AN OPTIMAL DATA AGGREGATION SCHEME FOR WIRELESS SENSOR NETWORK USING QOS PARAMETERS WITH EFFICIENT FAILURE DETECTION AND LOSS RECOVERY TECHNIQUE

A. Raja Basha[*], *C. Yaashuwanth*[†]

Abstract: WSN: Wireless Sensor Networks play a significant part in its modern era but its limited power supply acts as a blocking stone in it growth. In order to save energy in WSN the concept of aggregator node is introduced, where the aggregator node would act as a mid-point between the source and destination node during the data transmission. The data aggregation process creates major problems like excess energy expenditure, and delay. In the process of eliminating or reducing the delay and energy expenditure, the researchers have been handled in different ways. Applications like environment monitoring, target tracking, military surveillance and health care require reliable and accurate information. Many researchers have proposed data aggregation techniques to enhance the latency, average energy consumption and average network lifetime. However, these techniques are not sufficient to address situations like node failure and loss recovery. This paper proposes to build a solid wireless sensor system which concentrate on efficient optimal data aggregation along with additional QoS metrics such as failure detection and loss recovery. The first contribution of this paper is to propose an Improved Wolf Optimization (IWO) algorithm for clustering. The clustering process includes an efficient cluster formation like, Cluster Head (CH), and Sub Head (SH) selection. The second contribution of this paper is inclusion of failure detection and loss recovery. The former is developed based on Multi-criteria Moths-Flame Decision-making (MMFD) model and the latter is achieved through SH. SH node will act as the backup node for cluster head when failure instances are detection. CH recovers the lost data through SH, which minimize the additional delay of backup node selection process and save much more energy. The results are simulated using network simulator 2 tool and it is compared with existing techniques. The Network Simulator – 2 results disclose that the findings are better than the available existing methodologies.

Key words: *optimal data aggregation, improved wolf optimization (IWO), failure detection, data loss recovery*

[*]Adam Raja Basha – Corresponding author; Department of Electronics and Communication Engineering, Audisankara College of Engineering and Technology (Autonomous), Gudur, AP, India, E-Mail: arajabasha@gmail.com

[†]Calpakkam Yaashuwanth; Department of Information Technology, Sri Venkateswara College of Engineering (Autonomous), Sriperumbudur, Tamilnadu, India

Neural Network World 5/2019, 301–324

Received: August 15, 2018
Revised and accepted: August 28, 2019

DOI: 10.14311/NNW.2019.29.019

1. Introduction

In any sensor network deployments the number of sensing nodes and that varies in number from few hundred and goes up to many hundreds in order to perform a specific operation. Those nodes are in-expensive and very small in size. Those sensor nodes are built with very limited inbuilt memory and a transceiver antenna for communication purpose along with a processing unit which is capable to do limited intelligent tasks. The possible ways to guarantee the successful data reception in the event of, the sensor nodes are failed during the operation is the bottle neck of WSN [1]. In the common fabrication methods, the sensing nodes are built with an integrated processor. However, the real time complication like, overall cost, scalability in any extend, tolerances level during node failures and so on, are in need to be addressed in justifiable level. These are important factors in realization of wireless sensor network [2]. The sensors networks are prone to frequent failure, because of many applications. There is a need to deploy the sensors in harsh and contaminated environments such as battlefield, tough weather conditions etc., and deployed sensor might suffer from many faults due to environmental impacts such as lightning, dust and moisture. These things will reduce the quality of wireless communications and possibly divert the sensor nodes from its desirable operations [3]. These failures might be the cause for data failure and functional failures in WSN and also, the defect arises out of hardware components of sensors networks due toits low cost components as well as nano scale components, which also negatively impacts the desirable network operations [4, 5]. In addition to these complications, the software "bug" also significantly impacts the network operations [6–8]. Data failures would lead to incorrect response from the network manager, and the faulty nodes are responsible for inaccurate routing which leads to heavy energy loss into the system,and also data's through intermediate faulty nodes leads to unpredictable losses in the wireless sensor network.

Many faultfinding methods have been established to identify the failures in the wireless sensor network. A distributed solution for a canonical task [9] is used to detect environmental events failure in WSN. A fault-tolerant based clustering technique which is based on clock synchronization scheme, in this method, the clustering is done at each and every rounds of the clock synchronization [10]. The mechanism based on probabilistic Bayesian decoding technique [11] is effectively employed in detection and resolving the faulty sensor readings from the WSN deployment by examining the values of spatial correlation, which understands the impressions of the nearby sensors and this, is the base for the technique. And this is the first protocol developed for fault tolerance in a single-hop network, and it brought improvements in energy efficiency and also provided the solution for fault tolerant and problems in permutation routing [12]. The first paper which classified the fault [13] issued the fault-tolerant fusion rule to make the decision about the failures from the local sensors, in this methodology, the decision fusion is combined with an integrating channel decoding for fault-tolerant classification.

The combined technique has brought a new rule of fusion, where the local decision rule integrates with soft decision decoding with introduction of no redundancy [14].

To identify the faults in the wireless sensor network a fault management system is placed which will identify the faults either online and offline and classify the same [15]. An anomaly detection approach [16] has a passion of fusion which will fuse the gathered data from the different sensing nodes with the utilization of principal component analysis. To detect the dissimilarities in WSN CESVM and QSSVM are used which are pronounced as cantered hyper ellipsoidal support vector machine and quarter-sphere support vector machines. These two techniques are very effective in identifying the dissimilarities from wireless sensor networks [17]. The approach which addresses the very limited diverse level of fault tolerance without using the extensive number of relay node is demonstrated as one of the impacting methods [18]. A data aggregation method with fault tolerant capability has achieved the aggregation tree repair activity without any initiated operation only with the help of local information and also automatically reschedules the nodes to achieve interference free data aggregation [19]. A recursive subspace tracking technique is used to detect the fault in online itself, and this detection approach would also track the fault with the utilization of OPASTA: orthonormal projection approximation subspace tracking algorithm to minimize the arithmetic complexity [20]. This OPASTA is efficiently used to calculate the Eigen-vector and Eigen-values at the places where the fault occurred. This paper proposes a Quality of Service contributed Optimal Data Aggregation (Q-ODA) technique with efficient failure detection and loss recovery. The modified moth flame decision-making algorithm (MMFDM) is used to achieve the failure detection and an improved wolf optimization (IWO) algorithm is used for clustering with selected backup node (SH) is used to achieve the loss recovery.

The proposed Q-ODA technique consists of two fold. First, the efficient clustering is performed by Improved Pair Detection (IPD) algorithm using the position and velocity of sensor nodes. Those QoS metrics from each and every cluster member are collected for the purpose of computation of the Rank of the sensor nodes. The node which has the highest rank in the cluster is selected as Cluster Head (CH), the node which has the second highest rank in the cluster is selected as Sub Head (SH). The QoS parameters are time varying functions and hence the ranks are flexible with parameter which make CH and SH are time varying and depend on the Rank. The failure detection is performed by the multi-criteria moth-flame based decision-making (MMFD) model with the prior knowledge of statistical distributions of sensing data i.e. delivery rate, loss rate, delivery time, overhead of received data at receiver end. With the analysis based on this result, we may classify the received aggregated data as normal and fault. On this fault detection process, when the detected fault is caused by the cluster member (SH) then, the proposed technique finds the nearby best Rank node for further process. When the detected fault is caused by the H node means, the proposed technique automatically choose SH as H node.

