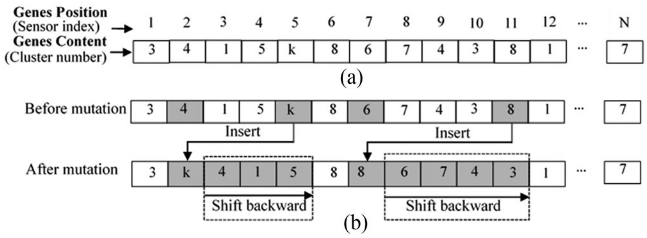


Fig. 6. Representation and mutation: (a) Chromosomal representation; and (b) shift mutation.

The initialization of the population would have a substantial influence on the cluster optimization speed and results, because all the chromosomes are generated based on the initial ones in the parent population. The initialized clusters based on K-means++ in this paper can provide better first chromosome in the initial parent population P. In the existing studies, the initial chromosome is generally generated randomly, while its genes are randomly scattered in the solution space. In contrast, the genes in the chromosome obtained by K-means++ are more concentrated and have smaller energy consumption. This initial chromosome with higher fitness values (function values) will accelerate the optimization process. Subsequently, the other $m-1$ chromosomes in the parent population are generated using shift mutation methods based on the first chromosome [28]. As shown in Fig. 6(b), two pairs of points are selected at random, then the rear points are inserted ahead of the front point and the front points are all shifted backward. The number of point pairs is determined by the mutation probability p_m.

2) Population Optimization: In the population optimization phase, we try to generate the optimal chromosomes within an acceptable timeframe. The offspring Q with m chromosomes is generated based on the parent population through crossover operation and the roulette wheel selection operation.

The individual selection probability is defined as

$$p(c_i) = \frac{g(c_i)}{\sum_{i=1}^{m} g(c_i)} \tag{28}$$

where c_i is the ith chromosome and $g(c_i)$ is its fitness value, denoted as

$$\max g(c_i) = \left(\frac{1}{f_1(E_{\mathrm{CH}i})}, \frac{1}{f_2(E_{\mathrm{CH}i}, E_{\mathrm{NCH}i})} \right). \tag{29}$$

Theoretically, greater residual energy renders larger probability of selection. The two-point crossover and shift mutation are adopted in the following steps to create the new offspring Q. Subsequently, the offspring Q and the parent population P are combined into a new population R. The NSGA-II method is used to update the parent population P from R. Finally, the best chromosomes in the updated parent population P will be selected as the optimal clusters. The detailed process is shown in Table II.

TABLE II
STEPS OF THE NSGA-II METHOD

NSGA-II Algorithm Population Optimization

Input: Algorithm parameters, the population scale *m*, the iteration times T;
Output: Pareto optimal solution set P;
Step 1: Create m initial chromosomes as the initial population P based on K-means++;
Step 2: Generate m chromosomes in offspring Q based on the traditional genetic algorithm;
Step 3: Combine the parent with offspring in population R;
Step 4: Fast-nondominated-sort $R(m_i)$
All chromosomes in R are allocated into several layers according to dominant relationships. The nondominant chromosomes in R are allocated to the first layer; the nondominant chromosomes in the remaining population (removing the chromosomes in the first layer from R) are allocated into the second layer, and so on. The chromosomes in the same layer are assigned the same Pareto value;
Step 5: Crowding distance assignment $D(m_i)$
The crowding distance is defined as the sum value of multi-objective functions. This is important because it allows the ranking of chromosomes in the same layer;
Step 6: Elitism selection
Elitism selection is the selection of *m* chromosomes from R based on $R(m_i)$ and $D(m_i)$. This strategy can be used to retain good individuals and improve the overall level of population evolution;
Step 7: Judge if the terminal criteria are satisfied. If so, output the solution; if not, return to ***Step 2.***

E. Steady Communication Phase

After the optimal clusters and the most energy-efficient transmission route are determined, the system steps into the steady communication phase. In this phase, the infrastructure condition information is inspected, synthesized, and transmitted from the sensors to the sink node. The steady communication process is executed as follows. First, the inspected data packets from non-CH sensors in each cluster are sent to the corresponding CH. Then, all the received data packets are synthesized by their respective CHs, which can reduce information redundancy and minimize energy consumption. Finally, the synthesized information packets at all CHs are transmitted to the sink node.

As shown in Fig. 3, the transmission rounds will repeat as long as the sensors have sufficient energy. In this optimal communication protocol designed above in each round, the communication between the sensors and sink node will operate in the energy-efficiency ways. The valid lifetime of the railway infrastructure monitoring wireless system will be extended ultimately.

V. SIMULATION VALIDATION AND ANALYSIS

In this section, the proposed protocol is validated via computer simulations with Python 3.6.2. Comparative case studies are carried out to demonstrate the superiority of the proposed scheme.

A. Simulation Configuration

In the paper, a rectangular region ($L \times W$: 500 m $\times$ 50 m) along the rail is taken as the RIWMS monitoring area. The monitoring region is divided into ten small square regions ($L \times W$: 50 m $\times$ 50 m) evenly, with ten sink nodes located above them. The information inspected by the sensors in the small square monitoring regions is sent to the corresponding sink nodes and then transmitted to the BS.

TABLE III
PARAMETERS OF THE OPTIMIZATION MODEL

No.	Parameter	Describe	Value
1	$E_{\rm ele}$	Electronic energy	50 nJ/bit
2	$\xi_{\rm fs}$	Amplifier energy (d^2)	10 pJ/bit/m^2
3	$\xi_{\rm mp}$	Amplifier energy (d^4)	0.0013 pJ/bit/m^4
4	$E_{\rm in}$	Initial energy	0.02 J
5	N	Sensor number	200
6	l (bit)	Data packets size	150
7	$k_{\rm opt}$	Clusters number	8
8	Sink node	Location	(25, 50)
9	Sensor nodes	Square monitoring region	[(0 25), (0 –25), (50 –25), (50 25)]
10	m	Population size	100
11	p_m	Mutation rate	0.01
12	p_c	Crossover rate	0.85
13	t	Termination conditions	400
14	α	Distance adjust parameter	0.6

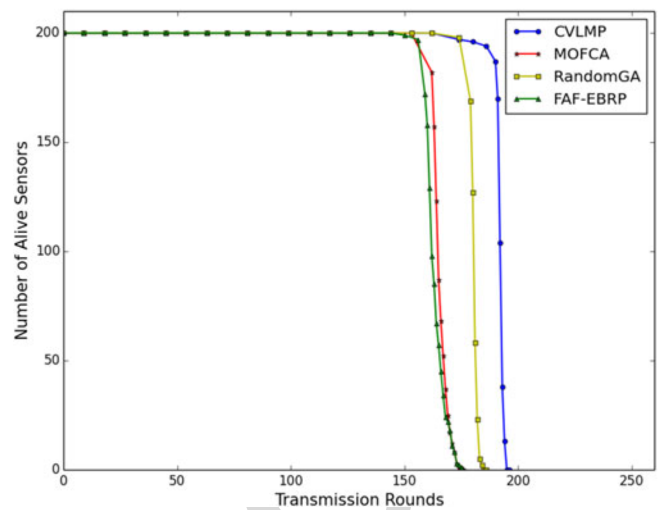


Fig. 7. Comparison of number of alive sensors.

This paper focuses on the protocol optimization of the communication between the sensors and sink nodes. We pick one sink node and 200 sensors in the corresponding small square monitoring region as the simulation objects. In this simulation scenario, the rail is considered as the *X*-axis, and its terminal point, which is far away from the BS, is set as the origin. The parameters used in the simulations are shown in Table III.

Based on (7), (9), and the simulation environment parameters in Table III, we can get $E[d^2_{\rm toSK\it i}] = 3264$ m^2 and initial optimal clusters number is $k_{\rm opt} = 7.75$. Since the number is a positive integer, we set $k = 8$ for the experiments in this paper. Moreover, the residual energy of the sensors declines with the communication. To reduce the load on the CHs and guarantee connectivity of the communication system, we increase the clusters number at the rate of 1 once the energy of the existing CHs is not sufficient to transmit all the information from their own clusters to the sink node. The crossover and mutation rates were experienced from the range of (0.75,0.95) and (0.005,0.02), respectively [28]. In this paper, they are selected as $p_c = 0.85$ and $p_m = 0.01$.

B. Simulation Results and Analysis

For a railway infrastructure wireless monitoring system, four aspects should be considered to ensure the stability and continuity.

1) The system valid lifetime

In this paper, the valid lifetime of the RIWMS is represented by the number of alive sensors and the total residual energy of all sensors after several transmission rounds.

1) The balance of energy consumption across all sensors

Balancing the energy consumption aims to avoid the death of some sensors due to energy exhaustion.

The variance of the remaining energy of all sensor nodes is used to reflect the balance of energy consumption, defined as

$$\mathrm{VRE} = \frac{\sum_{k=1}^{M} \left(E_{{\rm RE}k}(r) - \overline{E}_{\rm RE}(r)\right)^2}{M} \tag{30}$$

where M is the number of sensor nodes in the first monitoring region; $E_{{\rm RE}k}(r)$ is the remaining energy of sensor node k at time t; and $\bar{E}_{\rm RE}(r)$ is the average remaining energy of all sensor nodes.

Furthermore, the inspection data received at the sink node should be another major criterion to support the infrastructure condition estimation and prediction.

The comparisons among the proposed CVLMP, FAF-EBRP [10], and MOFCA [11] are conducted based on the aforementioned four criteria. To ensure the accuracy of the comparisons, the CHs setting and clusters generation for the two protocols are all implemented strictly according to the algorithm flow in [10] and [11]. Moreover, the CVLMP, MOFCA, and FAF-EBRP are all performed under the unique simulation environment (i.e., the railway wireless monitoring system) to guarantee the fair comparison. Additionally, we compare the solution based on the traditional NSGA-II, which generates the initial population randomly.

We compare and justify the performances of the proposed CVLMP with other two protocols FAF-EBRP and MOFCA in four criteria, as shown in Figs. 7–10.

1) From Fig. 7, we see that all sensors died after 190 transmission rounds using the CVLMP, whereas using the FAF-EBRP, the sensors death begins at 150 rounds and decreases quickly to 0 by 175 rounds. Using the MOFCA, the tendency of sensors death is similar to but a bit more than using the FAF-EBRP. The efficiency of the protocols is compared in Fig. 8 in terms of the total residual energy. Using CVLMP, FAF-EBRP, and MOFCA, the total residual energy declines are all smooth, while the most residual energy is using the CVLMP. Thus, the CVLMP is proved to be the longest-running protocol.
2) The variance of the residual energy in the CVLMP is smaller than those of the other two, as shown in Fig. 9, revealing that the CVLMP remains the most balanced energy consumption of all sensors. The CVLMP can keep all the sensors alive to the utmost and ensure the connectivity of the entire network.

