



# Article An Accuracy-Aware Energy-Efficient Multipath Routing Algorithm for WSNs

Feng Dan<sup>1</sup>, Yajie Ma<sup>1,\*</sup>, Wenqi Yin<sup>1</sup>, Xian Yang<sup>2</sup>, Fengxing Zhou<sup>1</sup>, Shaowu Lu<sup>1</sup>, and Bowen Ning<sup>1</sup>

- <sup>1</sup> Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; danfeng@wust.edu.cn (F.D.); yinwq980825@163.com (W.Y.); zhoufengxing@wust.edu.cn (F.Z.); shawn2013@wust.edu.cn (S.L.); ningbwen@wust.edu.cn (B.N.)
- <sup>2</sup> Alliance Business School, University of Manchester, Manchester M13 9PL, UK; xian.yang@manchester.ac.uk
- \* Correspondence: mayajie@wust.edu.cn

Abstract: In the fields of industrial production or safety monitoring, wireless sensor networks are often content with unreliable and time-varying channels that are susceptible to interference. Consequently, ensuring both transmission reliability and data accuracy has garnered substantial attention in recent years. Although multipath routing-based schemes can provide transmission reliability for wireless sensor networks, achieving high data accuracy simultaneously remains challenging. To address this issue, an Energy-efficient Multipath Routing algorithm balancing data Accuracy and transmission Reliability (EMRAR) is proposed to balance the reliability and accuracy of data transmission. The multipath routing problem is formulated into a multi-objective programming problem aimed at optimizing both reliability and power consumption while adhering to data accuracy constraints. To obtain the solution of the multi-objective programming, an adaptive artificial immune algorithm is employed, in which the antibody initialization method, antibody incentive calculation method, and immune operation are improved, especially for the multipath routing scheme. Simulation results show that the EMRAR algorithm effectively balances data accuracy and transmission reliability while also saving energy when compared to existing algorithms.

**Keywords:** artificial immune algorithm; data accuracy; energy-efficient; multipath routing; wireless sensor networks

# 1. Introduction

Wireless sensor networks (WSNs) are composed of many small, energy-limited, and low-cost sensor nodes deployed in a certain area, forming a multihop self-organizing network system through wireless communication. In fields like industrial production or safety monitoring, WSN applications have high requirements for data quality. In these applications, the reliability of data transmission and the accuracy of received data directly affect the response and handling of user applications to emergencies such as fire or toxic gas leakage. However, in environments like factories or outdoors, the wireless channels used in WSN communication exhibit the characteristics of unreliability, time variation, and susceptibility to interference [1,2]. These characteristics pose challenges in ensuring the reliability of the transmission path and the accuracy of received data, leading to widespread concerns in recent years [3].

An important technology for ensuring transmission reliability lies in the deployment of highly reliable routing algorithms [4–6], with multipath routing standing out as a widely adopted strategy to improve the reliability of data transmission in WSNs. In contrast to single-path routing, multipath routing augments data transmission reliability by establishing multiple paths between source and destination nodes. To achieve greater throughput, enhanced failure resilience, faster mobility and reduced latency, coding-based multipath routing schemes have garnered increasing research attention in recent years [7–9].



Citation: Dan, F.; Ma, Y.; Yin, W.; Yang, X.; Zhou, F.; Lu, S.; Ning, B. An Accuracy-Aware Energy-Efficient Multipath Routing Algorithm for WSNs. *Sensors* 2024, 24, 285. https:// doi.org/10.3390/s24010285

Academic Editors: Bartłomiej Płaczek and Marcin Bernaś

Received: 17 October 2023 Revised: 26 November 2023 Accepted: 28 November 2023 Published: 3 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). These routing schemes typically segment the original data into multiple coding packets in uniform size, containing some redundant information. The coding packets are then distributed across multiple paths for transmission. Intermediate nodes process the data through techniques such as data aggregation or compression to reduce the number of packets. Once a certain quantity of coding packets is collected, the destination node recovers the original data through decoding and data fusion [10]. However, it is worth noting that the aggregation of multiple coding packets at intermediate or sink nodes along the path may lead to a reduction in data accuracy [11,12]. This reduction in accuracy can have repercussions on the result of data analysis or decision-making. Consequently, maintaining data accuracy poses a challenge for coding-based multipath routing.

To address this challenge, current research efforts can be categorized into two main approaches. One approach focuses on improving the routing schemes to reduce the number of data copies by reasonably selecting or minimizing the transmission paths. In the literature [13], a combination of multipath routing and a tree-based aggregation algorithm was employed. During path construction, a node is selected and remains as the first node of aggregation, and convergence proceeds step by step from this node to the destination. In the literature [14], the authors combined the clustering method with multipath routing and utilized spatio-temporal correlation to aggregate sampled data at the source node to reduce the volume of transmitted data. The second approach to improving aggregation algorithms is to guarantee data accuracy by reducing the accuracy loss. Fan and Chen proposed a Scalable Counting (SC) algorithm to guarantee data accuracy [11]. The SC algorithm inherits the idea of the Linear Counting (LC) algorithm, which incorporates a hash function with arrays to reduce the accuracy loss caused by aggregation calculation. In the literature [15], a routing tree is established for data transmission. It calculates the maximum allowable data accuracy loss of each subtree in proportion to the size of the routing tree. The data compression ratio at each node is then determined based on the maximum allowable accuracy loss to ensure accuracy. However, there is a trade-off between data accuracy and data transmission reliability [11]-more coding packets can enhance the reliability of data transmission, whereas multiple aggregations of copies during routing can reduce data accuracy at the destination. As aggregation instances increase, the loss of data accuracy becomes more pronounced, which makes the accuracy-aware data collection algorithms less effective.

To tackle the aforementioned challenge, a multipath routing algorithm EMRAR (Energy-efficient Multipath Routing algorithm balancing data Accuracy and transmission Reliability) is proposed in this paper. The core concept behind EMRAR is to jointly optimize transmission reliability and data accuracy while considering energy consumption. Recognizing that data accuracy is influenced by the number of intermediate aggregation nodes, our approach aims to minimize this count. To achieve this goal, the measurement methods for transmission reliability and data accuracy are defined first. Taking transmission reliability and path energy consumption as the optimization objectives, an accuracy constraint-based optimization model is formulated to transform the routing problem into a multi-objective optimization problem with the help of a fitness function to assess the solutions and identify the optimal Pareto solution set for the algorithm. To enhance optimization performance, the Improved Adaptive Artificial Immune Algorithm (IAAIA) [16] is employed. To adapt IAAIA to multipath routing, multiple adaptive immune operators are introduced, which not only retain the search capabilities of the artificial immune algorithm but also harness the adaptive nature of immune operators to avoid "premature" convergence. Experimental results show that the proposed routing algorithm can efficiently find a solution to multipath within a reasonable convergence time, meeting the criteria for transmission reliability and data accuracy.

The main contributions of this work include:

1. A multipath routing algorithm EMRAR is proposed. Unlike existing schemes that either construct a congestion-free transmission route to guarantee data accuracy at the expense of reliability or solely ensure data accuracy at the source node, EMRAR

can concurrently guarantee transmission reliability and data accuracy at the destination node. It accomplishes this by efficiently controlling the number of paths and aggregation nodes to minimize accuracy loss.

- 2. A constraint-based multi-objective optimization formula is constructed. This formula considers both transmission reliability and path energy consumption as optimization objectives under the constraints of accuracy loss and node residual energy. This approach ensures a balanced trade-off between these critical factors.
- 3. The IAAIA algorithm is adopted and customized. To address the multi-objective optimization problem and apply it to the multipath routing scheme, the adaptation includes the design of an antibody initialization method and enhancements to the antibody incentive calculation method and immune operation. These modifications optimize the algorithm's effectiveness in the context of multipath routing.