The below phrases will elaborate the contributions done on this paper:

1. The clustering is done using Improved Pair Detection method where the fundamental pair detection algorithm is modified with two scenarios to provide better pair selection and this process is based on the Rank system which is energy efficient and less selection process

2. Q-ODA the proposed technique shall utilize the simple decision-making model rather than the Support Vector Machine which is used for the failure detection in SFDLA [31]

3. Does not require any additional algorithm for backup node selection process since the Q-ODA chooses SH from available H node without requiring any new algorithm. SH acts as H in the event of header node failure

4. The proposed Q-ODA is computed using network simulator 2 tool and the results are compared with the existing methodologies namely DFTR [21], PDAFT [22], FIDA [24], EMDC [28], MDFU [30] and SFDLA [31] and the simulation results emphasis our proposed work an edge over than the above methodologies specifically in relations with energy consumption, End-to-End delay, data delivery ratio, data accuracy, network lifetime, failure rate, and cost of data aggregation.

In this paper work, the contributions of the fellow researchers are elaborated in Section 2. The problem statement is done at Section 3. The proposed Q-ODA scheme along with mathematical model derivation is explained in detail at Section 4. The detailed analysis of network simulator outputs and its performance comparisons are explained in Section 5. At Section 6 this paper concludes with the result.

2. Related works

A DFTR: Distributed energy efficient and fault tolerant routing algorithm is developed by Azharuddin et al. [21]. In their work, they have achieved the focussed energy efficiency through the selection process of one hop to another hop during data transmission. The routing process works on the proposed algorithm which ensures data byte transfer from successive node in case of cluster head failure. The compromising points of this method are cost function. And it is computed from the distance between the gateway to BS (Base Station) and the gateway of next-hop to BS along with the cost estimation of BS and the gateway residual energy of the next-hop which is complex and less attractive

Chen et al. [22] came up with a scheme called privacy-preserving data aggregation for secure smart grid communication with fault tolerance (PDAFT). They have utilized an encryption method known as Homomorphism Parlier Encryption to secure the sensitive information of the user, because of this over cover makes the base station about no clue about the originating node of the received data. Their work missed to demonstrate any strong cover to individual node since they have compromised a few servers at the Base station control unit.

The major privacy issue faced by the mobile nodes in the wireless sensor network is addressed by Chen et al. [23]. They proposed a private data aggregation technique they have efficiently handled the sensing problems faced by mobile crowd. They have sensibly handled the connections and disconnections of the mobile nodes through the group management protocol which also deals with the data-integrity verification process in which it had addressed the retreat of data weakness of specific data range. A specific buffering methodology is proposed by the authors to

deal with the fault tolerance of a future message. Over head during computation as well communication difficulties are over come at a desired level while deploying this methodology.

Wu et al. [24] developed a technique called FIDA: fault influence domainan alysis scheme or wireless sensor network. They have proposed a domain information system to convey the failure details to the system in the event of fault occurrences. They have grouped the main administration into main and sub system where the sub system keeps getting the guidelines to arrange resource management in the event of failure. These arrangements effectively contribute to overcome the impact of failure. The simulation with discrete event demonstrates that this domain based system has considerable impact on failure handling in WSN.

Cheraghlou et al. [25] have proposed failure identification and retrieval mechanisms for numerous fault levels in wireless sensor network. In their work they have included ways to addresses the fault arises out of node to node communication and fault at each and every node level as a results the networks fault tolerance level and communication failure tolerance level increased considerably along with three to five fold increase in life span of the overall sensor network is visualized. But all these are not achieved with a cost while executing this protocol the information interchange is enormously increased which results in much more energy expenditure than saved because of this method and the end results in overall life span of the WSN is reduced.

Xu et al. [26] have proposed an optimization algorithm to investigate the packet loss rate during data transmission and queuing waiting delay under self-similar data movement flow of WSN. Their proposed work is based on no dominated sorting genetic algorithm. Their functional concentration is self-similar QoS parameter and they have named it like B is the function cache and the channel data transfer rate is C and they have done the computation of B and C at different quality of service constrained environments. The outcome of this method is special administrative mechanism which effectively handled the queuing delay and data byte loss of the wireless sensor network.

Mahdi et al. [27] have proposed a weighted data aggregation routing to achieve maximum possible data aggregation using hop-tree. Based on the local states of each and every node the hop-tree is build and it keeps updated based on the same through which the adaptive performance of the node shall be obtained for event-driven WSNs. The energy saving is achieved by this methodology through finding shortest path between the sink and source node through finding the ideal point where the route gets overlapped. Based on the distribution, comprehensive weights and adaptive cost the next hop is selected for data byte transmission which they called triple cost function.

Henna et al. [28] have proposed an ideal solution to the excess energy spent during the relay action of the fault tolerance processes of wireless sensor networks. They have come up with approximation which could be proved mathematically and they have called it EMDC: Energy Efficient Maximum Disjoint Coverage. Through this approximation they have shorted out the problem of target coverage in fault tolerance without much energy expenditure. And also it is the continuous work of Adam Raja Basha [34–36].

Guan et al. [29] have presented one of the best performing fault tolerance methods for smart grid environment with highest user security. They named it as Secret Sharing of Aggregated Data with Fault Tolerance. The outcome of this proposed algorithm is the base station or control unit receives the aggregated data with complete user privacy and their method affectively functions even the situation of differential attack on node during the data aggregation process. The security analysis and performance evaluation of this scheme meet the safety obligations of wireless sensor network. The results are comparatively has better outcome than the other methodologies.

Almeida et al. [30] have presented a technique called MDFU: Mass Distribution with Flow-Updating protocol using the Flow Updating concept to classic Mass Distribution in WSN. The experimental evaluation of the mass distribution based FU protocol becomes the first proof of its validation as well as the time required for convergence is an evident that the stochastic message loss which will produce very low overhead. The heuristic adjustment of MDFU provides the proportionality with the loss in message rate with fixed deviation. It has better results than the many of the flow based and mass distribution based protocols proposed for wireless sensor network. The working difference of this proposed work in comparisons with other related technologies are listed in Tab. I.

Reference	Technique	Type followed	Contributed parameters
[21]	Distributed Fault Tolerance Routing	Cluster/chain based	Cost
[22]	Privacy-preserving Data Aggregation for Fault Tolerance	Non-cluster based	User Data Encryption
[24]	Fault Influence Domain Analysis	Non-cluster based	Fault Tolerance
[28]	Energy efficient Maximum Disjoint Coverage	Non-cluster based	Delay, Energy consumption
[30]	Mass Distribution with Flow-Updating	Non-cluster based	data accuracy
[31]	S SVM-based failure detection and loss recovery	Cluster based	failure rate, Packet drop
ours	optimal data aggregation (ODA)	Cluster based	Energy consumption, delay, delivery ratio, data accuracy, network lifetime, failure rate, cost and data aggregation

Tab. I *Working difference comparison.*

3. Problem methodology and network model

3.1 Problem methodology

Kamalesh et al. [31] have developed a protocol called SFDLA: Support vector machine Failure Detection and Loss recovery Algorithm in which they have done the data aggregation by deploying the Support Vector Machine (SVM) for execution of faultfinding and loss recovery in wireless sensor network. The cluster head opted by the performance metrics as node connectivity, which have separated clusters and the nodes s based on their location information. The cluster member in each and every cluster would occupy as maximum node connectivity selected as by the cluster head. When the clustering node obtains the binary data's from the source node, it will recognize the node failures of the source location based on the received data by categorizing the faulty data and using SVMSFDLA technique minimizes the packet drop, delay, transmission overhead and increases reliability.

From the existing papers [21–30], the authors have concentrated only on to detect the faults rather than the corresponding loss recovery system. SFDLA technique [31] overcomes that problem by the combined contributions, but the employed SVM-based failure detection consumed much more power. They have utilized a parameter correlation coefficient for fault classification which is not suitable for critical network. The cluster formation and backup node selection is also not effectively described in SFDLA technique. Moreover, the designated standby node, at any time may turn into a malicious and surely affects the activities of the sensor node data's [31]. To overcome the said complications of the wireless sensor network, the proposed QoS concentrated optimal data aggregation (Q-ODA) technique become vital for the efficient sensor network. Most of the works considered the general parameters, energy and delay are the quality metrics, in this paper, include the additional parameters such as network lifetime, fault tolerance, and cost function for quality check.