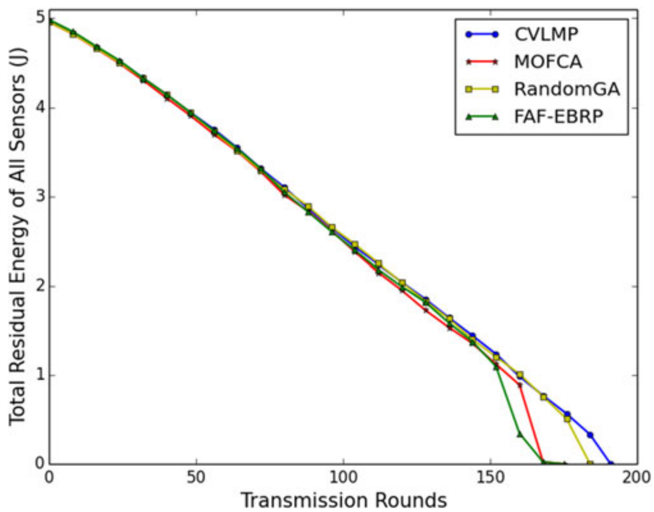


Fig. 8. Comparison of the total residual energy of all sensors.

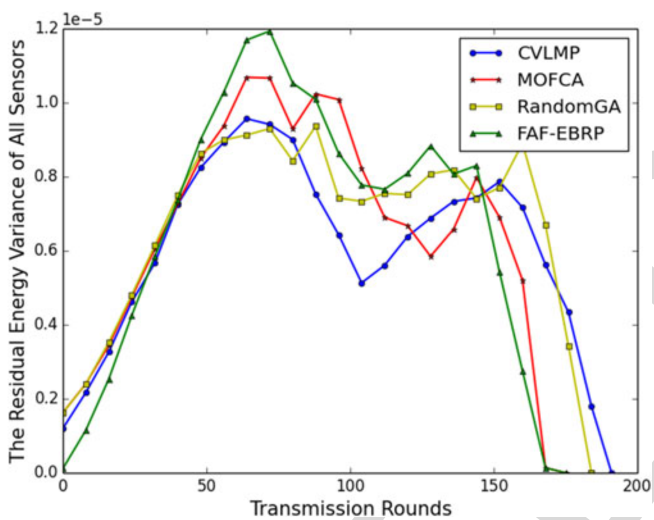


Fig. 9. Comparison of the residual energy variance of all sensors.

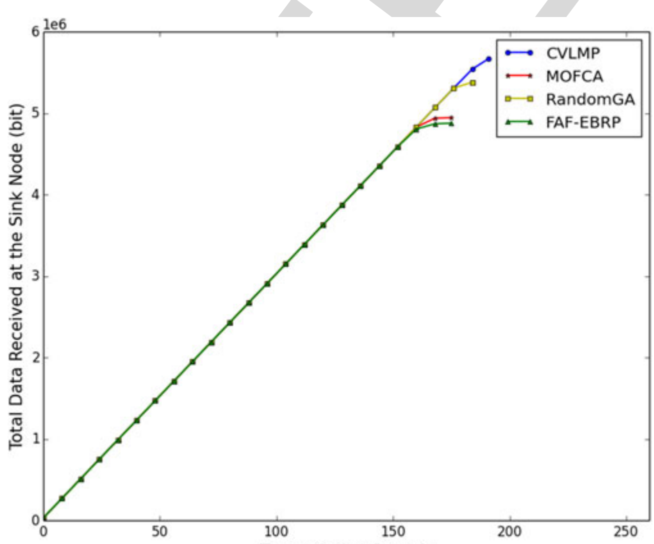


Fig. 10. Comparison of total data received at the sink node.

3) As shown in Fig. 10, the total data received at the sink nodes using the CVLMP are about 1.2 times of that using FAF-EBRP and MOFCA. Therefore, the protocols proposed in this paper render significantly more transmission rounds and received data in comparison with the other two methods, and this is crucial to maintaining the stable railway infrastructure monitoring and condition analysis.
4) Additionally, Figs. 7–10 reveal that the performances of the system using the CVLMP, which initializes the population by K-means++, are better than those using the random initialization population. Moreover, we found that the performances are not stable using the random GA due to the uncertain initialization population with limited iterations. More advanced control and monitoring schemes with robustness need to be studied in the future to optimize the system performance further [29], [30].

VI. Conclusion

Wireless railway infrastructure condition monitoring network is vital to the railway industry. Safe and efficient railway operations require a sufficient lifetime of the sensor network. This paper proposes a novel CVLMP to maximize the lifetime of the monitoring system. The optimization models are used to rotate the CHs and optimize the clusters before each transmission round starts, so as to minimize the total energy consumption and to balance the energy consumption among sensors.

Simulation results demonstrate that the superiority of the proposed CVLMP is threefold.

1) Compared with FAF-EBRP and MOFCA models in the literature, the CVLMP can effectively prolong the lifetime of the monitoring system by 23%, all else being equal.
2) The CVLMP has superior performance in optimally conserving the total residual energy of all the sensors and balancing energy consumption among sensors.
3) The monitoring data received at the sink node are more than those using the other two methods.

References

[1] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, "Wireless sensor networks for condition monitoring in the railway industry: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1088–1106, Jun. 2015.
[2] D. Barke and W. K. Chiu, "Structural health monitoring in the railway industry: A review," *Struct. Health Monit.*, vol. 4, no. 1, pp. 81–93, Mar. 2005.
[3] H. C. Tsai *et al.*, "Railway track inspection based on the vibration response to a scheduled train and the Hilbert–Huang transform," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 229, no. 7, pp. 815–829, Apr. 2015.
[4] K. Chebrolu *et al.*, "Brimon: A sensor network system for railway bridge monitoring," in *Proc. 6th Int. Conf. Mobile Syst., Appl., Serv.*, 2008, pp. 2–14.
[5] E. Berlin and K. V. Laerhoven, "Sensor networks for railway monitoring: Detecting trains from their distributed vibration footprints," in *Proc. IEEE Int. Conf. Distrib. Comput. Sens. Syst.*, 2013, pp. 80–87.
[6] G. M. Shafiullah, S. A. Azad, and A. B. M. Shawkat Ali, "Energy-efficient wireless MAC protocols for railway monitoring applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 649–659, Jun. 2013.
[7] K. Ovsthus and L. M. Kristensen, "An industrial perspective on wireless sensor networks—A survey of requirements, protocols, and challenges," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 3, pp. 1391–1412, Jan. 2014.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, 2000, pp. 1–10.

[9] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[10] D. Zhang, G. Li, K. Zheng, X. Ming, and Z.-H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 766–773, Feb. 2014.

[11] S. A. Sert, H. Bagci, and A. Yazici, "MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks," *Appl. Soft Comput.*, vol. 30, pp. 151–165, May 2015.

[12] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3066–3078, Aug. 2008.

[13] Z. Han, J. Wu, J. Zhang, L. Liu, and K. Tian, "A general self-organized tree-based energy-balance routing protocol for wireless sensor network," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 2, pp. 732–740, Apr. 2014.

[14] G. Han, L. Liu, J. Jiang, L. Shu, and G. Hancke, "Analysis of energy-efficient connected target coverage algorithms for industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 135–143, Feb. 2017.

[15] H. Zhang and C. Liu, "A review on node deployment of wireless sensor network," *Int. J. Comput. Sci. Issues*, vol. 9, no. 6, pp. 378–383, Nov. 2012.

[16] A. Cenedese, M. Luvisotto, and G. Michieletto, "Distributed clustering strategies in industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 228–237, Feb. 2017.

[17] H. Mahboubi, K. Moezzi, A. G. Aghdam, K. Sayrafian-Pour, and V. Marbukh, "Distributed deployment algorithms for improved coverage in a network of wireless mobile sensors," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 163–174, Feb. 2014.

[18] D. Zhang *et al.*, "A new clustering routing method based on PECE for WSN," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, Dec. 2015, Art. no. 162.

[19] F. Ren, J. Zhang, T. He, C. Lin, and S. K. Das Ren, "EBRP: Energy-balanced routing protocol for data gathering in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 12, pp. 2108–2125, Dec. 2011.

[20] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct. 2004.

[21] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach," *Eng. Appl. Artif. Intell.*, vol. 33, pp. 127–140, Aug. 2014.

[22] S. Kurt, H. U. Yildiz, M. Yigit, B. Tavli, and V. C. Gungor, "Packet size optimization in wireless sensor networks for smart grid applications," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2392–2401, Mar. 2017.

[23] F. Khan, "An initial seed selection algorithm for k-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application," *Appl. Soft Comput.*, vol. 12, no. 11, pp. 3698–3700, Nov. 2012.

[24] P. Bholowalia and A. Kumar, "EBK-means: A clustering technique based on elbow method and k-means in WSN," *Int. J. Comput. Appl.*, vol. 105, no. 9, pp. 17–24, Nov. 2014.

[25] D. Wei, Y. Jin, S. Vural, K. Moessner, and R. Tafazolli, "An energy-efficient clustering solution for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3973–3983, Nov. 2011.

[26] E. A. Zimichev, N. L. Kazanskii, and P. G. Serafimovich, "Spectral-spatial classification with k-means++ particional clustering," *Comput. Opt.*, vol. 38, no. 2, pp. 281–286, Jul. 2014.

[27] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliab. Eng. Syst. Safety*, vol. 91, no. 9, pp. 992–1007, Sep. 2006.

[28] Z. J. Wang and H. Hu, "Improved precast production–scheduling model considering the whole supply chain," *J. Comput. Civil Eng.*, vol. 31, no. 4, Jul. 2017, Art. no. 04017013.

[29] P. Li, R. X. Li, Y. Cao, D. Y. Li, and G. Xie, "Multi-objective sizing optimization for island microgrids using triangular aggregation model and levy-harmony algorithm," *IEEE Trans. Ind. Informat.*, Nov. 2017, to be published. Doi: 10.1109/TII.2017.2778079.

[30] P. Li, R. Dargaville, Y. Cao, D. Y. Li, and J. Xia, "Storage aided system property enhancing and hybrid robust smoothing for large-scale PV systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2871–2879, Nov. 2017.

Xiaoping Ma received the M.S. degree in 2014 from Beijing Jiaotong University, Beijing, China, where he is currently working toward the Ph.D. degree.

He is studying at the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His research interests include traffic information intelligent, sensing, and service technologies.

Honghui Dong received the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007.

He is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. His current research interests include pattern recognition and intelligent systems, as well as transportation science and engineering.

Xiang Liu received the Ph.D. degree from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2013.