The organization of the rest of this paper is arranged as follows. In Section 2, the existing research on multipath routing is discussed, with a particular focus on examining various multipath construction methods. In Section 3, the system model is established to obtain the reliability and energy optimization objectives while adhering to the accuracy loss constraint. In Section 4, the solution for the multi-objective optimization is proposed. The antibody population update strategy is designed, and the fitness function is constructed to evaluate the antibody individuals. In Section 5, the time complexity and message complexity of the algorithm are analyzed. In Section 6, the simulations are carried out to compare EMRAR with existing algorithms. Finally, the conclusions and future work are given in Section 7.

## 2. Related Work

#### 2.1. Multipath Routing in WSNs

Multipath routing techniques have found widespread application in enhancing the delivery of multimedia content, providing fault-tolerance routing, and supporting QoS across diverse network environments, including multihop Local Area Networks (LANs), Wireless Area Networks (WANs), the Internet, as well as ad hoc networks and WSNs [17]. The advantages of various multipath routing protocols for WSNs and their associated benefits were presented in [18]. Both best-effort and real-time multipath routing protocols for Wireless Multimedia Sensor Networks (WMSNs) are investigated in [19] from a network application perspective.

The performance of the Multipath Routing Protocol (MRP) is heavily reliant on both the quantity and quality of selected paths. Consequently, MRPs are designed with distinct principles, primarily based on whether the paths have intersection nodes or links, and can generally be categorized into three types: node-disjointed multipath, link-disjointed multipath, and intersection-based multipath. Disjointed multipaths are particularly popular because there is no node or link overlap among paths, which greatly mitigates the impact of node failure or link disconnection on the overall routing functionality. This characteristic offers substantial advantages in prolonging network lifetime and enhancing fault tolerance. Extensive research has been devoted to discovering multiple disjointed paths among multiple nodes over the past two or three decades. Although some algorithms only provide two disjointed paths [20–23], others can offer more paths [24,25].

Generally, routing protocols used for real-time applications typically impose stringent demands for various QoS metrics such as delay, throughput, and reliability. In [26], by finding two disjointed paths with the shortest transmission distance between source and destination nodes, the number of relay nodes is limited, which improves the reliability and fault tolerance of the network. Addressing the high maintenance costs associated with disjointed multipath routing, Ref. [27] proposes an enhanced AODV protocol. Recognizing that delay-sensitive data, integrity-sensitive data, and ordinary data impose varying QoS requirements on routing, Ref. [28] allocates reserved disjointed multipath to different data types according to the QoS capabilities of different paths, ensuring high delivery rates for integrity-sensitive data and low latency for delay-sensitive data. To mitigate

energy consumption and prolong the lifetime, an adaptive multipath routing scheme based on balanced energy consumption is proposed in the literature [29]. This approach jointly considers the path energy, distance, hops, and the lowest energy node, dynamically adjusting the amount of data transmission among multiple paths according to the change in energy information entropy of each path. Meanwhile, in [30], the authors present a multipath routing algorithm to prolong the lifetime by effectively balancing the offered load among forwarding nodes.

However, building completely disjointed multipaths in sensor networks can be challenging in practical scenarios. As demonstrated in [31], the effort to find a multipath routing solution with more than two disjointed paths can incur significant overhead, encompassing computational complexities, energy consumption, and storage demands, especially in sensor networks with only limited resources. Ref. [32] proves that it is NP-hard for a source node to identify more than three disjointed paths. To address this challenge, Ref. [27] optimizes disjointed multipath routing by setting aggregation nodes in networks. The aggregation nodes allow paths to intersect, therefore enhancing routing scalability. Furthermore, Ref. [8] augments the number of paths to improve transmission reliability by constructing braided paths around disjointed multipaths from source to destination. It becomes evident that compared to disjointed multipath routing, intersection-based multipath routing can provide more paths and is more straightforward to implement.

#### 2.2. Data Accuracy Guarantee in WSNs

According to the locations within the network where data accuracy guarantee tasks are performed, existing research can be categorized into two distinct schemes: transmission-based and source node-based schemes.

For the transmission-based scheme, as exemplified by literature [15], a congestion control algorithm is introduced to address the problem of substantial data accuracy loss in existing congestion control algorithms. This algorithm operates under the premise that nodes transmit data upward to the sink through a routing tree. Given the upbound error at the destination node, the maximum allowable error for each subtree can be calculated based on the tree's topological attributes. When congestion occurs, nodes employ lossy compression to compress data within the maximum tolerable error range, which can mitigate the congestion while preserving data accuracy to a certain extent at the destination node. Similarly, in the literature [33], accuracy loss is reduced by integrating a clustering algorithm with a routing algorithm, therefore diminishing the number of data copies in the network. This algorithm leverages a minimum spanning tree-based routing algorithm to transmit data copies to the destination node. In this scheme, cluster heads aggregate copies within their respective clusters to reduce the number of copies. However, the tree-based routing approach in this context is less reliable.

In the source node-based scheme, existing algorithms try to ensure accuracy during the data collection. For instance, in the literature [14], spatial-temporal correlations of sampling data are harnessed to reduce the volume of transmitted data and ensure the data accuracy. This approach then integrates multipath routing with a clustering algorithm to ensure accuracy during transmission. However, it cannot address the issue of accuracy loss caused by data aggregation in multipath transmission. In [34], the authors proposed an AUV-aided hybrid data collection scheme based on the Value of Information (VoI). It tries to keep the VoI of urgent information generated at the source node by assigning it a higher transmission priority. Nevertheless, it attenuates the VoI of normal data, potentially compromising the accuracy of all data.

It can be seen that the transmission-based scheme struggles to deliver highly reliable transmissions to the destination, while the source node-based scheme faces challenges in eliminating the accuracy loss during transmission or may focus solely on specific data types. In a large-scale sensor network employing multipath routing to ensure data transmission reliability and considering the constraints of limited resources, adopting intersection-based multipath routing appears reasonable. Therefore, data aggregation at intersection nodes is

inevitable, and consequently, controlling the number of aggregation nodes becomes crucial to reduce the accuracy loss caused by multiple aggregations.

## 3. Multi-Objective Optimization Model for EMRAR

To design the EMRAR algorithm, we formulate the task of identifying a multipath set as a multi-objective optimization subject to constraints, which jointly optimize the objectives of reliability and energy consumption while adhering to the constraint of accuracy loss, which translates to reducing the number of intersection nodes. The construction of the optimization model is detailed as follows.

The coding scheme will not be specified and analyzed as it falls outside the scope of the focus of this work.

For ease of expression, the terms "node", "sensor", or "sensor node" will be interchangeably used to refer to a sensor node in the subsequent sections of this paper.

Consider a wireless sensor network system with *n* sensor nodes randomly distributed in the monitoring area. Due to the limitation of the communication distance, all nodes send their collected data to the sink via a multihop mode. To facilitate our discussion, we establish the following definitions based on Graph Theory.

- Weighted Undirected Graph *G*: a WSN can be illustrated as a weighted undirected graph *G* = {*V*, *E*}, where *V* = {*v*<sub>1</sub>, *v*<sub>2</sub>, ..., *v*<sub>n</sub>} is the set of sensor nodes, and *E* = {*e*<sub>1</sub>, *e*<sub>2</sub>, ..., *e*<sub>k</sub>} is the set of edges between two sensor nodes.
- Neighbor: Suppose the communication range of each sensor node is R. Let the distance between sensor nodes be *i*∈*V* and *j*∈*V* be *d*(*i*, *j*). If *d*(*i*, *j*) ≤ *R*, then nodes *i* and *j* are neighbors, and there is an edge between them. The neighbor relationship can be represented as *i*∈*N*(*j*) or *j*∈*N*(*i*), where *N*(*i*)⊂*V* and *N*(*j*)⊂*V* are the neighbor sets of *i* and *j*, respectively.
- Multipath Set *MP*(*i*, *S*): Suppose a source node *i*∈*V* can set up multipath to the sink *S*∈*V*. Then, the set of all the paths from *i* to *S* can be represented as *MP*(*i*, *S*). |*MP*(*i*, *S*)| is the number of paths of *MP*(*i*, *S*). Each path in *MP*(*i*, *S*) can be represented as *mp*<sub>i</sub>(*i*, *S*).

j = 1, ..., |MP(i, S)|.  $|mp_i(i, S)|$  is the number of nodes in path  $mp_i(i, S)$ .