3.2 Network model

We have considered the wireless sensor network which has homogeneous in nature and energy-constrained high density sensors nodes. They are distributed in random manner over the specific area shown in Fig. 1. In this proposed work the Rank computation is plays an important role based on which the clustering head is elected for this purpose. Our proposed algorithm has grouped the sensing nodes and collected its QoS metrics for Rank computation. The data is forwarded from source node to destination node or base station through the head (H) node also called aggregated node, that makes the communication to/from cluster takes place. The main role would be played by aggregated node to avoid the data redundancy. It has achieved through obtaining all the binary data information from nearby sensor nodes and grouped it together as a single data stream as well as it will be routed to BS through a single route as a single data bytes through which enormous energy is saved as well the redundancy of data is avoided which results in better network performance.

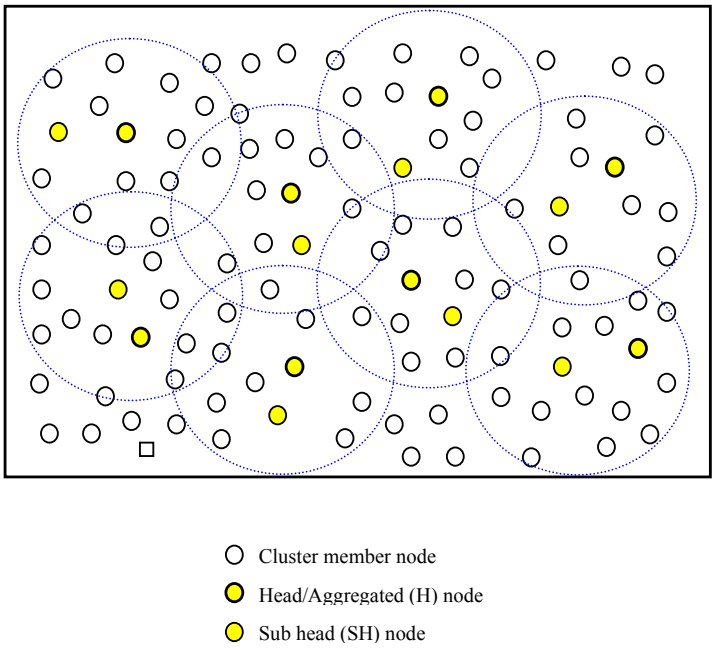

Fig. 1 *Network model of proposed Q-ODA technique.*

4. Proposed QoS concentrated optimal data aggregation technique

The proposed QoS concentrated Optimal Data Aggregation (Q-ODA) technique consists of two tires, first the clustering algorithm present in Section 4.1 and the fault detection described in Section 4.2.

4.1 Clustering using improved pair detection (IPD) algorithm

In WSN, the end user will receive the final aggregated data through the BS (Base Station). The main advantage of this aggregated data forwarding method brought many advantages into the network. Specifically it reduces a much more data transmission over head which results in much more energy savings in the wireless sensor network. The concept called cluster which is a group of sensing nodes on a specific geographic horizon and is much more helpful in data aggregation process. We may define Clustering as a group of nodes in connection with few mechanisms. The core need of WSN is its life term improvement which will be result of clustering since the lifetime of the WSN is act as a vital tool to analyse the performance of any WSN. Another handful merit of clustering is its scalable property along with energy expenditure minimizations in the wireless sensor network. Construction of the clustering also comprises the passing on the part to the node on the basis of their perimeters.

Clustering may cause abrupt packet drop in the network. Hence, it is highly important to overcome which is done in the proposed work with the implementation of Improved Pair Detection (IPD). IPD easily overcomes such a scenario, even the critical mission durations; in this implemented approach we have consider the $1+N$ path protections for a secured and reliable network. By using $1+N$ path protection strategies, the ingress node considers the $1+N$ disjoint paths to send similar copies of the burst. Hence this will secure the network from the fear of simultaneous link failures of N links.

This proposed buffer less IPD network has N sets of nodes which are considered with a maximum of T numbers of trunks for the given network. Considering an ingress/egress node, at the edge of the network, with M trunks (where $M \epsilon T$) to support F_M number of nodes with each and every cluster would supporting W number of wavelengths through themin which each wavelength is considered to support S sub-wavelength channels.

- Case 1: With distinct number of sub wavelength channels, each node is carrying different number of wavelengths. In which each trunk has different number of node.
- Case 2: With same number of sub wavelength channels and each node is carrying the same number of wavelengths. In which each trunk has same number of nodes.

Number of channels carried by M is given as,

$$C_M = \begin{cases} F_M \times W_M \times S_M & \text{full wavelength conversion scenario} \\ F_M \times W_M & \text{no wavelength conversion scenario} \end{cases} \quad \text{for case 1,} \quad (1)$$

$$C = \begin{cases} F \times W \times S & \text{full wavelength conversion scenario} \\ F \times W & \text{no wavelength conversion scenario} \end{cases} \quad \text{for case 2.} \quad (2)$$

For j pair of source and destinations in the given network, the route with minimum number of hops is used as the primary path P^j_{primary}. The route with minimum number of hops, excluding the primary path, is considered to be having a protected path of $P^j_{\text{protected}}$ for the considered Source-Destination pair in the considered Optical Burst Switching (OBS) network. Excluding the P^j_{primary} path, the first protected path is considered as $P^j_{\text{protected}_1}$. Similarly for the next protected path both the P^j_{primary} path and $P^j_{\text{protected}_1}$ path is excluded and the next path with least number of hops is considered as $P^j_{\text{protected}_2}$(Second protected path). This process will continue until $P^j_{\text{protected}_N}$total paths are evaluated. When a premium burst is received at the ingress node, it will send the same through its primary path towards the destination, P^j_{primary} and copies the same burst through each protected path $P^j_{\text{protected}_1}, P^j_{\text{protected}_2}, \ldots, P^j_{\text{protected}_N}$ simultaneously. Similarly when a regular burst arrives at the ingress node, the same will only be sent to the destination through the available primary path and no protected paths are employed in transmission.Poisson process with rates μ^j_{premium} and μ^j_{regular} are considered for burst arrival for both the type of users respectively.

The traffic offered by each source-destination pair, for $j \in \psi$ (where ψ is the total number pair in the network considered), the primary path experiences premium traffic $P^j_{\text{primary}} = \frac{\mu^j_{\text{premium}}}{\text{mean value}}$ and protected path experiences traffic of $P^j_{\text{protected}_N} = \frac{\mu^j_{\text{premium}}}{\text{mean value}}$.

Hence $P^j_{\text{primary}} = P^j_{\text{protected}_N} = \frac{\mu^j_{\text{premium}}}{\text{mean value}}$ (i.e. both the paths are considered to be mutually independent). Similarly for the same given number of source-destination pair, the regular traffic offered to the primary path is denoted as $P^j_{\text{regular}} = \frac{\mu^j_{\text{regular}}}{\text{mean value}}$. The burst reaches the ingress node at first will randomly chooses one wavelength among all the trust available at path between the source-destination pair. Only if all channels in all the given wavelengths are found busy, the burst is blocked to propagate through the given trunk. And if any of the channels in any of the wavelength is found free, it is employed for the propagation of the burst. When the centre positions are obtained precisely then the cluster is called as the best configured cluster. The assigned centre points are helpful in increasing the horizons of the clusters. The function of objective is derived from the Eq. (3).

$$F = \sum_{i=1}^{N} \sum_{j=1}^{M} \|x_{i,j} - C\|^2 . \tag{3}$$

In the Eq. (3) N represents the nodes count in the cluster, the total clusters count is represented by M. The i-th node belongs to the j-th cluster. For each and every search agent characterizes a set of M centres and it provides a hint that how fit was this agent. It is understandable that the least value of F is lies with the fittest search agent. In clustering, each and every cluster has its corresponding individual nodes of that particular horizon. It is the rule of Thumb is that the assignment is always done with the cluster which has the minimum distance with centre. The nodes are assigned to clusters based only on minimum distance to centre. After cluster formation, BS collects QoS metrics such as end-to-end delay, energy consumption, network lifetime, fault tolerance, and cost function from every sensor node in the network.