He is with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ, USA. His research interests include railway safety and risk management, transportation system analytics and optimization, as well as intelligent railway transportation systems.

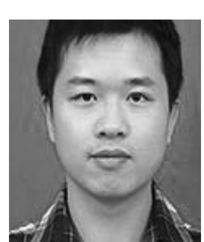
Limin Jia received the Ph.D. degree from China Academy of Railway Sciences, Beijing, China, in 1991.

He is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. His current research interests include safety science and engineering, control science and engineering, transportation engineering, safety technology, and engineering and system science.

Guo Xie received the Ph.D. degree from Nihon University, Tokyo, Japan, in 2013.

He is with the School of Automation and Information Engineering, Xi'an University of Technology, Xi'an, China. His current research interests include safety and reliability of railway systems, intelligent information processing, and fault diagnosis.

Dr. Xie was the recipient of the Japanese Government Scholarship from the Japanese Ministry of Education, Culture, Sports, Science and Technology.



Zheyong Bian received the B.S. and M.S. degrees in industrial engineering from Tianjin University, Tianjin, China, in 2016. He is working toward the Ph.D. degree in civil engineering at Rutgers University, New Brunswick, NJ, USA.

His research interests include operations research, logistics optimization, and freight transportation.

Queries

Q1. Author: Please check whether the edit made to the sentence "Exact analytical methods face difficulty in . . . " retains the intended sense.

Q2. Author: Please complete and update Ref. [29].

Q3. Author: Please provide the areas of study in which Xiaoping Ma received the M.S. degree, and Honghui Dong, Xiang Liu, Limin Jia, and Guo Xie received the Ph.D. degrees.

An Optimal Communications Protocol for Maximizing Lifetime of Railway Infrastructure Wireless Monitoring Network

Xiaoping Ma, Honghui Dong, Xiang Liu, Limin Jia, Guo Xie, and Zheyong Bian

***Abstract*—A wireless monitoring network is an effective way to monitor and transmit information about railway infrastructure conditions. Its lifetime is significantly affected by the energy usage among all sensors. This paper proposes a novel cluster-based valid lifetime maximization protocol (CVLMP) to extend the lifetime of the network. In the CVLMP, the cluster heads (CHs) are selected and rotated with the selection probability and energy information. Then, the clusters are determined around the CHs based on the multi-objective optimization model, which minimizes the total energy consumption and balances the consumption among all CHs. Finally, the multi-objective model is solved by an improved nondominated sorting genetic algorithm II. The simulation results show that, compared with two other strategies in the prior literature, our proposed CVLMP can effectively extend the valid lifetime of the network as well as increase the inspected data packets received at the sink node.**

***Index Terms*—Cluster, energy-efficiency, K-means++, lifetime, nondominated sorting genetic algorithm (NSGA) II, railway monitoring network.**

I. Introduction

With the rapid increase of operating speed and mileage, operational safety of railways has attracted wide attention from both academia and industry. A large portion of railway accidents is caused by infrastructure failures. Thus, rail infrastructure inspections are an indispensable solution to obtaining related information regarding infrastructure conditions (e.g., rail defects, track geometry failure) to ensure railway safety [1], [2]. The track inspection vehicle has been a prevalent technology for railway infrastructure inspection. However, track inspection vehicles are usually used to periodically inspect railway infrastructure and can hardly provide real-time information regarding railway conditions [3] by detecting emergency conditions. Moreover, frequent inspection requires intensive human resources and high capital cost. As a supplementary approach for periodic track inspection, industrial wireless sensor networks (IWSNs) emerge as a preferable technology to inspect the operational condition of the railway infrastructure [4]–[6]. In addition, in order to identify infrastructure defects quickly by means of monitoring the system in real time, IWSNs appear to be an efficient method for monitoring railway infrastructures, especially for complex, extensive railway transportation systems, such as the high-speed rail networks.

In IWSNs, a large quantity of sensors is installed on railway infrastructures (e.g., rails, roadbeds, and bridges) to monitor and transmit condition information [7]. The efficiency of IWSNs greatly depends on the availability and reliability of these sensor nodes. However, the processing ability, communication bandwidth, and energy storage of these sensor nodes are often limited. Therefore, an effective communication protocol is needed to utilize the limited sensing resources. Previous literature demonstrated that the cluster-based routing protocol is one of the energy-saving strategies [8]–[20] that minimizes the total energy consumption and/or balances the energy consumption among all sensors.

To optimally utilize IWSNs, this paper proposes a novel cluster-based valid lifetime maximization protocol (CVLMP), which generates the optimal global clusters dynamically in each transmission round. In the CVLMP, the K-means++ algorithm is used to generate initial clusters based on the deployment of the sensors, and then, the initialized clusters are optimized dynamically in the following rounds. During the optimization process, the cluster heads (CHs) are dynamically selected and rotated from all sensors in the monitoring area to balance their energy consumption. To achieve the target, each CH's rotation is performed based on the probability of CH candidacy, the ratio of residual energy, and predicted energy consumption. Then, the genetic algorithm (GA) is used to optimize clusters with the objective of minimizing the total energy consumption and

Manuscript received June 13, 2017; revised September 7, 2017 and October 24, 2017; accepted November 23, 2017. This work was supported in part by the National Science and Technology Major Project under Grant 2016YFB1200100, in part by the State Key Laboratory under Grant RCS2016ZT018, and in part by the National Natural Science Foundation of China under Grant U1534208 and Grant 51207007. Paper no. TII-17-1257. *(Corresponding author: Limin Jia.)*

X. Ma is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, and also with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ 08817 USA (e-mail: xm54@scarletmail.rutgers.edu).

H. Dong and L. Jia are with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: hhdong@bjtu.edu.cn; jialm@vip.sina.com).

X. Liu and Z. Bian are with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ 08817 USA (e-mail: xiang.liu@rutgers.edu; zheyongbian@gmail.com).

G. Xie is with Shaanxi Key Laboratory of Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an 710048, China (e-mail: guoxie@xaut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2017.2785786

balancing the consumption of each sensor. Experimental simulations are designed to test the performance of the proposed CVLMP. The experimental results show that the CVLMP outperforms the typically used approaches, including the energy-balanced routing method based on forward-aware factor (FAF-EBRM) and multi-objective fuzzy clustering algorithm (MOFCA) [10], [11].

The rest of the paper is organized as follows. Section II reviews the related literature and formulates the problem. The overall scheme of the energy-saving protocol is described in Section III. The novel CVLMP is presented in Section IV. Experimental simulations are analyzed in detail in Section V, and Section VI presents the key conclusions.

II. Related Prior Work and Problem Formulation

As previously mentioned, IWSNs have been widely applied in railway transportation systems, especially in infrastructure condition monitoring, to improve operational safety. IWSNs transmit real-time monitoring information for remote fault diagnosis of critical parts of bridges, tunnels, trains, and rails [4], [5]. However, most existing studies focus on the technological applications of IWSNs in railway systems, whereas the reliability and availability of IWSN information transmission have not received adequate attention. To address this problem, Shafiullah optimized the energy consumption among the sensors installed in the rail cars [6], given a relatively small sensor network. The energy efficiency optimization in a large network with many sensors in railway monitoring systems has rarely been studied, while the number of sensors continues to increase, owing to their wide application in railway inspection.

Some studies focus on improving the energy efficiency of IWSNs using optimization methods. For example, transmission power adjustment technologies help to avoid the interference by controlling the signal coverage area and thus reduce the energy consumption [12]–[14]. Optimizing the deployment of the sensor nodes is useful to reduce the energy consumption as well [15], [17]. The consumption of limited energy is reduced by optimizing the transmitted data packets size [18]. Besides, hierarchization (clustering) is an ideal solution to reducing and balancing the energy consumption. To this end, CH selection/rotation and cluster generation/optimization schemes are two major procedures to establish such a protocol.

For the CH selection and rotation scheme, some existing methods prescribe that those nodes that have not been regarded as CHs in previous rounds might have a higher probability to be selected as CHs [8], [9]. Zhang proposed an FAF-EBRM algorithm for CH selection and rotation scheme using a multihop strategy to balance the energy consumption [10]. In this strategy, the next-hop node is determined by the link weight and forward energy density. This method aims to minimize wasted energy caused by the backward transmission of information. The position of the sensors is considered in the MOFCA [11]. Similarly, the residual energy and distances between sensor nodes are considered in CH selection in predictive energy consumption efficiency, energy-balanced routing protocol (EBRP), and hybrid, energy-efficient, distributed clustering approach [18]–[20]. Kuila applied the particle swarm optimization heuristic algorithm to solve similar problems [21]. However, in these approaches, CH selection and rotation are achieved within each local cluster, which is determined in the initial stage. This may result in a local optimization rather than the global optimization for CH selection and rotation.

For cluster generation, Heinzelman proposed two clustering strategies, LEACH and LEACH-centralized [8], [9], which divide all sensor nodes evenly into k-clusters. Cenedese proposed a distributed clustering strategy to generate the clusters [22]. Khan used the K-means method to classify the sensor nodes according to their relative positions [23]. However, these methods did not consider the impact of cluster size on CH energy consumption. Thus, some improved clustering methods are developed by determining the size of each cluster based on the distances from the sensor nodes to the sink node [24], [25]. However, in those studies, the cluster size remains constant once the clusters are determined, and thus may result in unbalanced sensor node energy consumption in a dynamic operational environment.

To solve the above-mentioned problems, an optimal cluster-based protocol is proposed in this paper for the RIWMS. The lifetime of the system is maximized by minimizing the total energy consumption and balancing the consumption among all sensors simultaneously. The intended contributions of this study can be summarized as follows.

1) The impact factors are comprehensively considered in the modeling of CHs selection and rotation, including the predicted energy consumption of the tentative CHs, the candidacy probability, together with the residual energy of sensors. This addresses the limited consideration of energy consumption after CHs selection in existing studies, which can effectively reduce the selected probability as CHs for incompetent sensors.
2) The cluster scales and the CH-to-members correspondence are dynamically optimized at the beginning of each transmission round, simultaneously considering both the energy consumption minimization and balance, which expands the single focus in previous studies.
3) A hybrid algorithm by blending K-means++ with a nondominated sorting genetic algorithm II (NSGA-II) is developed to achieve the multi-objective optimization model. Using K-means++ can improve the population initialization of NSGA-II so as to accelerate the optimization progress and enhance the quality of the solutions.
4) The proposed CVLMP model enhances the utilization efficiency of the limited energy, which satisfies the real-time and consecutive inspection requirements of the railway wireless monitoring system. This will promote its application in the railway monitoring system and increase the safety of the railway operation.