#### 3.1. Optimization Objects

1. Reliability object

Reliability is evaluated by the transmission success rate (TSR) of the data transmitted from a source node to the sink by multiple paths. It is assumed that bit errors and packet losses are the primary factors impacting the transmission success rate. The TSR can be calculated as follows.

For bit error

Suppose the data transmission between neighbor nodes is a single hop, and the bit error rate (BER) in the single-hop wireless channel at node *k* is  $e_k$ . Then, the probability of correctly transmitting one bit to the next hop of node *k* is  $(1 - e_k)$ . Then, the probability of correctly transmitting a packet with *L* bits on one hop at node *k* is

$$p_k = (1 - e_k)^L \tag{1}$$

For  $\forall j$ , as  $mp_j(i, S)$  has  $|mp_j(i, S)|$  hops, then the probability of correctly transmitting a packet from *i* to *S* on  $mp_j(i, S)$  is

$$p_{BER_j} = \prod_{k=1}^{|mp_j(i, S)|} p_k$$
 (2)

Then, the average probability of correct transmission on BER is

$$p_{BER} = \frac{1}{|MP(i, S)|} \sum_{j=1}^{|MP(i, S)|} p_{BER_j}$$
(3)

#### • For packet loss

Suppose the packet loss rate (PLR) for a path  $mp_j(i, S)$  is  $q_j$ , then the probability of successfully transmitting a packet from *i* to *S* on  $mp_j(i, S)$  is  $(1 - q_j)$ . Hence, the average probability of successful transmission of PLR is

$$p_{PLR} = \frac{1}{|MP(i, S)|} \sum_{j=1}^{|MP(i, S)|} (1 - q_j)$$
(4)

For a packet, if it does not arrive at the sink properly, either bit error or packet loss happens. If both happen, it can be treated as packet loss. Then, the average successful transmission probability  $p_h$  is

$$p_h = max\{p_{BER}, p_{PLR}\}$$
(5)

To meet the demand for reliability, more paths are required. Suppose the paths required to deliver *M* coding packets to the sink is  $|MP(i, S)| (|MP(i, S)| \le M)$ . For each packet, the successful transmission probability is  $p_h$  and the failure probability is  $(1 - p_h)$ . The *TSR* for transmitting these *M* packets on |MP(i, S)| paths is

$$TSR = \sum_{m=1}^{M} C^{m}_{|MP(i,S)|} (1-p_{h})^{(|MP(i,S)|-m)} p_{h}^{m}$$
(6)

To optimize *TSR*, we have

$$Obj1: Max(TSR) \tag{7}$$

# 2. Energy consumption object

In a sensor network, energy consumption is a critical factor of concern. Having a greater number of paths increases reliability, but it also results in higher energy consumption. Suppose the energy consumption for packet transmission from one node to its neighbor is  $E_p$ . Then, the energy consumption for transmitting a packet on path  $mp_i(i, S)$  is

$$E_j = E_p \sum_{k=0}^{|mp_j(i, S)|-1} p_k$$
(8)

Therefore, on the multipath transmission with |MP(i, S)| paths, the total energy consumption is

$$E_m = \sum_{j=1}^{|MP(i, S)|} E_j = E_p \sum_{j=1}^{|MP(i, S)|} \sum_{k=0}^{|mp_j(i, S)|-1} p_k$$
(9)

To optimize  $E_m$ , we have

$$Obj2: Min(E_m) \tag{10}$$

## 3.2. Constraints

1. Accuracy constraint

Accuracy is used to evaluate the variance between the data collected and sent by the source node and the corresponding data received at the sink. Data aggregation predominantly takes place at path intersections during data transmission. For a given path, as explained in the previous section, the fewer intersections it has with other paths (or less aggregation occurs), the smaller the decrease in accuracy caused by data aggregation during data transmission. To quantify this accuracy decrease, we introduce three metrics: NAL (Node Aggregation Loss), PAL (Path Accuracy Loss), and MAL (Multipath Accuracy Loss), defined as follows:

• NAL: An NAL  $\delta(0 \le \delta \le 1)$  means that the approximation transmission error on an aggregation node on path  $mp_i(i, S)$  is within a factor of  $\delta$ .

PAL: Suppose a path *mp<sub>j</sub>(i, S) (j 1, ..., |MP(i, S)|)* has c<sup>(i)</sup><sub>j</sub> intersections with other paths in *MP(i, S)*. If the accuracy loss for each aggregation is δ, then for each path *mp<sub>j</sub>(i, S)*, the PAL Δ<sup>(i)</sup><sub>j</sub> is:

$$\Delta_i^{(i)} = 1 - (1 - \delta)^{c_j^{(i)}} \tag{11}$$

• MAL: Suppose the data collected and sent by sensor node *i* is  $x_i$ . After multipath transmission, the data received by sink *S* is  $x_i'$ . If the transmission is successful on each path in MP(i, S), then the MAL  $\Delta^{(i)}$  for MP(i, S) from *i* to *S* can be calculated as:

$$\Delta^{(i)} = \frac{|x'_i - x_i|}{x_i} = \frac{\sum_{j=1}^{|MP(i,S)|} \Delta_j^{(i)}}{|MP(i,S)|}$$
(12)

Hence, for a predefined maximum MAL value  $\Delta$ , the accuracy constrain can be represented as:

$$\Delta^{(i)} = \frac{\sum_{j=1}^{|MP(i,S)|} \Delta_j^{(i)}}{|MP(i,S)|} \le \Delta$$
(13)

To satisfy Formula (13), the multipath routing algorithm must try to find multiple paths with fewer intersection nodes. Thus, it also gives the constraint on the value of  $c_i^{(i)}$ .

# 2. Node energy constraint

As data transmission and aggregation all consume energy, only the nodes with sufficient residual energy can participate in multipath routing. Suppose the residual energy for a node *k* in a path  $mp_j(i, S)$  (j = 1, ..., |MP(i, S)|) is  $E_k$   $(k = 1, ..., |mp_j(i, S)|)$ . The initial energy of node *k* is  $E_{max}$ , and the predefined minimum residual energy required for a relay node in a path is  $E_{min}$ . Accordingly, we have

$$E_{min} \le E_k \le E_{max}, k = 1, ..., |mp_i(i, S)|$$
 (14)

#### 3.3. Multi-Objective Optimization Function

To design the multi-objective function, the reliability object Formula (7) needs to be transferred to an equivalent minimization function as

$$Max(TSR) \equiv Min(1/TSR)$$
(15)

Then, the evaluation function can be designed as

$$Obj = \omega_1 \frac{1}{TSR} + \omega_2 E_m \tag{16}$$

where  $\omega_1 + \omega_2 = 1$ . Hence, the multi-objective optimization problem can be formulated as

$$\begin{cases} Min(Obj) = Min(\omega_1 \frac{1}{TSR} + \omega_2 E_m) \\ Subject to \\ \omega_1 + \omega_2 = 1 \\ \frac{\sum_{j=1}^{|MP(i,S)|} [1 - (1 - \delta)^{c_j^{(i)}}]}{|MP(i,S)|} \le \Delta \\ E_{min} \le E_k \le E_{max}, k = 1, ..., |mp_j(i,S)| \end{cases}$$

$$(17)$$

It can be seen that the problem of balancing data accuracy and data transmission reliability has been modeled as a multi-objective optimization. In multi-objective optimization, different objectives exhibit varying value ranges and dimensions, which can potentially lead to conflicts among optimization objectives. For instance, optimizing one objective may inadvertently compromise the performance of another one. Consequently, in the context of multi-objective and multi-constraint optimization problems, the optimization objectives manifest as convex functions with optimal solutions. Generally, there is no universal optimal solution that satisfies all objective functions. Therefore, the introduction of non-inferior solutions also referred to as satisfactory solutions or Pareto optimal solutions, becomes imperative. The outcome of multi-objective optimization is to generate a set of mutually balanced solutions where the target values of these solutions are non-inferior to one another while still meeting the specified constraints.