Each and every nodes current level of battery is at the time of initializing any simulation is called as Energy consumption (x_1) of that node, in simple terms it is also known as initial energy. During information interchange each node spares some energy for that particular interaction interchange, which we call transmit energy and receiving energy respectively. The basic quality of service parameter is the End-to-end delay (x_2) which is in simple terms defined as the average time duration in which the data packets reaches its destination which also comprises all kind of delays like route discovery process delay, packet waiting in queue delay. The time duration consumed by the network to finish a particular task is called Network lifetime (x_3). Complete draining of Battery, interferences due to environmental impacts and physical damage to sensing nodes are the common scenarios which cause the sensing node gets failed, which are not affect the overall task i.e. called Fault tolerance (x_4). The capacity to retain the sensor node activities even the scenarios of node failure encounter is called as Fault tolerance of a network in simple terms the capacity to work even failure of some nodes. Cost function (x_5)

represents the maximum number of paths required to satisfy the fault tolerance and reliability. In this proposed work, we have used the Rosenbrock function for Rank computation of nodes as follows:

$$F(x) = \sum_{i=0}^{3} \left[\left(\left(x_{i+1} - x_i^2 \right) (10 \times 10) \right) + (x_i - 1)(x_i - 1) \right]; i = 0, 1, \ldots, n \quad (4)$$

Now, select head (H), sub head (SH) node as

$$H = \max_i \{F(x)\}, \quad (5)$$

$$SH = \max_{i-1} (F(x)). \quad (6)$$

The final simplified format of the IPD (improved pair detection) has been displayed in the below Algorithm 1 in Detailed way

Algorithm 1 Clustering using improved pair detection (IPD) algorithm.

Input: number of populations, control variables
Output: cluster formation, head (H), sub head (SH) selection

Initialize the burst of populations
for each search **do**
  Compute C_M and C using Eq. 1.
  while the $C_M$not assigned C **do**
    Assign node to it is nearest C
  end while
end for
for $i = 1$ to n **do**
  for $j = 1$ to k **do**
    $X_1(P_i, C_j)$ = Compute (energy consumption between node i and j)
    $X_2(P_i, C_j)$ = Compute (delay between node i and j)
    $X_3(P_i, C_j)$ = Compute (network lifetime between node i and j)
    $X_4(P_i, C_j)$ = Compute (fault tolerance between node i and j)
    $X_5(P_i, C_j)$ = Compute (cost between node i and j)
    Calculate Rank (R) using Rosenbrock function (11)
  end for
end for
for $i = 1$ to n **do**
  $H = \max_i(F(x))$
  $SH = \max_{i-1}(F(x))$
end for
Return: Cluster, CH, and Level

4.2 Failure detection using multi-constraints moth-flame decision-making (MMFD) model

The moth-flame flies having a special character that it has a unique navigation technique which it exhibits during night time. In simple terms it will keep on a

particular angle with moon throughout its path during the long night time travels even the moth flame displays such characteristics with man-made light sources that it will try to maintain a straight path which results in spiral route because of artificial light source based on which the moth-flame optimizer (MFO) [32] is developed which is a metaheuristic optimization technique. To put those thing mathematically the matrix format is employed where X denotes a group or set of moth flies by keeping all into an array format which represented by X_a. After the naming the corresponding fitness values also stored in the same order. Lames are the second constituents of the algorithm and it has denoted by F matrix, the similar fashion F_a array represents their fitness values respectively.

In this paper, we have utilized this optimizer as decision making model with multi-constraints inputs related to the fault classification, named as MMFD model. The multiple constraints are loss rate, delivery time, and overhead at the time of data forward from one node to others. Here, we have utilized the same Rosenbrock function for fault level computation. When the output level is low the data forwarded node denotes as normal node (consider below $\tilde{0}.5$), otherwise, the node treat as faulty.

$$\text{Node} = \begin{cases} \text{normal}; & 0 < \mathrm{F}(x) < 0.5 \\ \text{faulty}; & \text{otherwise} \end{cases}. \tag{7}$$

In this proposed work, by deploying above technique any fault in the network is identified since the cluster head and members of the cluster may faces the situation of node fault or node failures. And it will compute the backup node from the neighbour of fault occurred node with minimum fault level. When the system identifies faults in H nodes then this technique automatically chooses the highest rank SH into H since SH will act as a backup node. The MMFD model starts with the initialization process and the problem of optimization derived from Eq. (6)

$$\text{MMFD (Multiconstraints Mot Flame Decision)} = (\mathrm{K},\mathrm{S},\mathrm{T}), \tag{8}$$

where K represents the function of random population $\mathrm{K} \to \{X, X_a\}$, S represents the fault level of data forwarding sensor nodes around search space $\mathrm{S} \to X$. The T represents the termination criteria $\mathrm{T} \to \{\text{True,False}\}$. Once the execution starts, the function S runs iteratively, till the time the termination criteria return back to true status. The behaviour of forwarding data will be updated with respect to a destination direction as follows:

$$X_i = \mathrm{s}(X_i, F_i), \tag{9}$$

where s represents the spiral function, i and j represents the i-th data, j-th destination respectively. The spiral's starting point and ending point should be placed such that the forwarded data as the origin and the placement of destination as terminal point. Where it has to be fixed like search space as dead end and all kind of fluctuations in the spiral path should not exceed it. MMFD's logarithmic value of a spiral is defined by Eq. (10)

$$\mathrm{s}(X_i, F_i) = |F_j - M_i| \, e^{bt} \cos(2\pi r) + F_j, \tag{10}$$

where $|F_j - M_i|$ specifies the remoteness of the i-th data aimed at the j-th node, and the denotation r indicates the random number and the symbol b represents

logarithmic spiral shape. The fault level of forwarded data with respect to destination is defined from Eq. (10). And also parameter r is used to define the number of faults present in the forwarded data. In order to upsurge the assortment of fault level computation against premature convergence and accelerate the convergence speed, we have improved the MMFD model by the Levy-flight. It has the projecting stuffs to upsurge the diversity of fault levels, consecutively, and this could create this excellently jump out of the local optimum. The new fault level of forwarded data updated as follows:

$$X_i^2 = X_i^1 + u \; sign \; [r - 0.5] \oplus \text{Levy} \; (\text{O}) \,, \tag{11}$$

where the uniform distribution is confirmed by the random parameters t and u, $sign\ [r - 0.5]$ is taken as 1, 0, and -1. Levy-flight computes the fault detection rate with the computed step lengths and the jumps conform to a Levy distribution as follows.

$$\text{Levy}\,(\text{O}) \approx \frac{\left[\frac{\Gamma(1+\text{O})\times\sin\left(\pi\times\frac{\text{O}}{2}\right)}{\Gamma\left(\left(\frac{(1+\text{O})}{2}\right)\times\text{O}\times 2^{\frac{(\text{O}-1)}{2}}\right)}\right]^{\frac{1}{\text{O}}} \times \mu}{|\nu|^{\frac{1}{\text{O}}}} \tag{12}$$

The complete fault detection is formalized in Algorithm 2 which gives detailed in sight about how the work is done.