III. Overall Scheme of the Energy-Efficiency Protocol

A. Overall Structure

In RIWMS, the information transmission network is composed of three layers, as shown in Fig. 1, sensor layer

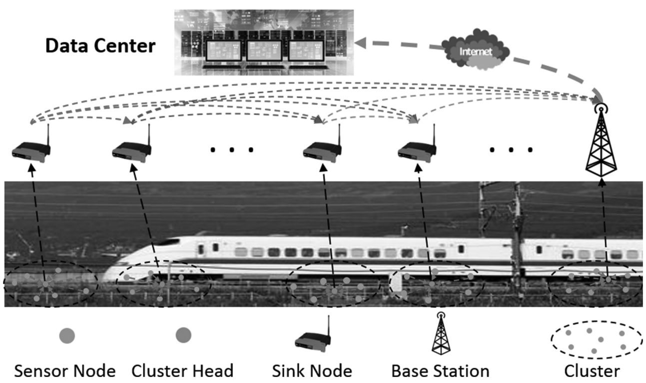


Fig. 1. Schematic structure of the RIWMS.

(infrastructure condition monitoring), communication layer (information transmission), and data processing layer (fault diagnosis and prediction). Since the base station (BS) is located at one end of the monitoring region, it is hard for the sensors to send information to the BS directly. To address this problem, the monitoring region is divided into several small regions, and the sink node is equipped to collect the information from each small region locally. The information transmission progress is shown as follows. First, the information inspected by the sensors is sent to CHs and then forwarded and collected at the local sink nodes. Then, the collected information is transmitted among the sink nodes and forwarded to the BS. Finally, the infrastructure information arrives at the data center from BS via internet. The sink nodes, whose energy storage, computing, and communication ability are all superior to the sensors, act as the local BS in RIWMS to collect the information from the nearby sensors.

B. Overall Framework

We adopt and optimize a cluster-based protocol to reduce the energy consumption in the information transmission between the sensors and sink nodes in the small monitoring region. In this protocol, all the sensors in the small monitoring region are divided into several clusters. One of the sensors in each cluster is selected as the CH to receive the information from other non-CH sensors. Due to the small scale of the monitoring region and high real-time demands, the single-hop protocol is adopted in the communication between the sensors and the corresponding sink node.

As shown in Fig. 2, the scheme is composed of three layers: the input information layer, cluster optimization strategy layer, and output information layer.

In the input layer, the input information, including the energy and location information, is centered around the sensors deployed in the monitoring region.

In the strategy layer, the CHs and clusters are generated with the objective of minimizing the total energy consumption and balancing the consumption among all sensors, and the approach will ultimately extend the RIWMS lifetime. First, the number of the clusters, which is essential to the K-means++, is calculated in advance. Then, the K-means++ algorithm is adopted to initialize the clusters and CHs. The basic idea of this approach

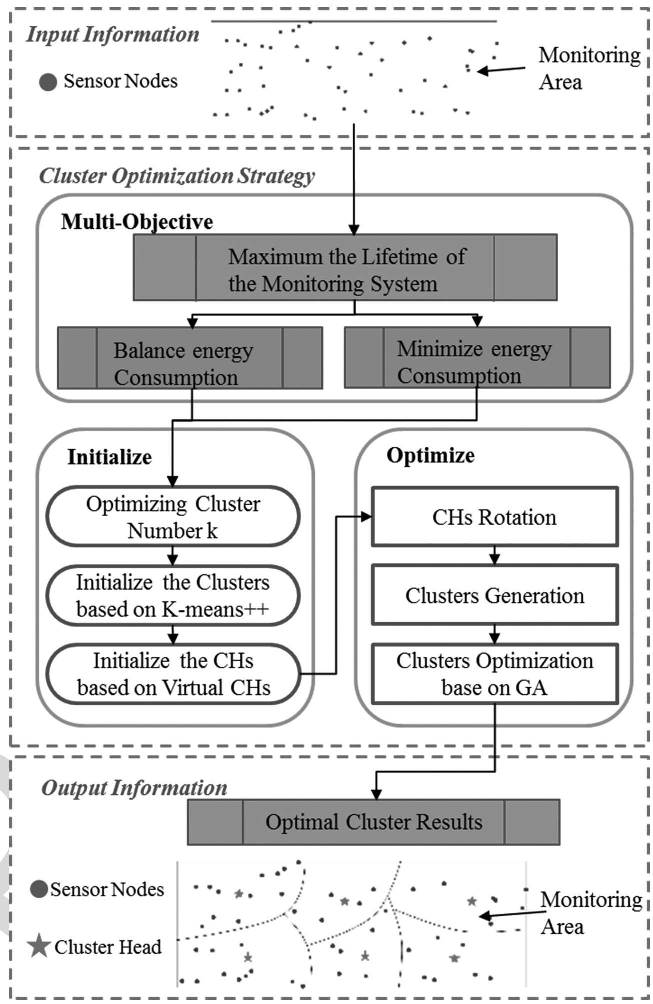


Fig. 2. Overall scheme of RIWMS sensor cluster optimization.

is to cluster the sensors that are close to each other in order to reduce the energy consumption. However, for the complex RIWMS, the CHs and clusters generated by the K-means++ are not necessarily the optimal solutions, and thus, we develop a more effective method in which clusters and CHs are optimized in the following steps. In each optimization round, the CHs are rotated and selected at first, and then, clusters are generated around them. The candidacy probability, residual energy, and predicted energy consumption of all sensors are all considered to determine the CHs. Furthermore, the clusters are obtained by solving a multi-objective optimization model using the GA.

In the output layer, the sensors are allocated to the clusters according to the optimal solutions. The transmission of the inspected information based on the optimal solutions is expected to extend the lifetime of the RIWMS significantly.

IV. Proposed Cluster-Based Valid Lifetime Maximization Protocol

The CVLMP is proposed and introduced in detail in this section, including cluster initialization, CH rotation and selection, and cluster optimization by solving a multi-objective model

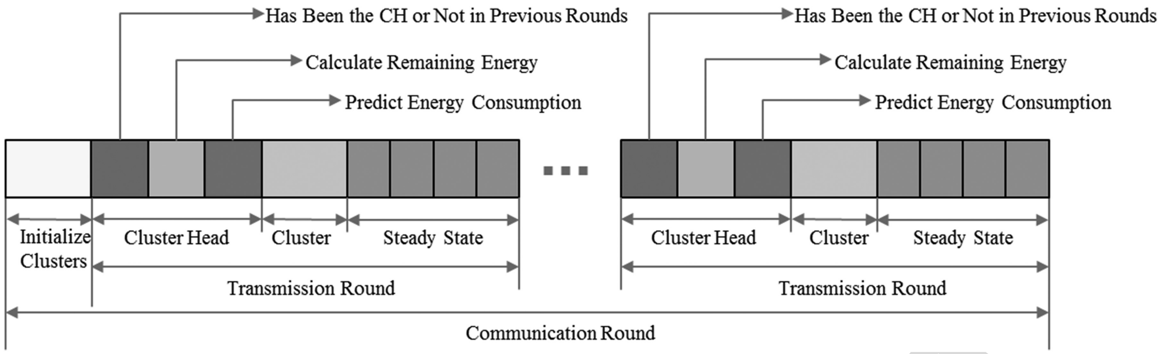


Fig. 3. Structure of the communication round of the proposed CVLMP.

using the NSGA-II. The operational process of the proposed CVLMP is shown in Fig. 3.

The protocol is executed by several communication rounds, including one initialization clustering round and several transmission rounds. In the initialization round, all sensor nodes are divided into several clusters using K-means++. In the transmission rounds, three phases (CH selection and rotation, cluster generation and optimization, and steady transmission) are executed repeatedly.

A. Clusters Initialization by the K-Means++ Algorithm

In the cluster-based protocol, the energy consumption of non-CHs is determined by the distance to the CH and the size of data packets. In this section, the K-means++ algorithm is adopted to initialize the clusters, and they will accelerate the cluster optimization in the following stages.

1) Optimal Number of Clusters: For the K-means++ algorithm, the number of the clusters k should be determined in advance, as it will affect the total energy consumption of all the N sensors in the small monitoring region. Take a short section of the rail and the corresponding infrastructure as monitoring objects; the monitoring field could be considered as a rectangular region. Then, the rectangular region is divided into several small square regions, as mentioned above. The coordinates of the small region are ($x \in [-M/2, M/2]$,$y \in [-(1+\alpha)M, -\alpha M]$). α is the distance coefficient, which is used to adjust the vertical distance between the square region to the sink node. The sink node is located at the origin SK$(0,0)$. The joint probability density of the sensor nodes coordinates is $\rho(x,y) = 1/M^2$.

The total energy consumption of all the sensor nodes in one transmission round is calculated as

$$E_{\text{total}} = \sum_{i=1}^{k} \left(E_{\text{CH}}^{i} + \sum_{j=1}^{n-1} E_{\text{Non-CH}}^{ji} \right) \tag{1}$$

where E_{CH}^{i} is the energy consumed by the ith CHs, $E_{\text{Non-CH}}^{ji}$ represents the energy consumption of the jth sensor node in ith cluster, and n is the number of the sensors in each cluster. We assume that all the N sensors are divided into k clusters equally. Hence, there are $n = N/k$ sensors in each cluster, including one CH sensor and $n-1$ non-CH sensors.

The energy of the CH sensors is consumed in three phases: data packet reception ($E_{\text{CH}-Rx}^{i}$), aggregation ($E_{\text{CH}-Dx}^{i}$), and transmission ($E_{\text{CH}-Sx}^{i}$). The energy consumption of the CHs can be written as

$$\begin{aligned} E_{\text{CH}}^{i} &= E_{\text{CH}-Rx}^{i} + E_{\text{CH}-Dx}^{i} + E_{\text{CH}-Sx}^{i} \\ &= (n-1) \times l \times E_{\text{ele}} + n \times l \times E_{\text{DA}} \\ &\quad + (l \times E_{\text{ele}} + l \times \xi_{\text{mp}} \times d_{\text{toSK}i}^{4}) \\ &= l \times (n \times E_{\text{ele}} + n \times E_{\text{DA}} + \xi_{\text{mp}} \times d_{\text{toSK}i}^{4}) \end{aligned} \tag{2}$$

where E_{ele} denotes the electronics energy coefficient, ξ_{mp} is the amplifier energy coefficients for the multipath fading model, l is the size of data packet, and $d_{\text{toSK}i}$ is the distance from the ith CH sensor to the sink node. It is calculated as

$$d_{\text{toSK}i} = \sqrt{x(i)^2 + y(i)^2}. \tag{3}$$

The non-CHs just transmit their data packets to the corresponding CHs, and the energy consumption is calculated as

$$E_{\text{Non-CH}}^{ji} = l * E_{\text{ele}} + l * \xi_{\text{fs}} * d_{\text{toCH}ji}^{2} \tag{4}$$

where ξ_{fs} is the amplifier energy coefficients for the free space model, and $d_{\text{toCH}ji}$ represents the distance from the jth sensor to the CH in the ith cluster, which is calculated as

$$d_{\text{toCH}ji} = \sqrt{(x(j) - x(i))^2 + (y(j) - y(i))^2}. \tag{5}$$

Statistically, the opportunity of each sensor to be the CH is equal. The expected total energy consumption, as shown in (1), can be rewritten as

$$\begin{aligned} E[E_{\text{total}}] &\approx k * E(E_{\text{CH}}^{i}) + N * E[E_{\text{Non-CH}}^{ji}] \\ &= l * (2N * E_{\text{ele}} + N * E_{\text{DA}} + k * \xi_{\text{mp}} * E[d_{\text{toSK}i}^{4}] \\ &\quad + N * \xi_{\text{fs}} * E[d_{\text{toCH}ji}^{2}]). \end{aligned} \tag{6}$$