## 4. Multi-Objective Optimization Resolution

In this section, we tackle the solution to the constrained multi-objective optimization problem (16). In addition to classical algorithms, recent years have witnessed a surge in the popularity of intelligent algorithms, including evolutionary algorithms [16,35]. In this context, an Improved Adaptive Artificial Immune Algorithm (IAAIA) [16] is adopted, which is an enhancement of the classic Artificial Immune Algorithm (AIA) [36]. IAAIA not only inherits the search characteristics of the AIA algorithm but also exhibits robust global search capabilities. Furthermore, it improves convergence speed, accuracy, and stability, all of which are achieved without introducing additional computational complexity to the algorithm.

The IAAIA framework encompasses several key steps, including antigen and antibody construction, initial antibody group formation, affinity calculation, immune operations, antibody group updates, and criteria for algorithm termination. However, IAAIA is primarily geared towards mathematical optimization and cannot be directly applied to the optimization of routing schemes in this work. To apply IAAIA to multi-objective optimization (16), we must customize the antibody initialization method, antibody incentive calculation method, and immune operations while retaining its advantages of fast convergence and precise computation. Detailed calculations and modifications are described in the following subsections.

#### 4.1. Antibody Coding

In the immune algorithm, antibody represents the feasible solution to the optimization problem. In the routing scheme, an antibody represents a group of paths from the source node to the destination node, one of which is a gene of the antibody. Hence, suppose the candidate multipath set from source node *i* to sink *S* is *C*, which is generated by the following Algorithm 1.

The number of paths in *C* is *L*. Then the antibody has *L* genes. Extract l ( $l \le L$ ) paths from *C* randomly to form a subset  $C' \subseteq C$ . Let *C'* be the initial antibody population. Then, for the *L* genes, there are *l* non-zero genes and L - l zero genes, where each non-zero gene corresponds to a path from the source node *i* to the sink node *S*. Because the hops of each path may be different, variable length coding scheme is adopted in this algorithm. As an example, shown in Figure 1, source node *i* has multipath to the destination *S* via nodes set {1, 2, 3, 4, 5, 6, 7}. Suppose an antibody is composed of three genes. As a path is a gene, the coding of the antibody is { $3 \rightarrow 4 \rightarrow 5$ ,  $1 \rightarrow 2$ ,  $6 \rightarrow 7 \rightarrow 2$ }. Based on this coding scheme, the number of joint nodes in this antibody and the hops of each path can be calculated, and then the accuracy loss, energy consumption, and reliability of the antibody can also be calculated.

## 4.2. Fitness Function

In the immune algorithm, an antigen represents the problem to be optimized, and the fitness function serves as a foundational tool for computing the individual fitness values. These fitness values gauge the compatibility or closeness between an antibody and an antigen, a concept akin to the affinity observed in genetic algorithms. As an antibody is composed of *L* genes, the number of non-zero genes in each antibody may be different, and then the corresponding energy loss  $E_m$  and transmission success rate *TSR* are

different. According to Formula (16), the individual fitness function of the antibody can be expressed as:

$$FIT = 1 - Obj \tag{18}$$

Since the smaller the *Obj* value is, the better the optimization effect is, and the larger the fitness *FIT* is, the better the antibody performance is.

# Algorithm 1. Candidate multipath set generation

Input: network topology, source node *i*, sink node *S*, a PAL threshold  $\varepsilon$ , a MAL threshold  $\Delta$ ,  $C = \Phi$ Output: *C* Find all the paths  $mp_i(i, S)$  (j = 0, 1, 2, ...) from *i* to *S* that satisfy the node energy constraint Formula (14); for each *j* Add  $mp_i(i, S)$  to *C*; Calculate  $\Delta_j^{(i)}$  according to Formula (11); if  $\Delta_j^{(i)} > \varepsilon$ Calculate  $\Delta^{(i)}$  according to Formula (12); if  $\Delta^{(i)} > \Delta$ Remove  $mp_i(i, S)$  from *C*; end end end



Figure 1. An example of an antibody coding scheme.

#### 4.3. Antibody Simulation Calculation

Antibody simulation serves as a comprehensive metric for evaluating the quality of antibodies. It encompasses both fitness and antibody density, with the latter being a crucial measure of antibody diversity within the population. Suppose  $f_{aff}(C'_u, C'_v)$  be the affinity between antibodies  $C'_u \subset C'$  and  $C'_v \subset C'$ . As the affinity indicates the similarity between antibodies, the value of  $f_{aff}(C'_u, C'_v)$  can be measured by the number of the same paths in these two antibodies. Hence, the same paths set in  $C'_u$  and  $C'_v$  can be expressed as the intersection of two sets and  $f_{aff}(C'_u, C'_v)$  can be calculated as:

$$f_{aff}(C'_u, C'_v) = \left| C'_u \cap C'_v \right| \tag{19}$$

Set  $\rho(0 \le \rho \le 1)$  as the threshold of the similarity, then judge whether two antibodies are similar or not can be transferred to a 0/1 binary judgment

$$S(C'_{u}, C'_{v}) = \begin{cases} 0, \ f_{aff}(C'_{u}, C'_{v})/L \le \rho \\ 1, \ f_{aff}(C'_{u}, C'_{v})/L > \rho \end{cases}$$
(20)

The density of antibody  $C'_{\mu}$  can be calculated as:

$$f_{den}(C'_u) = \frac{1}{N} [1 + \sum_{v=1}^M S(C'_u, C'_v)]$$
(21)

where *N* is the size of the population, and M = N - 1 is the size of the population after removing antibody  $C'_{\mu}$ .

Then an antibody  $C'_u$  with higher affinity and lower concentration has higher simulation, which can be calculated as the composition of its fitness and antibody density:

$$f_{sim}(C'_{u}) = \beta \frac{FIT(C'_{u})}{\sum_{u=1}^{N} FIT(C'_{u})} - (1-\beta) \frac{f_{den}(C'_{u})}{\sum_{u=1}^{N} f_{den}(C'_{u})}$$
(22)

where  $\beta(0 < \beta < 1)$  is the weight coefficient of fitness, which is related to the number of iterations and can be calculated as:

$$\beta = 1 - \frac{(G_{max} - G_{min})}{G_{max}} \cdot \frac{G}{G_{max}}$$
(23)

where  $G_{max}$  and  $G_{min}$  are the maximum and minimum number of iterations, respectively. *G* is the current number of iterations.  $\beta$  represents the impact of previous-generation antibodies on contemporary antibodies. As the number of iterations increases, the value of  $\beta$  linearly decreases, which enables the algorithm to have a high global search ability in the early stage to obtain suitable antibodies while having a high local search ability in the later stage to improve convergence accuracy.

#### 4.4. Population Updating

The population update operation comprises three essential steps: antibody cloning, antibody mutation, and antibody updating. This operation continuously refreshes the antibody population until the specified termination condition is satisfied.

## 1. Antibody cloning

To speed up the convergence of the population, the antibody with a high incentive should be selected for cloning. The key step involves calculating the average antibody incentive level across the population and subsequently selecting antibodies with incentives exceeding this average for cloning.

## 2. Antibody mutation

Mutation operation means that in the cloned antibody population, the antibody mutates according to the probability  $p_m$  of producing a new antibody. This operation can improve the local search performance of the algorithm and find potential antibodies. The variation probability  $p_m$  can be expressed as:

$$p_m = \begin{cases} \frac{k_1(f_{max} - f_{d})}{f_{max} - f_{avg}}, f_d \ge f_{avg} \\ k_2, f_d < f_{avg} \end{cases}$$
(24)

where  $f_{max}$  and  $f_{avg}$  are the maximum value and average value of the antibody fitness, respectively.  $f_d$  is the fitness of the mutated antibody.  $k_1$  and  $k_2$  are constants.