μ and ν represent the standard normal distributions, $\text{O} = 0.5$, Γ represents the standard Gamma function. To sum up, the random walk with Levy-flight is used to add value point to the global search ability of this algorithm.

5. Simulation results

In this section, we have presented the NS-2 (Network Simulator 2) simulation results discussion of this proposed Quality Concentrated Optimal Data Aggregation (Q-ODA) technique and the results are compared with the performance of the existing methodologies.

5.1 Simulation parameter and setup

The simulation performed by randomly deployed sensor nodes with a size of $1000 \times 1000\,\text{m}^2$. The number of nodes is varied by 20, 40, 60, 80, 100, and 120. The radio range of sensor node is 50 m with the first order radio model. The BS is located in the left side corner of the sensor field. The data rate of each node is 512 bits/s. The initial energy level of each node is 10 J. The data packet size of each node is 64 bytes. The simulated traffic is constant bit rate (CBR). The simulation parameters are summarized in Tab. II. In the first scenario, we vary the number of nodes by 20, 40, 60, 80, 100, and 120. In the second scenario, we vary the number of faults by 5, 10, 15, 20, and 25 with fixed number of nodes as 120. The total simulation will take the time period of 1000 s. The performance of proposed Q-ODA technique is analysed with the prevailing techniques and specifically DFTR [21], PDAFT [22], FIDA [24], EMDC [28], MDFU [30] and SFDLA [31]. The comparisons are done in

Algorithm 2 Fault detection using multi-criteria moth-flame decision-making (MMFD).

Input: populations, search space, max. iteration
Output: fault classification and recovery

```
Initialize the loss rate, delivery time, and overhead of populations
Iteration = 0
for i = 0 to n do
    compute initial solution using Rosenbrock function Eq. (11)
    compute best, worst population
    for j = 0 to k do
        compute new solution using Eq. (17)
        if (new solution > initial solution) then
            solution = new solution
        else
            solution = initial solution
        end if
    end for
end for
if (iteration < max. iteration) then
    iteration = iteration +1
else
    Stop
end if
for x = 0 to m do
    if (solution > 5) then
        population = normal
    else
        population = fault
    end if
    for y = 0 to n do
        if (fault = population member) then
            select neighbour as backup population
        else
            backup = SH
        end if
    end for
end for
Return faulty detection and recovery
```

terms of energy consumption, delay, delivery ratio, data accuracy, network lifetime, failure rate, and cost of data aggregation.

5.2 Varying number of nodes

In this test, we have varied the number of nodes form 20 to 120 with the fixed faults as 5.

Parameter	Value
Size of the Network	1000×1000
Quantity of nodes	20, 40, 60, 80, 100, and 120
Traffic source	CBR
Radio range	50 m
Deployment type	Random model
Data rate of node	512 bits/s
Initial energy of node	10 J
Data packet size	64 bytes
Simulation time	1000 s

Tab. II *Simulation parameters for Scenario 1.*

Fig. 2 (Explains how long it takes to transfer single bit of information in binary form from one node to another node i.e., Delay while increasing number of Nodes) shows the delay comparison of proposed Q-ODA and existing techniques. The plot clearly depicts the delay of proposed Q-ODA technique is very low in terms of 42% than DFTR, 36 % than PDAFT, 37 % than FIDA, 34 % than EMDC, 35 % than MDFU and 41 % than SFDLA techniques. Fig. 3 (Explains percentage of data delivery rate from one node to another node while increasing number of Nodes) shows the packet delivery ratio comparison of proposed Q-ODA and

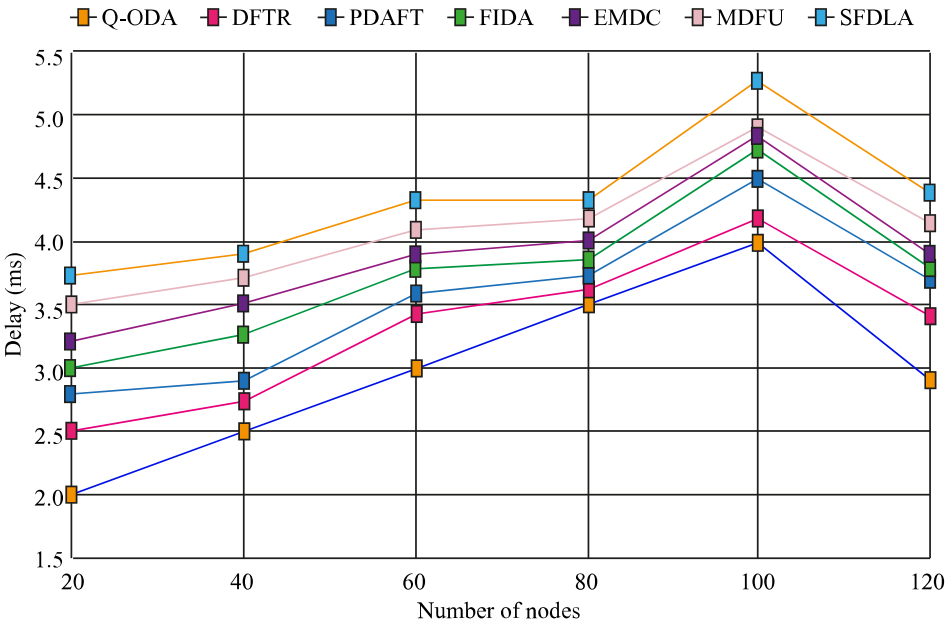

Fig. 2 *Delay comparison with varying nodes.*

existing SFDLA technique. The plot clearly depicts the packet delivery ratio of proposed Q-ODA technique is very high in terms of 32 % than DFTR, 30 % than PDAFT, 31 % than FIDA, 33 % than EMDC, 25 % than MDFU and 34 % than SFDLA techniques.

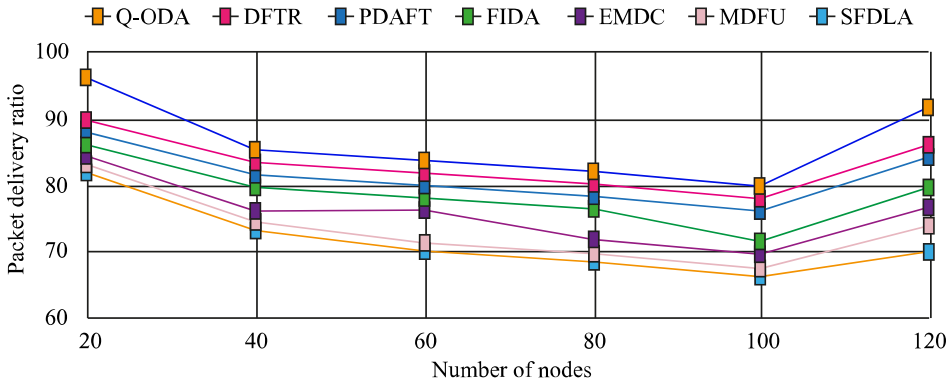

Fig. 3 *Packet delivery ratio comparison with varying nodes.*

Fig. 4 (Explains how much energy is consumed by individual nodes while engaging data transfer from one node to another node while increasing number of Nodes) shows the energy consumption comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the energy consumption of proposed Q-ODA technique is very low in terms of 43 % than DFTR, 22 % than PDAFT, 32 % than FIDA, 36 % than EMDC, 24 % than MDFU and 41 % than SFDLA technique.