The expected fourth (4th) power of $d^4_{\mathrm{toSK}i}$ is calculated as

$$\begin{aligned} E[d^4_{\mathrm{toSK}i}] &= \int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\left(\sqrt{x^2+y^2}\right)^4\rho(x,y)dxdy \\ &= \int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\left(x^2+y^2\right)^2\rho(x,y)dxdy \\ &= \left(0.0125+\frac{(\alpha+1)^3-\alpha^3}{18}+\frac{(\alpha+1)^5-\alpha^5}{5}\right)*M^4. \end{aligned} \tag{7}$$

The expected of $d^2_{\mathrm{toCH}ji}$ is expressed as

$$\begin{aligned} E[d^2_{\mathrm{toCH}ji}] &= \frac{1}{2}\int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2}\int_{-(\alpha+1)M}^{-\alpha M}\int_{-M/2}^{M/2} \\ &\quad \times\left(\sqrt{(x_j-x_i)^2+(y_j-y_i)^2}\right)^2 \\ &\quad \times \rho(x_j,x_i,y_j,y_i)dx_jdx_idy_jdy_i \\ &= \frac{M^2}{6k}. \end{aligned} \tag{8}$$

The initial optimal number of clusters is calculated by setting the derivative of $E[E_{\mathrm{total}}]$ in (6), with respect to k, to zero, as

$$k_{\mathrm{opt}} = \sqrt{\frac{\xi_{\mathrm{fs}}*N*M^2}{6\xi_{\mathrm{mp}}*d^4_{\mathrm{toSK}i}}}. \tag{9}$$

The optimal number of clusters k_{opt} is calculated by (9) when the energy of sensors is sufficient to act as CHs. However, the residual energy of the sensors will decrease as the communication travels. Once the residual energy is not enough to transmit all the information to the sink node as a CH in the current cluster scale, the number of clusters will increase to reduce the sensors in one cluster and decrease the load on the CHs. In this manner, the energy consumption of each CH will be reduced, and the information transmission will be resumed. The strategy is helpful to extend the system lifetime.

2) Initialize Clusters Based on K-Means++: After the number of clusters k is determined, the K-means++ is performed to initialize the clusters. In this algorithm, the selected initial cluster centers are spread out as much as possible [27], ensuring the dispersion of the clusters. The algorithmic details of the K-means++ are shown in Table I.

In step 4, the sensors far from the selected cluster centers will be selected as the new cluster center with high probability. In step 5, all the sensors will be divided into the k_{opt} clusters based on the selected cluster centers.

3) Initialize CHs Based on K-Means++: As shown in Fig. 4, the initial CHs are selected from the k clusters. The virtual CHs (the center points of clusters) are calculated by (10). Then, the sensors nearest to each virtual CH are selected as the initial CHs as follows:

$$\text{Virtual CHs } (X(k),Y(k)) = \left(\frac{\sum_{i=1}^{n_k}x(i)}{n_k},\frac{\sum_{i=1}^{n_k}y(i)}{n_k}\right) \tag{10}$$

TABLE I
CLUSTER INITIALIZATION BASED ON K-MEANS++

K-means++ Algorithm Cluster Initialization
Input: The N sensors deployed in a square monitoring region;
Output: The N sensors are divided into k_{opt} clusters;
Step 1: determine the number of the clusters k_{opt} based on (13).
Step 2: select one sensor as the first cluster center (CC_1).
Step 3: calculate the distance between sensors and selected cluster centers;
repeat
for $i = 1,2,\ldots,k$; cluster centers
for $j = 1,2,\ldots,N-k$; cluster members
$d_{\mathrm{toCC}ji} = \sqrt{(x(j)-x(i))^2+(y(j)-y(i))^2}$; calculate distance
until all the distances are calculated.
Step 4: select the kth cluster center (CC_k, $1 \le k \le k_{\mathrm{opt}}$)
repeat
for $j = 1,2,\ldots,N-k$
for $i = 1,2,\ldots,k$
$d_{\mathrm{toCC}j} = \min(d_{\mathrm{toCC}ji})$; select the minimum distance;
$\mathrm{sum}(d(x_{\mathrm{toCH}j})) = \sum_{j=1}^{N-k} d_{\mathrm{toCC}_j}$
$p_j = d(x_{\mathrm{toCH}j})/\mathrm{sum}(d(x_{\mathrm{toCH}j}))$; define the selection
probability of the cluster centers
until all the k_{opt} cluster centers are selected.
Step 5: Generate the initial clusters based on the k_{opt} CHs.
repeat
for $i = 1,2,\ldots,k_{\mathrm{opt}}$
for $j = 1,2,\ldots,N-k_{\mathrm{opt}}$
$d_{\mathrm{toCC}ji} = \sqrt{(x(j)-x(i))^2+(y(j)-y(i))^2}$
for $j = 1,2,\ldots,N-k_{\mathrm{opt}}$
for $i = 1,2,\ldots,k_{\mathrm{opt}}$
$d_{\mathrm{toCC}j} = \min(d_{\mathrm{toCC}ji})$
dividing the jth sensor into the ith cluster;
until all the sensors are grouped into the corresponding clusters.

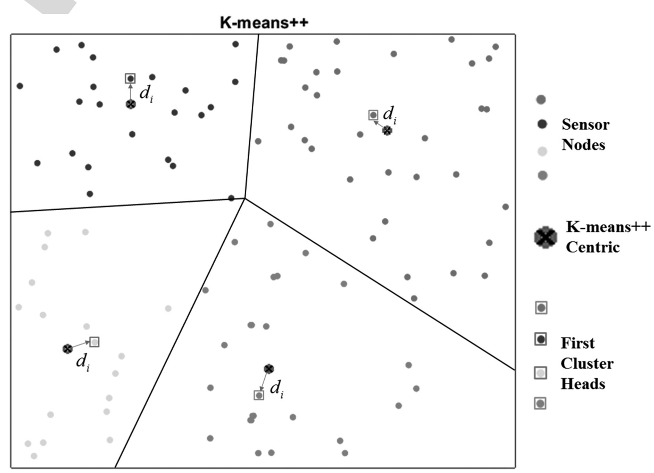


Fig. 4. Cluster initialization based on K-means++.

where n_k is the number of sensors in the kth cluster.

The initial clusters and CHs obtained by K-means guarantee that the sensors close to each other are divided into the same cluster, which can reduce the energy consumption of non-CHs. However, the size of the clusters is unpredictable, which is related to the balance of CHs energy consumption. To prolong the lifetime of the RIWMS, the initial CHs and clusters should be further optimized. The K-means++ method obtaining the initial CHs and clusters is not necessarily able to fully utilize the limited energy. They will be used as the initial clusters in the following cluster optimization.

B. CH Selection and Rotation Probability Model

In the cluster-based protocol, the energy consumption of CH is much more than that of non-CH sensors. Selecting and rotating the CHs before each transmission round starts will help to balance the energy among all sensors. In the CH selection and rotation, in addition to the candidacy probability and the sensors residual energy, the predicted energy consumption after the election is considered. This model could reduce the CH's selected probability of incompetent sensors.

1) Probability of CHs Candidacy: As described above, to balance the energy consumption among all sensors, each sensor should have a completely equal opportunity to operate as CH. This means that the N/k sensors in each cluster will have one chance to act as the CH in the following N/k rounds [8], [9].

The probability of CH candidacy considers whether the sensor is selected as CH in the last $r-1$ rounds. The probability is defined as [8]

$$p_{1i}(r) = \begin{cases} \frac{k}{N-k*r}, & C_i(r)=1 \\ 0, & C_i(r)=0 \end{cases} \tag{11}$$

where $r \in [1, \lfloor N/k \rfloor]$ means that when all sensor nodes have rotated as CHs in previous rounds, r will be reset to 1 and increase to $\lfloor N/k \rfloor$ in each subsequent round. $\lfloor N/k \rfloor$ means that r will always round down to the nearest integer. $C_i(r)$ represents the condition of the node i. If the node has been the CH in the previous r rounds, we set $C_i(r)=0$ and vice versa.

2) Ratio of Residual Energy: The residual energy of the sensors will vary after several transmission rounds, since

1) the energy consumptions of CHs and non-CHs are different;
2) the energy consumption of different CHs varies due to the different cluster scales and transmission distances; and
3) the different distances from non-CHs to the CHs will also lead to the different energy consumptions of non-CHs.

Therefore, in the CH selection and rotation phases, a higher ratio of residual energy yields a greater probability of CH selection. The ratio of the residual energy is defined as

$$p_{2i}(r) = \frac{\left|E^i_{\mathrm{Re}}(r) - E^{\min}_{\mathrm{Re}}(r)\right|}{\sum_{i=1}^{n}\left|E^i_{\mathrm{Re}}(r) - E^{\min}_{\mathrm{Re}}(r)\right|} \tag{12}$$

where $E^i_{\mathrm{Re}}(r)$ represents the residual energy of the ith sensor before the rth round, and $E^{\min}_{\mathrm{Re}}(r)$ is the minimum residual energy of all sensors in the cluster. Furthermore, the residual energy of the ith sensor before the rth round is computed as

$$E^i_{\mathrm{Re}}(r) = E^i_{\mathrm{Re}}(r-1) - E^i_{\mathrm{Co}}(r-1) \tag{13}$$

where $E^i_{\mathrm{Co}}(r-1)$ is the energy consumed by the ith sensor in the $r-1$th round. If the sensor is CH in the $r-1$th round, it is computed according to

$$E_{\mathrm{Co-CH}}(r-1) = E_{Rx}(r-1) + E_{Dx}(r-1) + E_{Tx}(r-1), \tag{14}$$

$$E_{Rx}(r-1) = (n_k - 1) * l * E_{\mathrm{ele}}, \tag{15}$$

$$E_{Dx}(r-1) = n_k * l * E_{\mathrm{DA}}, \tag{16}$$

$$E_{Tx}(r-1) = l * E_{\mathrm{ele}} + l * \xi_{\mathrm{mp}} * d^4_{\mathrm{toSK}}(r-1). \tag{17}$$

If the sensor is non-CH in the $r-1$th round, it is computed by

$$E_{\mathrm{Co-NCH}}(r-1) = l * E_{\mathrm{ele}} + l * \xi_{\mathrm{fs}} * d^2_{\mathrm{toCH}}(r). \tag{18}$$

3) Ratio of the Predicted Energy Consumption: The predicted energy consumption of the CHs after the election is another essential factor. The sensors with less energy consumption as CHs will be selected in the following transmission round with high probability. The ratio of the predicted energy consumption in the following $r+1$th round is defined as

$$p_{3i}(r) = \frac{\left|E^i_{\mathrm{Co}}(r+1) - E^{\max}_{\mathrm{Co}}(r+1)\right|}{\sum_{i=1}^{n_k}\left|E^i_{\mathrm{Co}}(r+1) - E^{\max}_{\mathrm{Co}}(r+1)\right|} \tag{19}$$

where $E^i_{\mathrm{Co}}(r+1)$ represents the predicted energy consumption of the ith sensor in the following $r+1$th round while it is selected as CH. $E^{\max}_{\mathrm{Co}}(r+1)$ is the maximum predicted consumption of the sensors in the cluster. The predicted energy consumption can be calculated by (14)–(17).