During this process, mutation introduces spontaneous random changes in antibodies. Its purpose is to maintain population diversity and prevent the optimization from falling into local optima. The process follows the coding rules of the algorithm as outlined below.

- Select an antibody from the antibody population with the mutation probability *p<sub>m</sub>*. Select one of the genes and select a node on this gene as the mutation node.
- The path from the source node *i* to the mutation node remains unchanged, and the mutation node searches the path to the sink node *S* again to find a new path.
- Get new antibodies.

# 3 Antibody updating

The population updating operation is to eliminate the antibodies with a low incentive degree in the antibody population and retain the first *N* antibodies with a high incentive degree to form a new population through the above operation.

The framework of the multi-objective optimization solution is illustrated in the flowchart in Figure 2.



Figure 2. The flowchart for solving multi-objective optimization function.

## 4.5. EMRAR Algorithm Procedure

The overall procedure of EMRAR is described by the following Algorithm 2. It illustrates the steps of establishing a multipath between a sensing node *i* and the sink node *S*. The |MP(i, S)| routes discovery can be realized by source routing protocols like DSR [37], in which request and reply messages will be sent among nodes between *i* and Shop by hop. Then, the values of  $p_{BER_j}$  and  $q_j$  can be obtained by node *i* from the reply messages, therefore calculating the values of  $p_{BER}$  and  $p_{PLR}$ .

The EMRAR algorithm operates round and round, similar to LEACH [38], with the goal of achieving a balanced energy consumption among nodes. However, there is a notable distinction between the two approaches. In LEACH, each cluster head enters a sleep state automatically in the next round, whereas in EMRAR, the states of nodes on the multipath in the current round do not enter a sleep state automatically in the subsequent round. Instead, the state is determined based on their residual energy during the initialization process in

the subsequent round. The duration for each round can be predefined, for example, 15 min, or adjusted to align with the realistic network environment.

Algorithm 2. Find optimal multipath by IAAIA		
Step 1	Network initialization: Each node exchanges residual energy, location, and other information with its neighbors to establish a neighbor table and determine relevant parameters.	
Step 2	Path discovery: when sensing node <i>i</i> needs to transmit data to sink node <i>S</i> , it sends path probing messages to neighboring nodes. After <i>S</i> receives all probing messages, the paths are initialized as a candidate path set if they satisfy the node energy constraint.	
Step 3	Encoding: <i>S</i> filters the paths by accuracy constraint; establishes antigen and antibody; encodes them to generate the initial antibody population.	
Step 4	Path selection: calculate the fitness of each path; select paths by elitist selection by calculating the simulations of them to form the antibody population.	
Step 5	Path update: update the paths within the population according to the population update rules.	
Step 6	Repeat: repeat Steps 4 and 5 until the algorithm completion condition is met to obtain the optimal path information.	
Step 7	Multipath establishment: <i>S</i> returns a response message along the optimal paths. Upon receiving the response message by node <i>i</i> , the multipath is established successfully.	
Step 8	Nodes that are not on the established multipath enter a sleep state till the next round of EMRAR.	

#### 5. Complexity Analysis

#### 5.1. Message Complexity

Suppose each node has  $\sigma$  neighbors. In the network initialization phase, each node exchanges information with its neighboring nodes once. Therefore, a total of  $2\sigma$  messages are required for a node to complete the initialization. Then, the message complexity is  $O(\sigma)$ . In the path discovery phase, a node also needs to send  $\sigma$  messages to its neighbors, and the messages will be forwarded to maximum  $\zeta$  hops. Then, the message complexity is  $O(\sigma\zeta)$ . After the destination node obtains the optimal paths, it will send back a response along each path in the multipath set. Then, the message complexity is O(1). Based on the above analysis, the message complexity is  $O(\sigma \zeta)$ .

# 5.2. Time Complexity

Suppose each node has  $\sigma$  neighbors, and at most  $\sigma$  paths will be initialized. The time complexity of this step is  $O(\sigma)$ . The time complexity of the IAAIA algorithm is determined by the immune operation and the number of iterations. The time complexity of crossover in each generation is  $O(\sigma^2)$ . The fitness values of all paths need to be calculated. Thus, the time complexity of mutation in each generation is also  $O(\sigma)$ . If the number of iterations is  $\eta$ , then the total time complexity of EMRAR is  $O(\sigma) + O(\eta \times \sigma) + O(\eta\sigma^2) = O(\eta\sigma^2)$ .

# 6. Evaluation

## 6.1. Simulation Configuration

In this simulation, the configuration of the main parameters is listed in Table 1.

Table 1. The configuration of main parameters.

Parameters	Configuration
СРИ	Core I5 processor
Main frequency	2.5 GHz
Sensing area range	$500 \text{ m} \times 500 \text{ m}$
maximum communication range	50 m
Initial energy of sensors	2 J
Bit error rate <i>e</i>	random(0, 1] $ imes$ 10 <sup>-6</sup>
Packet loss rate	$random(0, 1] \times 10^{-2}$
Node Accuracy Loss $\delta$	0.005
MAC protocol	802.15.4
Packet length	16 kbits

The sensor energy consumption adheres to the classic model described in [38] (the sensors are powered by batteries and do not benefit from alternative energy supply, such as wireless-powered communication technology [39,40]). The energy consumption by a node to transmit *k* bits of data over 1 meter is:

$$C_s(k,l) = E_{elec}k + \varepsilon_{amp}kl^2 \tag{25}$$

where the radio dissipates  $E_{elec} = 50 \text{ nJ/bit}$  to run the transmitter or receiver circuity and  $\varepsilon_{amp} = 0.015 \text{ nJ/(bit·m^2)}$  is the transmit amplifier.

For a receiving node, the energy consumption to receive *k* bits of data is:

$$C_r(k) = E_{elec}k \tag{26}$$

Hence, on the link *e* between two adjacent nodes, the energy consumption C(e) is the sum of the energy consumption for transmitting  $C_s$  and the energy consumption for receiving  $C_r$ , which can be expressed as:

$$C(e) = C_s(k, l) + C_r(k)$$
 (27)

Based on the energy consumption model, the following performance indexes are used to evaluate the performance of the EMRAR algorithm:

• Average Accuracy Loss (AAL): the relative error between data sent by the source node and the data received and recovered by the destination node. The AAL value in this work is the average loss of *m* data received during the simulation time, which can be calculated as:

$$AAL = \left(\frac{1}{m}\sum_{i=1}^{m} \frac{x_i - x'_i}{x_i}\right) \times 100\%$$
(28)

where  $x_i$  is the data sent out by the source node, and  $x'_i$  is the data received and recovered by the receiving node.

• Packet Delivery Rate (PDR): the ratio of the number of packets received by the destination node *Num<sub>R</sub>* to the number of packets sent by the source node *Num<sub>S</sub>*, which can be calculated as:

$$PDR = \frac{Num_R}{Num_S} \times 100\%$$
<sup>(29)</sup>

• Network Residual Energy (NRE): the ratio of the residual energy of all the nodes in the network to the initial energy of the network, which can be calculated as:

$$NRE = \frac{\sum E_{re}}{\sum E_{ini}} \times 100\%$$
(30)

The above indexes are evaluated under different densities of adjacent nodes (average number of neighbors of a node) or different data transmission rates. The experiments are performed in OMNet++ by means of the MiXiM framework. It uses the topology generator to generate random topologies for simulation each time. The experiments are repeated till the results can be achieved at a precision of 1% with a 90% confidence interval. The maximum number of iterations is set to 100.

# 6.2. The Effect of Accuracy Constraints on the Performance of the Algorithm

In this section, the impact of the accuracy constraint  $\Delta$  on the EMRAR algorithm is analyzed under different values of  $\Delta$  when the data rate is 50 kbps. In line with prior research [11,41], we set the accuracy constraint  $\Delta$  with 3 distinct values—1%, 3%, and

14 of 21



5%—in each simulation, representing incremental levels of allowable accuracy loss. The AAL, PDR, and NRE performances in different node densities are shown in Figure 3.