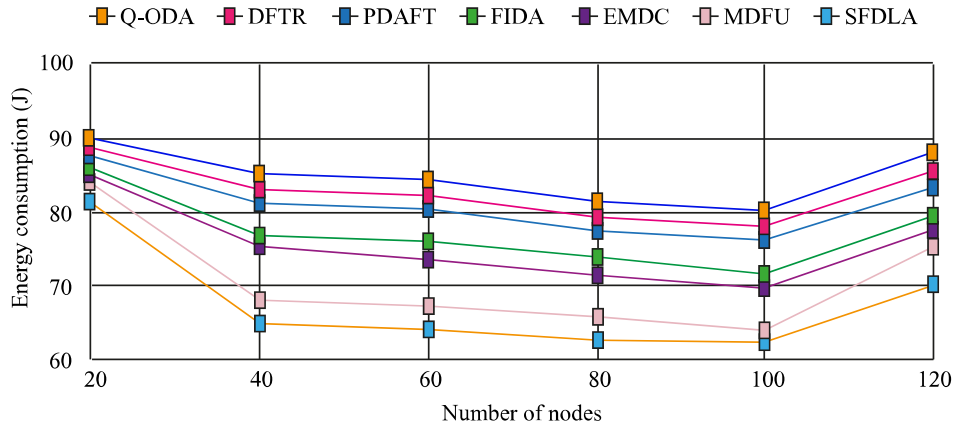

Fig. 4 *Energy consumption comparison with varying nodes.*

Fig. 5 (Explains Jitter presents during data transfer from one node to another node while increasing number of Nodes) shows the jitter comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the jitter of

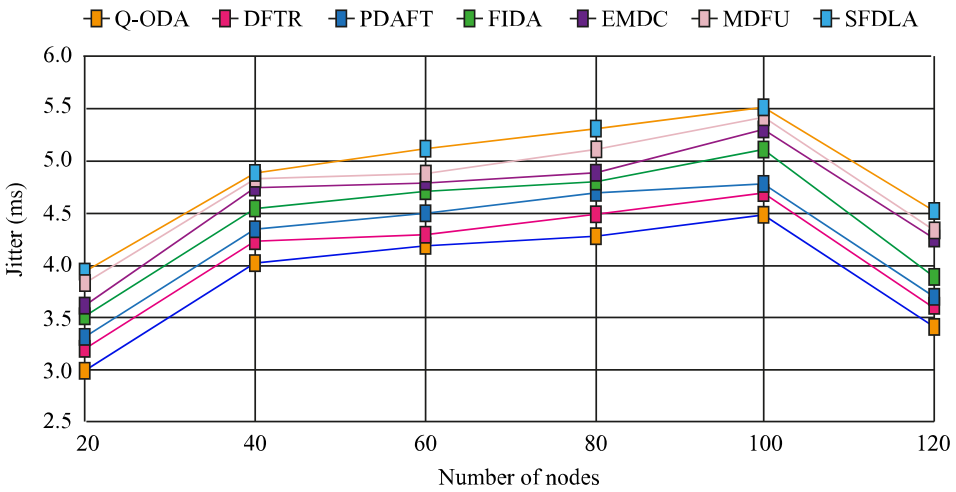

Fig. 5 *Jitter comparison with varying nodes.*

proposed Q-ODA technique is very low in terms of 32 % than DFTR, 33 % than PDAFT, 35 % than FIDA, 36 % than EMDC, 35 % than MDFU and 36 % than SFDLA technique.

Fig. 6 (Explains Calculate the battery left that is network lifetime of the deployed wireless sensor network while increasing number of Nodes) shows the net-

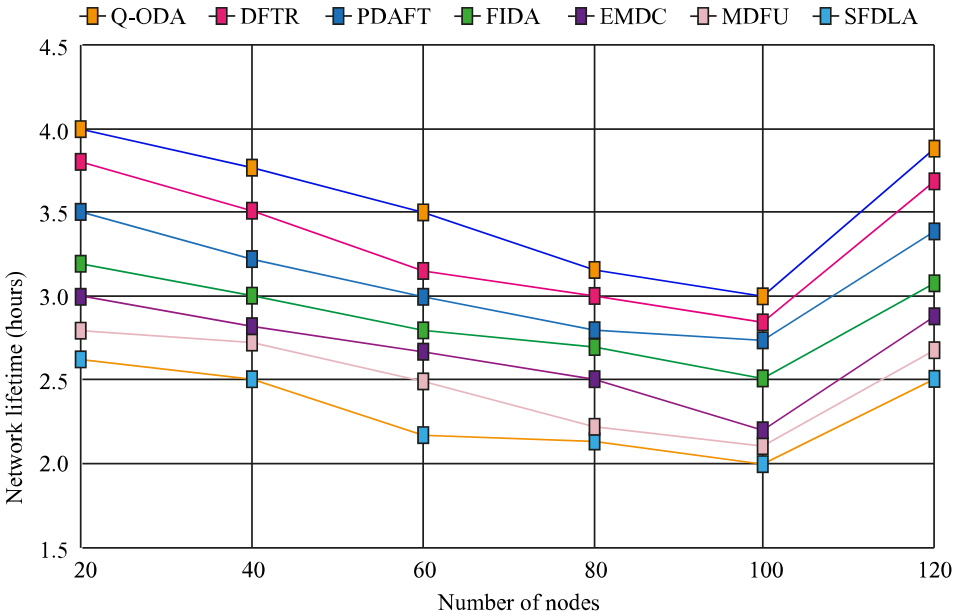

Fig. 6 *Network lifetime comparison with varying nodes.*

work lifetime comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the network lifetime of proposed Q-ODA technique is very high in terms of 23 % than DFTR, 25 % than PDAFT, 32 % than FIDA, 30 % than EMDC, 24 % than MDFU and 34 % than SFDLA technique. Fig. 7 (Explains Calculation of throughput of the network in simple terms estimation of delay present in the network while engaging the data transfer from one node another node) shows the throughput comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the throughput of proposed Q-ODA technique is very high in terms of 23 % than DFTR, 42 % than PDAFT, 42 % than FIDA, 46 % than EMDC, 44 % than MDFU and 31 % than SFDLA technique.

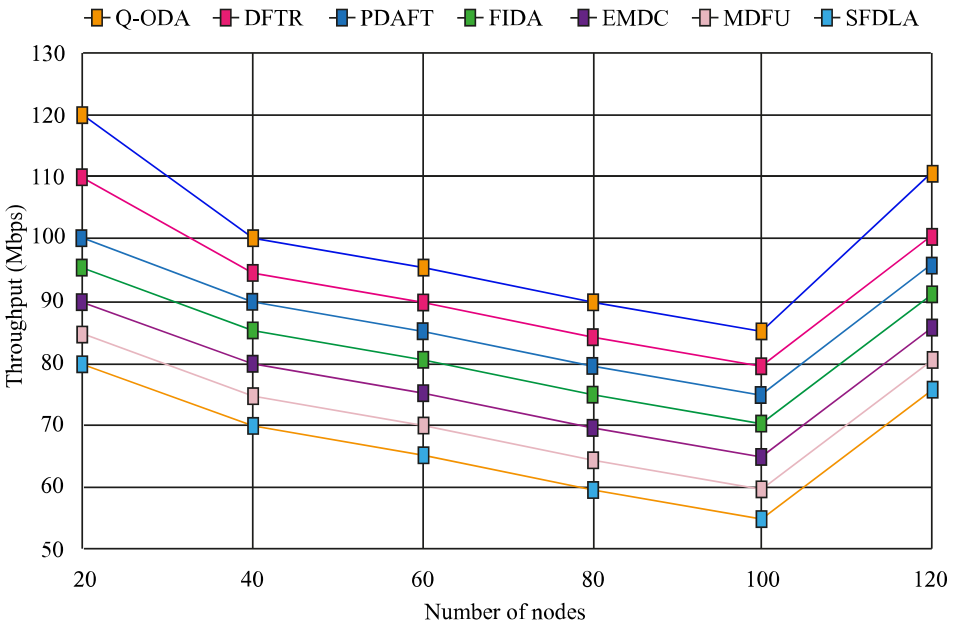

Fig. 7 *Throughput comparison with varying nodes.*

5.3 Varying number of faults

To the execution of the test the parameters mentioned in Tab. III have been incorporated. In this test, we vary the number of faults form 5 to 35 with the fixed node as 120. Fig. 8 (Explains how long it takes to transfer single bit of information in binary form from one node to another node i.e., Delay while increasing number of Fault Nodes) shows the delay comparison of proposed Q-ODA and existing techniques. The plot clearly depicts the delay of proposed Q-ODA technique is very low in terms of 22 % than DFTR, 23 % than PDAFT, 24 % than FIDA, 25 % than EMDC, 26 % than MDFU and 27 % than SFDLA techniques. Fig. 9 (Explains percentage of data delivery rate from one node to another node while increasing number of Fault Nodes) shows the packet delivery ratio comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the packet deliv-