The above-mentioned three factors are all considered in the CH selection and rotation probability model. The probability is defined as

$$\begin{aligned} & p_i(r) = \omega_1 p_{1i}(r) + \omega_2 p_{2i}(r) + \omega_3 p_{3i}(r) \\ & \text{S.T.} \quad \sum_{i=1}^{3} \omega_i = 1 \\ & 0 \le p_{1i}(r), p_{2i}(r), p_{3i}(r) \le 1 \end{aligned} \tag{20}$$

where $\omega_i, i = 1, 2, 3$ are the weighting coefficients, and they are used to adjust the importance of each factor to the model.

The most appropriate CHs are selected and rotated based on this model before each transmission round starts. They can be used to support the cluster generation and optimization in the following phases.

C. Cluster Generation and Optimization Probability Model

After the CHs are determined, the clusters will be updated accordingly to optimize the communication protocol and extend the system lifetime.

For the RIWMS, the death of any sensor may potentially lead to system instability or inspection failures. In this paper, the valid lifetime of the system is defined as the time when more than 90% of the sensors are alive. Hence, the optimal clusters should guarantee all sensors remain alive as long as possible.

The optimal clusters are generated based on the following two schemes. 1) Adjust the clusters scales to balance the energy consumption among CHs. 2) Optimize the correspondence among the CHs and non-CHs to minimize the total energy consumption.

1) Scale of Clusters: As mentioned before, the energy consumption of the CHs is related to d_{toSK} and the scale of clusters (number of non-CHs). d_{toSK} changes as the CHs rotate, and hence, the scales of clusters should vary correspondingly to balance the energy consumption among the CHs. The optimization model is defined as

$$f_1(E_{\mathrm{CH}i}(n_i)) = \min \frac{\sum_{i=1}^{k} \left(E_{\mathrm{CH}i}(n_i) - \overline{E}_{\mathrm{CH}}\right)^2}{k} \quad (21)$$

where $f_1(E_{\mathrm{CH}i}(n_i))$ aims to minimize the variance of the CHs energy consumption, which represents the energy consumption balance degree among all CHs. $E_{\mathrm{CH}i}$ represents the energy consumption of the CH in the ith cluster; $\overline{E}_{\mathrm{CH}}$ represents the average energy consumption for all CHs; and k is the number of clusters.

The energy consumed by the CH in the ith cluster is calculated based on (1), as

$$E_{\mathrm{CH}i}(n_i) = l * (n_i * E_{\mathrm{ele}} + n_i * E_{\mathrm{DA}} + \xi_{\mathrm{mp}} * d^4_{\mathrm{toSK}i})$$

$$\text{S.T.} \quad \sum_{i=1}^{k} n_i \leq N \quad (22)$$

where $n_i (i = 1, 2, \ldots, k)$ is the number of sensors in the ith cluster. The constraint implies that only the living sensors participate in the clusters generation and optimization.

The average energy consumption of all the CHs is calculated as

$$\bar{E}_{\mathrm{CH}} = \frac{\sum_{i=1}^{k} E_{\mathrm{CH}i}(n_i)}{k}. \quad (23)$$

2) CHs to Non-CHs Correspondence: Based on the above-mentioned model, the CHs have almost identical energy consumptions, while the energy consumption of each non-CHs sensor would vary while divided into different clusters. In the following model, the correspondences between the CHs and non-CHs will be optimized to minimize the energy consumption of non-CHs. The total energy consumption optimization model is given by

$$f_2(E_{\mathrm{NCH}i}, E_{\mathrm{CH}j}) = \min \left(\sum_{i=1}^{N-k} E_{\mathrm{NCH}i} + \sum_{j=1}^{k} E_{\mathrm{CH}i} \right), \quad (24)$$

$$E_{\mathrm{NCH}i} = l * E_{\mathrm{ele}} + l * \xi_{\mathrm{fs}} * \sum_j r_{ij} * d^2_{\mathrm{toCH}j}, \quad (25)$$

$$r_{ij} = \begin{cases} 1, \text{ sensor } i \text{ is allocated to cluster } j \\ 0, \text{ otherwise} \end{cases} \quad (26)$$

where $f_2(E_{\mathrm{NCH}i}, E_{\mathrm{CH}j})$ aims to minimize the total energy consumption of all sensors. $E_{\mathrm{NCH}i}$ is the energy consumption of the ith non-CH sensor and $E_{\mathrm{CH}j}$ is the energy consumption of the CH in the jth cluster. r_{ij} is the decision variable.

The scales of the clusters are optimized to balance the energy consumption among the CHs, while the total energy consumption is minimized by optimizing the correspondence between the CHs and non-CHs members. The optimization of the correspondence between the CHs and non-CHs will reversely affect the cluster scales. It is obvious that the cluster generation is a multi-objective optimization problem. It can be formulated as

$$\min f(E) = (f_1(E_{\mathrm{CH}i}), f_2(E_{\mathrm{CH}i}, E_{\mathrm{NCH}i})). \quad (27)$$

The following two constraints should be considered: 1) there are at least two nodes in each cluster, i.e., $2 \leq n_i \leq N$; and 2) the energy-exhausted, "dead" sensors should be excluded from cluster generation, i.e., $\sum_{i=1}^{k} n_i \leq N$.

The cluster generation and optimization is an NP-hard problem. Exact analytical methods face difficulty in obtaining the optimal solutions when the scale of the problem is large. Heuristic methods such as the GA are effective to solve such multi-objective optimization problems in practice. Q1

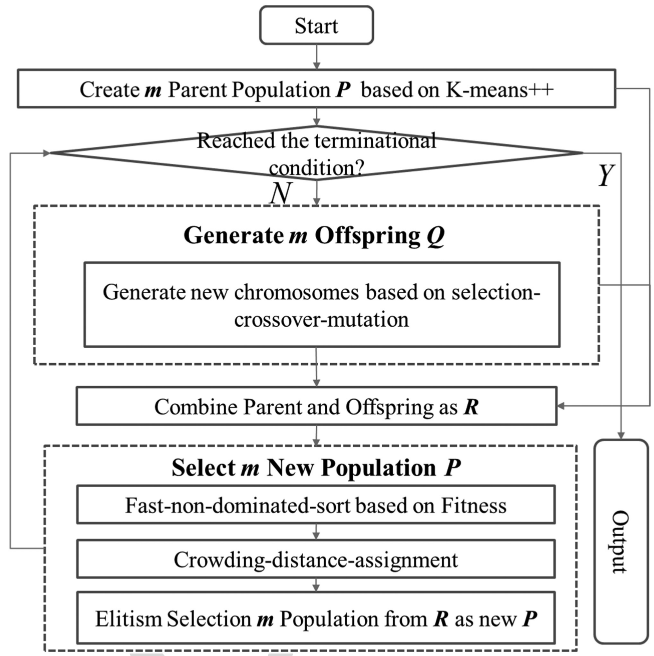


Fig. 5. Flowchart for cluster optimization using NSGA-II.

D. Cluster Optimization Based on the NSGA-II

The multi-objective optimization algorithm is adopted to optimize the clusters to minimize total energy consumption and balance the consumption among all CHs. The NSGA-II is one of the effective GA methods to solve multiple objective optimization problems [26]. The optimization process is described in Fig. 5. The optimal clusters can be obtained by the NSGA-II.

1) Population Initialization: The optimization process based on the NSGA-II is intended to obtain the best solution (chromosomes) based on the initial population. The population consists of m chromosomes. Each of them is composed of M genes, and their positions and contents denote the sensor index and cluster number [see Fig. 6(a)]. The sensor index $[1, 2, \ldots, N]$ will be assigned to sensors when they are deployed in the monitoring region. The cluster number $[1, 2, \ldots, k]$ represents which cluster the sensors belong to.

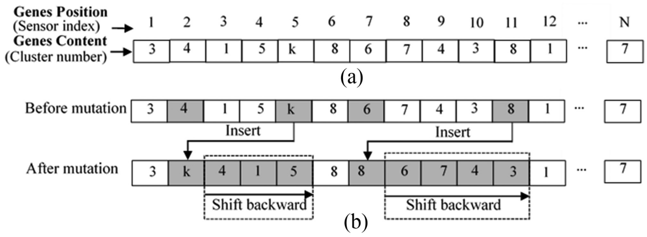


Fig. 6. Representation and mutation: (a) Chromosomal representation; and (b) shift mutation.

The initialization of the population would have a substantial influence on the cluster optimization speed and results, because all the chromosomes are generated based on the initial ones in the parent population. The initialized clusters based on K-means++ in this paper can provide better first chromosome in the initial parent population P. In the existing studies, the initial chromosome is generally generated randomly, while its genes are randomly scattered in the solution space. In contrast, the genes in the chromosome obtained by K-means++ are more concentrated and have smaller energy consumption. This initial chromosome with higher fitness values (function values) will accelerate the optimization process. Subsequently, the other $m-1$ chromosomes in the parent population are generated using shift mutation methods based on the first chromosome [28]. As shown in Fig. 6(b), two pairs of points are selected at random, then the rear points are inserted ahead of the front point and the front points are all shifted backward. The number of point pairs is determined by the mutation probability p_m.

2) Population Optimization: In the population optimization phase, we try to generate the optimal chromosomes within an acceptable timeframe. The offspring Q with m chromosomes is generated based on the parent population through crossover operation and the roulette wheel selection operation.