Figure 3. The effect of  $\Delta$  on AAL, PDR, and NRE under different adjacent node densities when the data transmission rate is 50 kbps. (a) AAL performance, (b) PDR performance, (c) NRE performance.

Figure 3a shows the AAL performance under different densities of adjacent nodes. From the figure, we can find 3 patterns:

- 1. A larger value of  $\Delta$  corresponds to a larger value of AAL. This correlation can be attributed to the fact that, with a fixed network size, a larger accuracy constraint  $\Delta$  implies a reduced constraint on the paths available for multipath routing. Consequently, the probability of path intersections increases, leading to more data aggregation and ultimately resulting in a higher AAL.
- 2. An increasing density of adjacent nodes leads to a lower AAL, signifying improved data accuracy. The reason for this is that a higher number of neighbors provides more available next-hop nodes for routing. The algorithm can choose the routes that lead to less aggregation towards the sink, consequently reducing accuracy loss.
- 3. The reduction in AAL becomes more obvious with a larger value of  $\Delta$ . The reason for this is that when there is a greater number of neighbors, a larger value of  $\Delta$  provides a relatively greater number of alternative nodes. Hence, it becomes easier to find better path choices that significantly reduce the accuracy loss.

Figure 3b shows the PDR performance under different densities of adjacent nodes. The shapes of all three curves are quite similar. For example, when  $\Delta = 3\%$ , the PDR increases by approximately 4.5% for every 5-unit increase in the number of adjacent nodes. This trend occurs because as the density of adjacent nodes increases, there are more available paths that meet the constraints. This results in a greater number of reliable paths for data transmission, fewer intersection nodes, and reduced congestion in the MAC layer. As a result, the PDR is improved. On the other hand, when the density of adjacent nodes is held constant, increasing  $\Delta$  relaxes the constraint on the path and increases the number of available paths, also resulting in improved PDR.

Figure 3c shows the NRE performance under different densities of adjacent nodes after 500 rounds. It can be seen that:

- 1. Across various values of density, the energy consumption at  $\Delta = 5\%$  consistently demonstrates superior performance compared to other values of  $\Delta$ . The reason for this is that when the accuracy constraint is relaxed, more intersection nodes are permitted in the multipath, allowing the establishment of shorter routes (with fewer hops) for data transmission, which can save the energy consumption of the whole network and prolong the lifetime.
- 2. With the increase in node density, the NRE exhibits a consistent increase across various values of Δ. The reason for this is that higher density implies a larger number of nodes within the area, and the percentage of active nodes relative to the total number of nodes decreases, resulting in an elevated NRE ratio.

The AAL, PDR, and NRE performances in different data transmission rates are shown in Figure 4, with an adjacent node density of 20.



**Figure 4.** The effect of  $\Delta$  on AAL, PDR, and NRE under different data transmission rates when the adjacent node density is 20. (a) AAL performance, (b) PDR performance, (c) NRE performance.

Figure 4a shows the AAL performance under different data transmission rates. It can be seen that as the data transmission rate increases from 50 kbps to 100 kbps, the AAL increases consistently across all  $\Delta$  values, signifying a reduction in data accuracy. This is because a higher data transmission rate necessitates the delivery of a greater number of data packets. Consequently, additional paths need to be established to facilitate their transmission. Hence, the path intersection will increase, and the data aggregation at intersection nodes will affect the accuracy of the data received at the sink. Furthermore, at the same data transmission rate, increasing the accuracy constraint  $\Delta$  results in a greater number of established paths, further elevating the likelihood of path intersections and subsequent data aggregation. This, in turn, leads to an increase in data accuracy loss.

Figure 4b shows the PDR performance under different data transmission rates. As we can see, the higher data transmission rate corresponds to lower PDR values, and  $\Delta = 1\%$  exhibits the lowest PDR compared to other values of  $\Delta$ . This is because higher transmission rates introduce a higher offered load on the network, leading to increased congestion and a subsequent reduction in PDR. Conversely, when the data transmission rate remains constant, PDR increases with the increasing  $\Delta$  values. This is because of the relaxation of accuracy constraints on the path, resulting in many available paths and subsequently improving the performance of data delivery.

Figure 4c shows the NRE performance under different data transmission rates. It can be seen that a higher transmission rate brings a larger volume of data being transmitted that will consume more energy by the sensor nodes, resulting in lower NRE values. Additionally, smaller values of  $\Delta$  also correspond to lower NRE values. The reason for this is that smaller values of  $\Delta$  restrict the utilization of intersection nodes within the multipath, potentially requiring the transmission paths with more hops. Therefore, a greater number of nodes must participate in data transmission to consume more energy and consequently lower NRE values.

# 6.3. Comparison with Other Algorithms

In this section, we conduct a comparative analysis of the EMRAR algorithm against three other algorithms: CADC proposed in [15], AMRBEC proposed in [29], and IQMRP proposed in [35]. Figure 5 presents a performance comparison of these four algorithms in terms of AAL, PDR, and NRE under different data transmission rates with a fixed  $\Delta$  value of 3% and an adjacent node density of 20. The NRE is calculated after running each algorithm for 500 rounds.

From these three figures, we can see that as the data transmission rate increases from 50 bps to 100 bps, all four algorithms exhibit a similar trend on AAL, PDR, and NRE. However, for different densities of adjacent nodes, EMRAR consistently outperforms the other three algorithms in most scenarios. The rationale behind this can be explained as follows.

For the AAL in Figure 5a, CADC establishes a tree-based topology rooted at the sink node. It gives upper bounds for estimation error and estimation distortion, which are used to evaluate the accuracy of the data. Although it employs lossy compression to mitigate congestion, it primarily focuses on meeting these upper bounds, potentially resulting in a relatively high AAL if the bounds are not stringent. In contrast, IQMRP adopts an Ant Colony Optimization (ACO) technique to choose the multipath, which, while reducing the number of intersection nodes to some extent, has limited effectiveness in minimizing AAL. For AMRBEC, it consistently establishes a disjointed multipath, and the data are only aggregated at the sink of the node, leading to the lowest data accuracy loss among the four algorithms.

For the PDR in Figure 5b, as the data transmission rate increases, the PDRs of all four algorithms decline because of heightened packet congestion. Among them, CADC does not adopt multipath routing. The single transmission route results in a lower transmission reliability compared to the others. IQMRP chooses multipath based on QoS parameters such as residual energy, bandwidth, and next-hop accessibility. It improves the PDR but with a predefined threshold for path quality evaluation, which is less adaptive. The metrics for AMRBEC to choose the multipath are similar to IQMRP. It is based on the residual energy, distance, hops, and the balance of energy consumption, which are less related to channel quality, leading to more pronounced errors or packet losses than those of IQMRP. EMRAR treats reliability as an optimization objective directly, which can improve the PDR effectively.



**Figure 5.** The comparison of three algorithms under different data transmission rates when  $\Delta = 3\%$  and the adjacent node density is 20. (a) AAL performance, (b) PDR performance, (c) NRE performance.

For the NRE in Figure 5c, CADC exhibits poor NRE performance due to two primary reasons. First, the routing algorithm is not power-aware, and second, it uses a tree-based topology to transmit data to the sink, which is known to have a disadvantage in balancing energy consumption at the nodes near the sink. IQMRP chooses the multipath by balancing the residual energy, bandwidth, and next-hop accessibility, with residual energy just one of the considerations. AMRBEC initially has the highest residual energy at a lower data transmission rate (<50 kbps), but it degrades as the data transmission rates increase. There are two reasons. First, it must establish the node-disjointed multipath, which involves selecting more nodes, and it is difficult to find shorter paths while adhering to the energy constraints. Second, AMRBEC establishes the multipath without considering the quality of the paths, such as error and packet loss. Thus, the retransmission reliability constraint. As a result, it has the best residual energy performance, especially in scenarios with higher offered loads.