Parameter	Value
Size of the Network	1000×1000
Quantity of faults	5, 10, 15, 20, and 25
Traffic source	CBR
Radio range	50 m
Deployment type	Random model
Data rate of node	512 bits/s
Initial energy of node	10 J
Data packet size	64 bytes
Simulation time	1000 s

Tab. III *Simulation parameters for Scenario 2.*

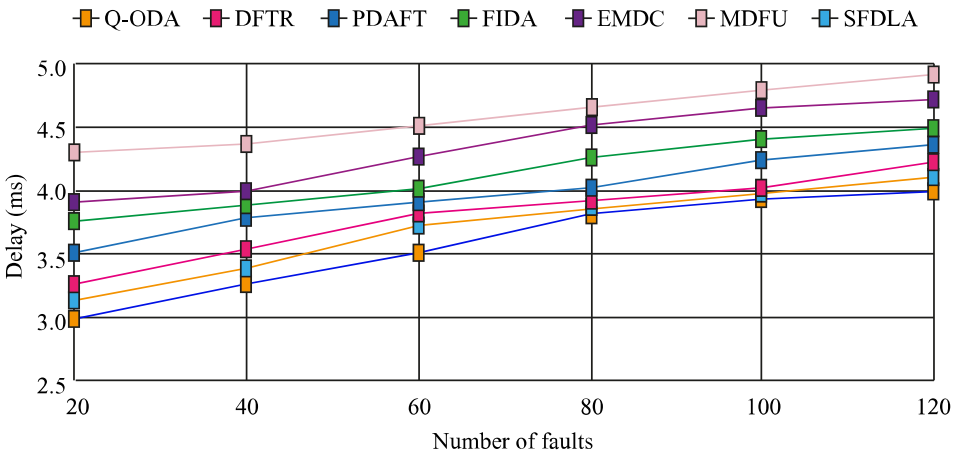

Fig. 8 *Delay comparison with varying faults.*

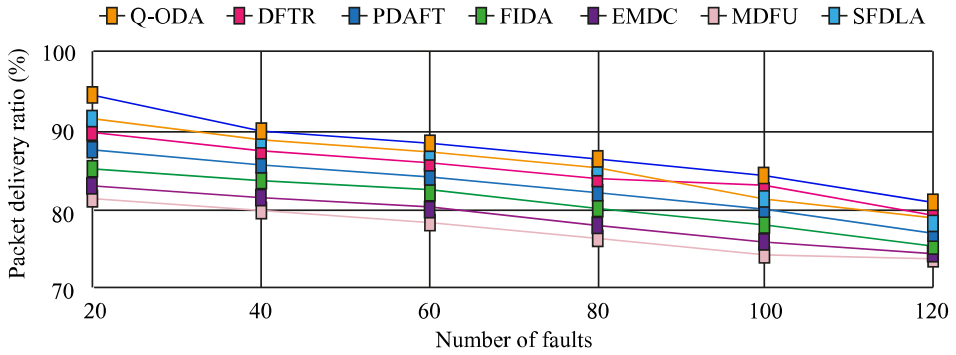

Fig. 9 *Delivery ratio comparison with varying faults.*

ery ratio of proposed Q-ODA technique is very high in terms of 31 % than DFTR, 32 % than PDAFT, 35 % than FIDA, 38 % than EMDC, 27 % than MDFU and 34 % than SFDLA techniques.

Fig. 10 (Explains how much energy is consumed by individual nodes while engaging data transfer from one node to another node while increasing number of Faulty Nodes) shows the energy consumption comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the energy consumption of proposed Q-ODA technique is very low in terms of 35 % than DFTR, 32 % than PDAFT, 26 % than FIDA, 28 % than EMDC, 29 % than MDFU and 37 % than SFDLA technique. Fig. 11 (Explains Jitter presents during data transfer from one node to another node while increasing number of Faulty Nodes) shows the jitter comparison of proposed Q-ODA and existing SFDLA technique. The

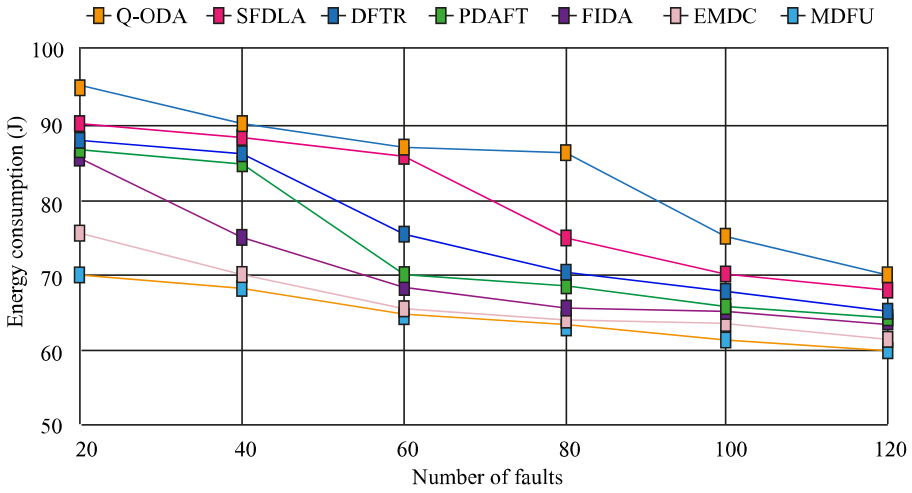

Fig. 10 *Energy consumption comparison with varying faults.*

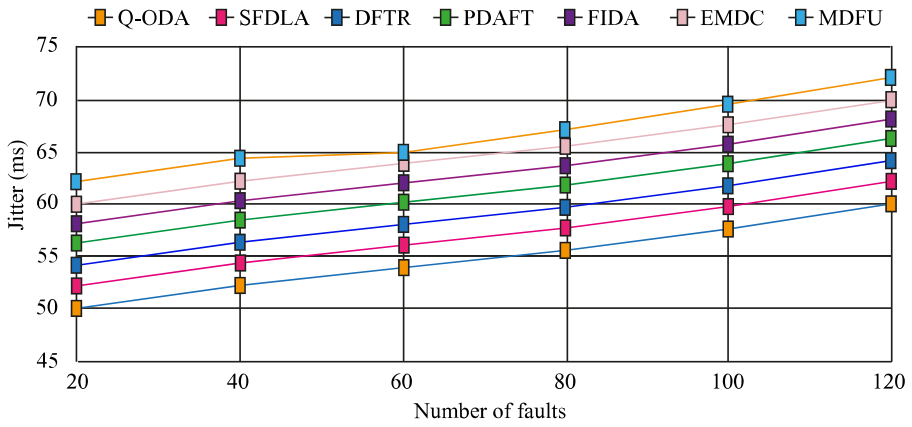

Fig. 11 *Jitter comparison with varying faults.*

plot clearly depicts the jitter of proposed Q-ODA technique is very low in terms of 32 % than DFTR, 33 % than PDAFT, 35 % than FIDA, 46 % than EMDC, 35 % than MDFU and 56 % than SFDLA technique. Fig. 12 (Explains Calculate the battery left that is network lifetime of the deployed wireless sensor network while increasing number of Faulty Nodes) shows the network lifetime comparison of proposed Q-ODA and existing SFDLA technique. The plot clearly depicts the network lifetime of proposed Q-ODA technique is very high in terms of 43 % than DFTR, 35 % than PDAFT, 42 % than FIDA, 40 % than EMDC, 44 % than MDFU and 45 % than SFDLA technique.