The individual selection probability is defined as

$$p(c_i) = \frac{g(c_i)}{\sum_{i=1}^{m} g(c_i)} \tag{28}$$

where c_i is the ith chromosome and $g(c_i)$ is its fitness value, denoted as

$$\max g(c_i) = \left(\frac{1}{f_1(E_{\mathrm{CH}i})}, \frac{1}{f_2(E_{\mathrm{CH}i}, E_{\mathrm{NCH}i})}\right). \tag{29}$$

Theoretically, greater residual energy renders larger probability of selection. The two-point crossover and shift mutation are adopted in the following steps to create the new offspring Q. Subsequently, the offspring Q and the parent population P are combined into a new population R. The NSGA-II method is used to update the parent population P from R. Finally, the best chromosomes in the updated parent population P will be selected as the optimal clusters. The detailed process is shown in Table II.

E. Steady Communication Phase

After the optimal clusters and the most energy-efficient transmission route are determined, the system steps into the steady

TABLE II
STEPS OF THE NSGA-II METHOD

NSGA-II Algorithm Population Optimization

Input: Algorithm parameters, the population scale *m*, the iteration times T;
Output: Pareto optimal solution set P;
Step 1: Create m initial chromosomes as the initial population P based on K-means++;
Step 2: Generate m chromosomes in offspring Q based on the traditional genetic algorithm;
Step 3: Combine the parent with offspring in population R;
Step 4: Fast-nondominated-sort $R(m_i)$
All chromosomes in R are allocated into several layers according to dominant relationships. The nondominant chromosomes in R are allocated to the first layer; the nondominant chromosomes in the remaining population (removing the chromosomes in the first layer from R) are allocated into the second layer, and so on. The chromosomes in the same layer are assigned the same Pareto value;
Step 5: Crowding distance assignment $D(m_i)$
The crowding distance is defined as the sum value of multi-objective functions. This is important because it allows the ranking of chromosomes in the same layer;
Step 6: Elitism selection
Elitism selection is the selection of *m* chromosomes from R based on $R(m_i)$ and $D(m_i)$. This strategy can be used to retain good individuals and improve the overall level of population evolution;
Step 7: Judge if the terminal criteria are satisfied. If so, output the solution; if not, return to ***Step 2.***

communication phase. In this phase, the infrastructure condition information is inspected, synthesized, and transmitted from the sensors to the sink node. The steady communication process is executed as follows. First, the inspected data packets from non-CH sensors in each cluster are sent to the corresponding CH. Then, all the received data packets are synthesized by their respective CHs, which can reduce information redundancy and minimize energy consumption. Finally, the synthesized information packets at all CHs are transmitted to the sink node.

As shown in Fig. 3, the transmission rounds will repeat as long as the sensors have sufficient energy. In this optimal communication protocol designed above in each round, the communication between the sensors and sink node will operate in the energy-efficiency ways. The valid lifetime of the railway infrastructure monitoring wireless system will be extended ultimately.

V. SIMULATION VALIDATION AND ANALYSIS

In this section, the proposed protocol is validated via computer simulations with Python 3.6.2. Comparative case studies are carried out to demonstrate the superiority of the proposed scheme.

A. Simulation Configuration

In the paper, a rectangular region ($L \times W:$ 500 m $\times$ 50 m) along the rail is taken as the RIWMS monitoring area. The monitoring region is divided into ten small square regions ($L \times W:$ 50 m $\times$ 50 m) evenly, with ten sink nodes located above them. The information inspected by the sensors in the small square monitoring regions is sent to the corresponding sink nodes and then transmitted to the BS.

TABLE III
PARAMETERS OF THE OPTIMIZATION MODEL

No.	Parameter	Describe	Value
1	E_{ele}	Electronic energy	50 nJ/bit
2	ξ_{fs}	Amplifier energy (d^2)	10 pJ/bit/m^2
3	ξ_{mp}	Amplifier energy (d^4)	0.0013 pJ/bit/m^4
4	E_{in}	Initial energy	0.02 J
5	N	Sensor number	200
6	l (bit)	Data packets size	150
7	k_{opt}	Clusters number	8
8	Sink node	Location	(25, 50)
9	Sensor nodes	Square monitoring region	[(0 25), (0 –25), (50 –25), (50 25)]
10	m	Population size	100
11	p_m	Mutation rate	0.01
12	p_c	Crossover rate	0.85
13	t	Termination conditions	400
14	α	Distance adjust parameter	0.6

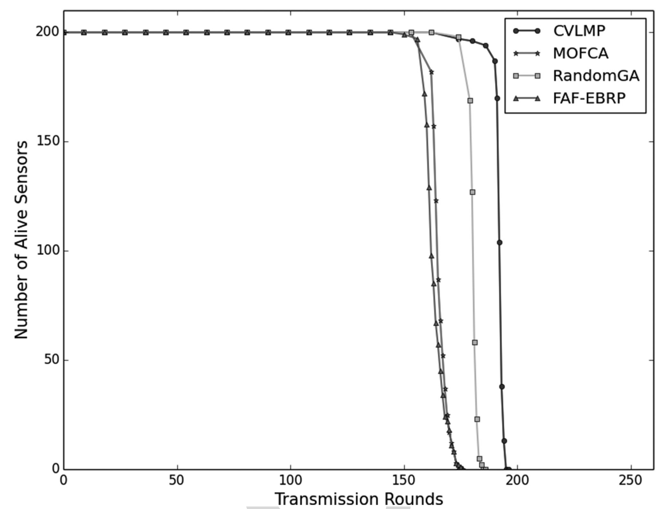


Fig. 7. Comparison of number of alive sensors.

This paper focuses on the protocol optimization of the communication between the sensors and sink nodes. We pick one sink node and 200 sensors in the corresponding small square monitoring region as the simulation objects. In this simulation scenario, the rail is considered as the *X*-axis, and its terminal point, which is far away from the BS, is set as the origin. The parameters used in the simulations are shown in Table III.

Based on (7), (9), and the simulation environment parameters in Table III, we can get $E[d^2_{\mathrm{toSK}i}] = 3264$ m^2 and initial optimal clusters number is $k_{\mathrm{opt}} = 7.75$. Since the number is a positive integer, we set $k = 8$ for the experiments in this paper. Moreover, the residual energy of the sensors declines with the communication. To reduce the load on the CHs and guarantee connectivity of the communication system, we increase the clusters number at the rate of 1 once the energy of the existing CHs is not sufficient to transmit all the information from their own clusters to the sink node. The crossover and mutation rates were experienced from the range of (0.75,0.95) and (0.005,0.02), respectively [28]. In this paper, they are selected as $p_c = 0.85$ and $p_m = 0.01$.

B. Simulation Results and Analysis

For a railway infrastructure wireless monitoring system, four aspects should be considered to ensure the stability and continuity.

1) The system valid lifetime

In this paper, the valid lifetime of the RIWMS is represented by the number of alive sensors and the total residual energy of all sensors after several transmission rounds.

1) The balance of energy consumption across all sensors

Balancing the energy consumption aims to avoid the death of some sensors due to energy exhaustion.

The variance of the remaining energy of all sensor nodes is used to reflect the balance of energy consumption, defined as

$$\mathrm{VRE} = \frac{\sum_{k=1}^{M} \left(E_{\mathrm{RE}k}(r) - \overline{E}_{\mathrm{RE}}(r)\right)^2}{M} \tag{30}$$

where M is the number of sensor nodes in the first monitoring region; $E_{\mathrm{RE}k}(r)$ is the remaining energy of sensor node k at time t; and $\bar{E}_{\mathrm{RE}}(r)$ is the average remaining energy of all sensor nodes.

Furthermore, the inspection data received at the sink node should be another major criterion to support the infrastructure condition estimation and prediction.

The comparisons among the proposed CVLMP, FAF-EBRP [10], and MOFCA [11] are conducted based on the aforementioned four criteria. To ensure the accuracy of the comparisons, the CHs setting and clusters generation for the two protocols are all implemented strictly according to the algorithm flow in [10] and [11]. Moreover, the CVLMP, MOFCA, and FAF-EBRP are all performed under the unique simulation environment (i.e., the railway wireless monitoring system) to guarantee the fair comparison. Additionally, we compare the solution based on the traditional NSGA-II, which generates the initial population randomly.

We compare and justify the performances of the proposed CVLMP with other two protocols FAF-EBRP and MOFCA in four criteria, as shown in Figs. 7–10.

1) From Fig. 7, we see that all sensors died after 190 transmission rounds using the CVLMP, whereas using the FAF-EBRP, the sensors death begins at 150 rounds and decreases quickly to 0 by 175 rounds. Using the MOFCA, the tendency of sensors death is similar to but a bit more than using the FAF-EBRP. The efficiency of the protocols is compared in Fig. 8 in terms of the total residual energy. Using CVLMP, FAF-EBRP, and MOFCA, the total residual energy declines are all smooth, while the most residual energy is using the CVLMP. Thus, the CVLMP is proved to be the longest-running protocol.
2) The variance of the residual energy in the CVLMP is smaller than those of the other two, as shown in Fig. 9, revealing that the CVLMP remains the most balanced energy consumption of all sensors. The CVLMP can keep all the sensors alive to the utmost and ensure the connectivity of the entire network.

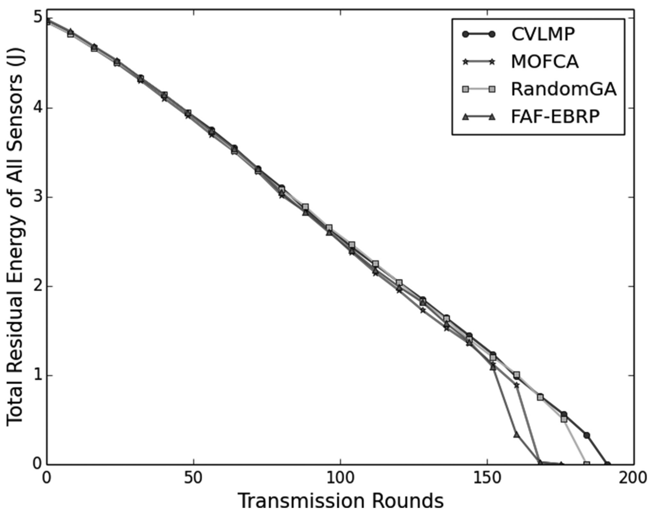


Fig. 8. Comparison of the total residual energy of all sensors.