Figure 6 provides the comparison of the convergence times under different densities of adjacent nodes. The convergence time is evaluated by the average time required for the network to establish a set of multiple paths from the source to the destination in IQMRP, AMRBEC, and EMRAR or a node to find a route to the sink on a tree-based topology in CADC. It can be seen that as the network size increases, the convergence time of each algorithm increases. However, CADC shows linear growth with increasing adjacent node density, while IQMRP, AMRBEC, and EMRAR show exponential increases. This difference

arises because the CADC algorithm only needs to find the shortest path from the source to the destination. It has a linear time complexity to the number of edges in the network. In contrast, IQMRP and EMRAR share similar time complexities—IQMRP adopts the classic ACO algorithm, while EMRAR adopts the IAAIA algorithm that improves the antibody excitation operator of the AIA algorithm and accelerates the convergence. Therefore, EMRAR boasts a shorter convergence time IQMRP. For AMRBEC, which shares the same time complexity as EMRAR, the convergence time curve closely resembles EMRAR's when the density of the adjacent nodes is larger than 10. However, at lower node densities, AMRBEC may struggle to find a multipath solution with disjointed nodes, resulting in considerably longer average converge times in the simulation. Therefore, AMRBEC is not suitable for scenarios with low node density.



**Figure 6.** The comparison of three algorithms on convergence time under different adjacent node densities.

Figure 7 shows the comparison of average end-to-end delay under different data transmission rates. End-to-end delay is evaluated by the average time taken for data to travel from the source to the destination. It can be seen that as the data transmission rate increases, resulting in a higher offered load in the network, the average end-to-end delay of each algorithm increases. Among them, both CADC and AMRBEC algorithms exhibit poorer end-to-end delay performance. For CADC, although it establishes a tree-based topology with the shortest path, data aggregation occurs at each branch node of the tree, leading to increased end-to-end delay caused by channel sharing at branch nodes. On the other hand, AMRBEC, which must establish node-disjointed multipath, faces challenges in establishing multiple shortest paths, occasionally resorting to longer paths. Hence, at lower data transmission rates, AMRBEC fails to outperform CADC in terms of end-toend delay. However, as the data transmission rate increases due to the node-disjointed multipath, AMRBEC performs better than CADC because of less data aggregation and queuing. IQMRP has the best performance in this comparison because it establishes a QoS-based multipath routing, with latency being one of the measurements. For EMRAR, it prefers paths with small accuracy loss to low-latency next-hop or shortest paths, resulting in an increased end-to-end delay in comparison with IQMRP.

The simulation results reveal distinctive strengths and weaknesses among the evaluated algorithms. CADC is superior in convergence time but poor in all other performance metrics. IQMRP shows the best performance in average end-to-end delay, but it has the largest convergence time and is relatively compromised in other aspects. AMRBEC has minimal accuracy loss and the highest residual energy when the data transmission rate is low. However, its performance is less satisfactory in other aspects, especially when the data transmission rate and node density are low. In contrast, our proposed algorithm, EMRAR,





Figure 7. The comparison of average end-to-end delay under different data transmission rates.

## 7. Conclusions

This paper proposed the multipath routing algorithm EMRAR, which aims to strike a balance between transmission reliability and data accuracy in wireless sensor networks. It employs a multi-objective optimization with constraints, utilizing an improved immune genetic algorithm, IAAIA, to obtain the optimal solution. Simulation results verify that the proposed EMRAR algorithm outperforms existing algorithms in terms of transmission reliability, data accuracy, and energy consumption. Convergence time and end-to-end delay are also discussed. The findings suggest that the EMRAR algorithm is well-suited for scenarios with high requirements in data accuracy and transmission reliability at the receiving end, where real-time computing or delay sensitivity is less critical. In addition, the work in this paper provides a valuable method to optimize multiple objectives simultaneously in multipath routing for sensor networks, with potential applications in addressing similar challenges in WSNs.

Our future research in this area will focus on applying the EMRAR algorithm to mobile ad hoc networks, such as VANET, where network topology dynamically changes due to vehicle movements. In such dynamic environments, calculating multiple path sets in real time to satisfy the evolving requirement is crucial. Hence, the impacts of node mobility and the real-time performance of the algorithm are issues that must be further considered.

**Author Contributions:** Conceptualization, F.D.; methodology, F.D., Y.M. and W.Y.; software, W.Y. and Y.M.; investigation, Y.M.; writing—original draft preparation, F.D. and W.Y.; writing—review and editing, Y.M. and X.Y.; supervision, F.Z.; project administration, B.N.; funding acquisition, F.Z. and S.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China: Key Technology Research of Early Fault Diagnose & Residual Lifetime (grant number 51975433) and Research on High-Speed and High-Precision Control Strategy of Linear Servo System under Periodic Disturbance (grant number 51975430).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

# Nomenclature

The following nomenclature lists the symbols used in this manuscript, along with their physical interpretations, respectively.

d(i, j)	The distance between sensor nodes <i>i</i> and <i>j</i>
N(i)	The neighbor sets of sensor node <i>i</i>
MP(i, S)	The set of all the paths from a source node <i>i</i> to the sink <i>S</i>
mp <sub>j</sub> (i, S)	A path $j$ in $MP(i, S)$
$e_k$	The bit error rate (BER) in the single-hop wireless channel at node $k$
$p_k$	The probability of correctly transmitting a packet on one hop at node $k$
$q_j$	The packet loss rate (PLR) for a path $mp_i(i, S)$
$p_{BER_i}$	The probability of correctly transmitting a packet from <i>i</i> to <i>S</i> on $mp_i(i, S)$
PBER	The average probability of correct transmission on BER
$p_{PLR}$	The average probability of successful transmission on PLR
$p_h$	The average successful transmission probability
$E_j$	The energy consumption for transmitting a packet on path $mp_i(i, S)$
$E_m$	The total energy consumption on $MP(i, S)$
δ	The node aggregation loss
$c_j^{(i)}$	The number of intersections that a path $mp_j(i, S)$ has with other paths in $MP(i, S)$
$\Delta_j^{(i)}$	The path accuracy loss
$x_i$	the data collected and sent by sensor node <i>i</i>
$\Delta^{(i)}$	The multipath accuracy loss
$C'_u$	An antibody

## References

- 1. Chang, H.; Lu, S.; Zheng, S.; Song, B.; Yang, J. Integration of Predictive Control and Interconnected Structure for Auto-tuning Velocity Controller. *IEEE-ASME Trans. Mechatron.* **2023**. [CrossRef]
- 2. Yan, S.; Gu, Z.; Park, J.H.; Xie, X. A delay-kernel-dependent approach to saturated control of linear systems with mixed delay. *Automatica* **2023**, *152*, 110984. [CrossRef]
- 3. Wu, F.; Luo, T.; Tan, H. Case studies of WSN-CPS applications. In *Cyber-Physical System Design with Sensor Networking Technologies*; Zeadally, S., Jabeur, N., Eds.; IET Press: London, UK, 2016; pp. 269–310.
- Ekmen, M.; Altın-Kayhan, A. Reliable and energy efficient wireless sensor network design via conditional multi-copying for multiple central nodes. *Comput. Netw.* 2017, 126, 57–68. [CrossRef]
- 5. Lin, Y.; Pan, C.; Yeng, C. Network reliability for multipath TCP networks with are transmission mechanism under the time constraint. *J. Stat. Comput. Sim.* 2018, *88*, 2273–2286. [CrossRef]
- 6. Prasad, R.V.; Rao, V.S.; Sarkar, C.; Niemegeers, I. ReNEW: A practical module for reliable routing in networks of energy-harvesting wireless sensors. *IEEE Trans. Green Commun. Netw.* 2021, *5*, 1558–1569. [CrossRef]
- Wu, R.; Ma, J.; Tang, Z.; Li, X.; Choo, K.-K.R. A generic secure transmission scheme based on random linear network coding. IEEE ACM Trans. Netw. 2022, 30, 855–866. [CrossRef]
- Li, Z.; Xu, M.; Liu, T.; Yu, L. A network coding-based braided multipath routing protocol for wireless sensor networks. *Wirel. Commun. Mob. Comput.* 2019, 2757601. [CrossRef]
- 9. Han, C.; Yin, J.; Ye, L.; Yang, Y. NCAnt: A network coding-based multipath data transmission scheme for multi-UAV formation flying networks. *IEEE Commun. Lett.* 2021, 25, 1041–1044. [CrossRef]
- Chen, G.; Li, H.; Han, S.; Zhong, Z.; Chan, E. Network coding-aware multipath routing in multi-hop wireless networks. J. Softw. 2010, 21, 1908–1919. [CrossRef]
- Fan, Y.; Chen, A. energy efficient schemes for accuracy-guaranteed sensor data aggregation using scalable counting. *IEEE Trans. Knowl. Data Eng.* 2012, 24, 1463–1477. [CrossRef]
- 12. Chang, H.; Lu, S.; Zheng, S.; Ma, Y.; Yang, C.; Song, B. Integrated thrust ripple identification and compensation for linear servo system using an MP algorithm. *ISA Trans.* **2023**, *142*, 594–605. [CrossRef] [PubMed]
- Atti, A.; Challal, Y.; Hadjidj, A.; Bouabdallah, A. Braided disjoint branch routing protocol for WSNs. In Proceedings of the IEEE International Conference on Broadband and Wireless Computing, Communication and Applications, Compiegne, France, 28–30 October 2013; pp. 106–113.
- 14. Chowdhury, S.; Roy, A.; Benslimane, A.; Giri, C. On semantic clustering and adaptive robust regression based energy-aware communication with true outliers detection in WSN. *Ad Hoc Netw.* **2019**, *94*, 101934. [CrossRef]