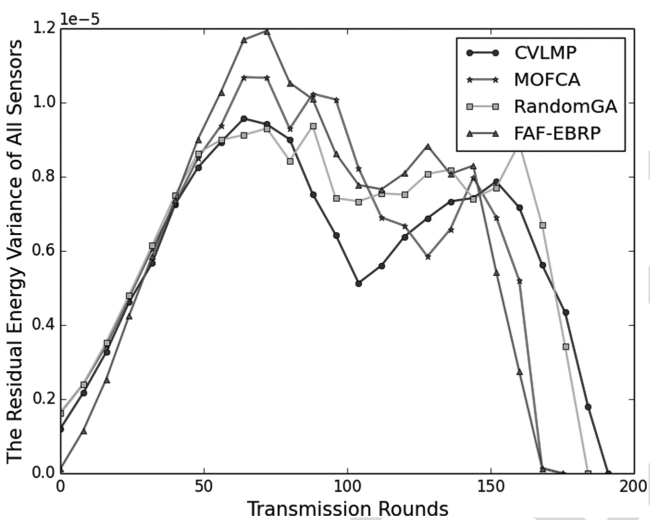


Fig. 9. Comparison of the residual energy variance of all sensors.

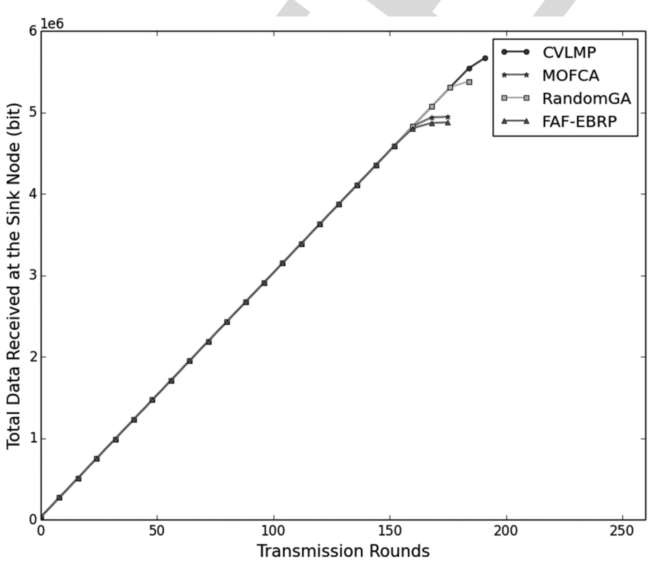


Fig. 10. Comparison of total data received at the sink node.

3) As shown in Fig. 10, the total data received at the sink nodes using the CVLMP are about 1.2 times of that using FAF-EBRP and MOFCA. Therefore, the protocols proposed in this paper render significantly more transmission rounds and received data in comparison with the other two methods, and this is crucial to maintaining the stable railway infrastructure monitoring and condition analysis.
4) Additionally, Figs. 7–10 reveal that the performances of the system using the CVLMP, which initializes the population by K-means++, are better than those using the random initialization population. Moreover, we found that the performances are not stable using the random GA due to the uncertain initialization population with limited iterations. More advanced control and monitoring schemes with robustness need to be studied in the future to optimize the system performance further [29], [30].

VI. Conclusion

Wireless railway infrastructure condition monitoring network is vital to the railway industry. Safe and efficient railway operations require a sufficient lifetime of the sensor network. This paper proposes a novel CVLMP to maximize the lifetime of the monitoring system. The optimization models are used to rotate the CHs and optimize the clusters before each transmission round starts, so as to minimize the total energy consumption and to balance the energy consumption among sensors.

Simulation results demonstrate that the superiority of the proposed CVLMP is threefold.

1) Compared with FAF-EBRP and MOFCA models in the literature, the CVLMP can effectively prolong the lifetime of the monitoring system by 23%, all else being equal.
2) The CVLMP has superior performance in optimally conserving the total residual energy of all the sensors and balancing energy consumption among sensors.
3) The monitoring data received at the sink node are more than those using the other two methods.

References

[1] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, "Wireless sensor networks for condition monitoring in the railway industry: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1088–1106, Jun. 2015.
[2] D. Barke and W. K. Chiu, "Structural health monitoring in the railway industry: A review," *Struct. Health Monit.*, vol. 4, no. 1, pp. 81–93, Mar. 2005.
[3] H. C. Tsai *et al.*, "Railway track inspection based on the vibration response to a scheduled train and the Hilbert–Huang transform," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 229, no. 7, pp. 815–829, Apr. 2015.
[4] K. Chebrolu *et al.*, "Brimon: A sensor network system for railway bridge monitoring," in *Proc. 6th Int. Conf. Mobile Syst., Appl., Serv.*, 2008, pp. 2–14.
[5] E. Berlin and K. V. Laerhoven, "Sensor networks for railway monitoring: Detecting trains from their distributed vibration footprints," in *Proc. IEEE Int. Conf. Distrib. Comput. Sens. Syst.*, 2013, pp. 80–87.
[6] G. M. Shafiullah, S. A. Azad, and A. B. M. Shawkat Ali, "Energy-efficient wireless MAC protocols for railway monitoring applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 649–659, Jun. 2013.
[7] K. Ovsthus and L. M. Kristensen, "An industrial perspective on wireless sensor networks—A survey of requirements, protocols, and challenges," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 3, pp. 1391–1412, Jan. 2014.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, 2000, pp. 1–10.

[9] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[10] D. Zhang, G. Li, K. Zheng, X. Ming, and Z.-H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 766–773, Feb. 2014.

[11] S. A. Sert, H. Bagci, and A. Yazici, "MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks," *Appl. Soft Comput.*, vol. 30, pp. 151–165, May 2015.

[12] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3066–3078, Aug. 2008.

[13] Z. Han, J. Wu, J. Zhang, L. Liu, and K. Tian, "A general self-organized tree-based energy-balance routing protocol for wireless sensor network," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 2, pp. 732–740, Apr. 2014.

[14] G. Han, L. Liu, J. Jiang, L. Shu, and G. Hancke, "Analysis of energy-efficient connected target coverage algorithms for industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 135–143, Feb. 2017.

[15] H. Zhang and C. Liu, "A review on node deployment of wireless sensor network," *Int. J. Comput. Sci. Issues*, vol. 9, no. 6, pp. 378–383, Nov. 2012.

[16] A. Cenedese, M. Luvisotto, and G. Michieletto, "Distributed clustering strategies in industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 228–237, Feb. 2017.

[17] H. Mahboubi, K. Moezzi, A. G. Aghdam, K. Sayrafian-Pour, and V. Marbukh, "Distributed deployment algorithms for improved coverage in a network of wireless mobile sensors," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 163–174, Feb. 2014.

[18] D. Zhang *et al.*, "A new clustering routing method based on PECE for WSN," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, Dec. 2015, Art. no. 162.

[19] F. Ren, J. Zhang, T. He, C. Lin, and S. K. Das Ren, "EBRP: Energy-balanced routing protocol for data gathering in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 12, pp. 2108–2125, Dec. 2011.

[20] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct. 2004.

[21] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach," *Eng. Appl. Artif. Intell.*, vol. 33, pp. 127–140, Aug. 2014.

[22] S. Kurt, H. U. Yildiz, M. Yigit, B. Tavli, and V. C. Gungor, "Packet size optimization in wireless sensor networks for smart grid applications," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2392–2401, Mar. 2017.

[23] F. Khan, "An initial seed selection algorithm for k-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application," *Appl. Soft Comput.*, vol. 12, no. 11, pp. 3698–3700, Nov. 2012.

[24] P. Bholowalia and A. Kumar, "EBK-means: A clustering technique based on elbow method and k-means in WSN," *Int. J. Comput. Appl.*, vol. 105, no. 9, pp. 17–24, Nov. 2014.

[25] D. Wei, Y. Jin, S. Vural, K. Moessner, and R. Tafazolli, "An energy-efficient clustering solution for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3973–3983, Nov. 2011.

[26] E. A. Zimichev, N. L. Kazanskii, and P. G. Serafimovich, "Spectral-spatial classification with k-means++ particional clustering," *Comput. Opt.*, vol. 38, no. 2, pp. 281–286, Jul. 2014.

[27] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliab. Eng. Syst. Safety*, vol. 91, no. 9, pp. 992–1007, Sep. 2006.

[28] Z. J. Wang and H. Hu, "Improved precast production–scheduling model considering the whole supply chain," *J. Comput. Civil Eng.*, vol. 31, no. 4, Jul. 2017, Art. no. 04017013.

[29] P. Li, R. X. Li, Y. Cao, D. Y. Li, and G. Xie, "Multi-objective sizing optimization for island microgrids using triangular aggregation model and levy-harmony algorithm," *IEEE Trans. Ind. Informat.*, Nov. 2017, to be published. Doi: 10.1109/TII.2017.2778079.

[30] P. Li, R. Dargaville, Y. Cao, D. Y. Li, and J. Xia, "Storage aided system property enhancing and hybrid robust smoothing for large-scale PV systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2871–2879, Nov. 2017.

Xiaoping Ma received the M.S. degree in 2014 from Beijing Jiaotong University, Beijing, China, where he is currently working toward the Ph.D. degree.

He is studying at the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His research interests include traffic information intelligent, sensing, and service technologies.

Honghui Dong received the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007.

He is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. His current research interests include pattern recognition and intelligent systems, as well as transportation science and engineering.

Xiang Liu received the Ph.D. degree from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2013.

He is with the Department of Civil and Environmental Engineering, Rutgers University, New Brunswick, NJ, USA. His research interests include railway safety and risk management, transportation system analytics and optimization, as well as intelligent railway transportation systems.

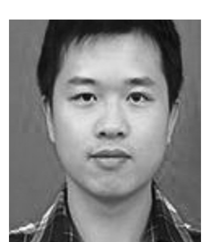
Limin Jia received the Ph.D. degree from China Academy of Railway Sciences, Beijing, China, in 1991.

He is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. His current research interests include safety science and engineering, control science and engineering, transportation engineering, safety technology, and engineering and system science.

Guo Xie received the Ph.D. degree from Nihon University, Tokyo, Japan, in 2013.

He is with the School of Automation and Information Engineering, Xi'an University of Technology, Xi'an, China. His current research interests include safety and reliability of railway systems, intelligent information processing, and fault diagnosis.

Dr. Xie was the recipient of the Japanese Government Scholarship from the Japanese Ministry of Education, Culture, Sports, Science and Technology.



Zheyong Bian received the B.S. and M.S. degrees in industrial engineering from Tianjin University, Tianjin, China, in 2016. He is working toward the Ph.D. degree in civil engineering at Rutgers University, New Brunswick, NJ, USA.

His research interests include operations research, logistics optimization, and freight transportation.

Queries

Q1. Author: Please check whether the edit made to the sentence "Exact analytical methods face difficulty in . . . " retains the intended sense.

Q2. Author: Please complete and update Ref. [29].

Q3. Author: Please provide the areas of study in which Xiaoping Ma received the M.S. degree, and Honghui Dong, Xiang Liu, Limin Jia, and Guo Xie received the Ph.D. degrees.