- 15. Zhuang, Y.; Yu, L.; Shen, H.; Kolodzey, W.; Iri, N.; Caulfield, G.; He, S. Data collection with accuracy-aware congestion control in sensor networks. *IEEE Trans. Mob. Comput.* **2019**, *18*, 1068–1082. [CrossRef]
- Meng, Y.; Wang, T.; Li, Z.; Cai, J.; Zhu, S.; Han, C. Improved adaptive artificial immune algorithm for solving function optimization problems. J. Beijing Univ. Aeronaut. Astronaut. 2021, 47, 894–903.
- Lou, W.; Liu, W.; Zhang, Y. Performance optimization using multipath routing in mobile ad hoc and wireless sensor networks. In *Combinatorial Optimization in Communication Networks*; Cheng, M.X., Li, Y., Du, D.-Z., Eds.; Springer: New York, NY, USA, 2006; Volume 18, pp. 117–146.
- 18. Masdari, M.M.M.; Tanabi, M.T.M. Multipath routing protocols in wireless sensor networks: A survey and analysis. *Int. J. Future Gener. Commun. Netw.* **2013**, *6*, 181–192. [CrossRef]
- 19. Hasan, M.Z.; Al-Rizzo, H.; Al-Turjman, F. A survey on multipath routing protocols for QoS assurances in real-time wireless multimedia sensor networks. *IEEE Commun. Surv. Tutor.* 2017, *19*, 1424–1456. [CrossRef]
- Bhandari, R. Optimal diverse routing in telecommunication fiber networks. In Proceedings of the Conference on Computer Communications, Toronto, ON, Canada, 12–16 June 1994; pp. 1498–1508.
- 21. Björklund, A.; Husfeldt, T. Shortest two disjoint paths in polynomial time. SIAM J. Comput. 2019, 48, 1698–1710. [CrossRef]
- 22. Gottschau, M.; Kaiser, M.; Waldmann, C. The undirected two disjoint shortest paths problem. *Oper. Res. Lett.* **2019**, 47, 70–75. [CrossRef]
- 23. Karaata, M. An algorithm for finding two node-disjoint paths in arbitrary graphs. J. Comput. Inf. Technol. 2020, 27, 1–14. [CrossRef]
- 24. Lopez-Pajares, D.; Alvarez-Horcajo, J.; Rojas, E.; Carral, J.A.; Martinez-Yelmo, I. One-shot multiple disjoint path discovery protocol (1S-MDP). *IEEE Commun. Lett.* **2020**, *24*, 1660–1663. [CrossRef]
- 25. Lopez-Pajares, D.; Rojas, E.; Carral, J.A.; Martinez-Yelmo, I.; Alvarez-Horcajo, J. The disjoint multipath challenge: Multiple disjoint paths guaranteeing scalability. *IEEE Access* 2021, *9*, 74422–74436. [CrossRef]
- Yang, N.; Chen, Y.; Wang, Y.; Li, Q. A multipath transmission protocol based on degree constrained shortest transmission tree. *Comput. Eng. Appl.* 2017, 53, 126–130.
- Sharma, A.; Tharani, L. Ant colony based node disjoint local repair in multipath routing in MANET network. *Wirel. Pers. Commun.* 2022, 127, 159–186. [CrossRef]
- Chen, Z.; Zhou, W.; Wu, S.; Cheng, L. An on demand load balancing multi-path routing protocol for differentiated services in MWSN. *Comput. Commun.* 2021, 1, 296–306. [CrossRef]
- 29. Liu, J.; Meng, X.; Li, S.; Cui, X.; Wu, H. An adaptive multipath routing method based on improved GA and information entropy. *IEEE Sens. J.* **2022**, *22*, 22264–22275. [CrossRef]
- 30. Yu, C.-M.; Ku, M.-L.; Wang, L.-C. BMRHTA: Balanced multipath routing and hybrid transmission approach for lifecycle maximization in WSNs. *IEEE Internet Things J.* 2022, *9*, 728–742. [CrossRef]
- 31. Qin, F.; Liu, Y. On-demand multipath routing for mobile ad hoc networks. In Proceedings of the 26th Annual IEEE Conference on Local Computer Networks, Tampa, FL, USA, 14–16 November 2001; pp. 64–70.
- 32. Liu, B.; Luo, J.; Shan, F.; Wei, L.; Shen, X. On finding maximum disjoint paths for many-to-one routing in wireless multi-hop network. *IEICE Trans. Inf. Syst.* 2014, 10, 2632–2640. [CrossRef]
- 33. Rajesh, L.; Mohan, H. Adaptive group teaching based clustering and data aggregation with routing in wireless sensor network. *Wirel. Pers. Commun.* **2021**, *122*, 1839–1866. [CrossRef]
- 34. Liu, Z.; Meng, X.; Liu, Y.; Yang, Y.; Wang, Y. AUV-aided hybrid data collection scheme based on value of information for Internet of underwater Things. *IEEE Internet Things J.* **2022**, *9*, 6944–6955. [CrossRef]
- 35. Alqahtani, A. Improve the QoS using multi-path routing protocol for Wireless Multimedia Sensor Network. *Environ. Technol. Innov.* **2021**, *24*, 101850. [CrossRef]
- Castro, D.; Zuben, V. Learning and optimization using the clonal selection principle. *IEEE Trans. Evolut. Comput.* 2002, *6*, 239–251. [CrossRef]
- Johnson, D.; Hu, Y.; Maltz, D. RFC4728: The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4. February 2007. Available online: https://www.rfc-editor.org/rfc/rfc4728.html (accessed on 16 October 2023).
- Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2000; pp. 3005–3014.
- Chu, Z.; Zhu, Z.; Zhou, F.; Zhang, M.; Al-Dhahir, N. Intelligent reflecting surface assisted wireless powered sensor networks for Internet of Things. *IEEE Trans. Commun.* 2021, 69, 4877–4889. [CrossRef]
- 40. Cao, K.; Ding, H.; Li, W.; Lv, L.; Gao, M.; Gong, F.; Wang, B. On the ergodic secrecy capacity of intelligent reflecting surface aided wireless powered communication system. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 2275–2279. [CrossRef]
- 41. Considine, J.; Hadjieleftheriou, M.; Li, F.; Byers, J.; Kollios, G. Robust approximate aggregation in sensor data management systems. *ACM Trans. Database Syst.* 2009, 34, 1–35. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.