Fig. 13 (Explains Calculation of throughput of the network in simple terms estimation of delay present in the network while engaging the data transfer from one node another Faulty node) shows the throughput comparison of proposed Q-ODA

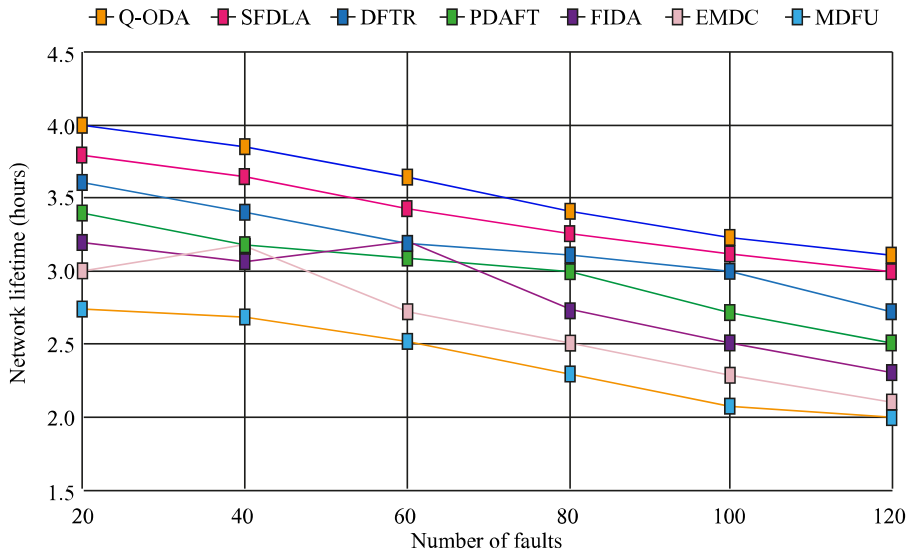

Fig. 12 *Network lifetime comparison with varying faults.*

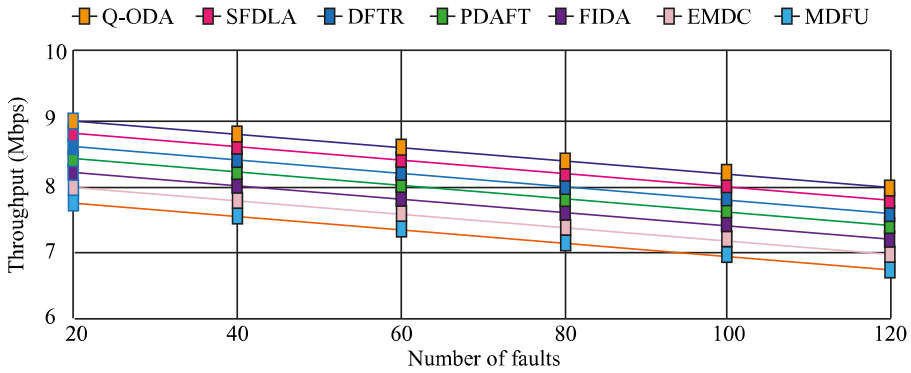

Fig. 13 *Throughput comparison with varying faults.*

and existing SFDLA technique. The plot clearly depicts the throughput of proposed Q-ODA technique is very high in terms of 34 % than DFTR, 36 % than PDAFT, 38 % than FIDA, 40 % than EMDC, 41 % than MDFU and 43 % than SFDLA technique.

6. Conclusion

This research work has proposed an optimal data aggregation (Q-ODA) with failure detection and loss recovery techniques. In this work, at first phase we have implemented an efficient aggregation scheme based on the Improved Pair Detection (IPD) algorithm which becomes the energy efficient clustering process with cluster head (H) and sub head (SH) selection in that the SH would act as backup node to recovers the lost data on the second phase, the multi-criteria moths-flame decision-making (MMFD) model is utilized to detects failure in the network. The result analyses have proved that the efficiency of this proposed Q-ODA technique in terms of delay, delivery ratio, energy consumption, jitter, network lifetime and throughput are outperformed over the existing techniques.

References

[1] BUTTYAN L., GESSNER D., HESSLER A., LANGENDOERFER P. Application of wireless sensor networks in critical infrastructure protection: challenges and design options. *Security and Privacy in Emerging Wireless Networks, IEEE Wireless Communications*, 2010, 17(5), pp. 44–49, doi: 10.1109/MWC.2010.5601957.

[2] CHOU C., IGNJATOVIC A., HU W. Efficient Computation of Robust Average of Compressive Sensing Data in Wireless Sensor Networks in the Presence of Sensor Faults, *IEEE Transactions on Parallel and Distributed Systems*, 2013, 24(8), pp. 1525–1534, doi: 10.1109/TPDS.2012.260.

[3] CLOUQUEUR T., SALUJA K., RAMANATHAN P. Fault tolerance in collaborative sensor networks for target detection. *IEEE Transactions on Computers*, 2004, 53(3), pp. 320–333, doi: 10.1109/tc.2004.1261838.

[4] HU K., IBRAHIM M., CHEN L., LI Z., CHAKRABARTY K., FAIR R. Experimental demonstration of error recovery in an integrated cyberphysical digital-microfluidic platform 2015. *IEEE Biomedical Circuits and Systems Conference (BioCAS), Atlanta, GA*, 2015, pp. 1-4, doi: 10.1109/BioCAS.2015.7348390.

[5] BALOUCHI F., BEVAN A., FORMSTOND R. *Detecting Railway Under-Track Voids using Multi-Train In-Service Vehicle Accelerometer 7th, IET Conference on Railway Condition Monitoring 2016 (RCM 2016)*, 2016. doi: 10.1049/cp.2016.1194.

[6] LIAN S., GUO X., GUO Z., ZHAO X. *Error Scene Restoration with Runtime Logs of Wireless Sensor Networks 2016 12th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*, 2016. doi: 10.1109/msn.2016.058.

[7] CHEN Z., SHIN K. Post-Deployment Performance Debugging in Wireless Sensor, *Networks 2009 30th IEEE Real-Time Systems Symposium*, 2009. doi: 10.1109/rtss.2009.47.

[8] TOBIAS N., BOLTON C., HESTER J., SITANAYAH L., SORBERSHOULDER ANGEL J. An Open Platform for Reprogramming Wayward Wireless Sensors. *IEEE Embedded Systems Letters*, 2016, 8(4), pp. 73–76, doi: 10.1007/s10878-010-9336-4.

[9] KRISHNAMACHARI B., IYENGAR S. Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks. *IEEE Transactions on Computers*, 2004, 53(3), pp. 241–250, doi: 10.1109/TDSC.2005.36.

[10] KUN SUN., PENGNING., WANG C. Fault-Tolerant Cluster-Wise Clock Synchronization for Wireless Sensor Networks. *IEEE Transactions on Dependable and Secure Computing*, 2005, 2(3), pp. 177–189, doi: 10.1109/TDSC.2005.36.

[11] QINGCHUN CHEN., KAM-YIU LAM., PINGZHI FAN. Comments on Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks. *IEEE Transactions on Computers,* 2005, 54(9), pp. 1182–1183, doi: 10.1109/TC.2005.140.

[12] DATTAA A. Fault-Tolerant Protocol for Energy-Efficient Permutation Routing in Wireless Networks. *IEEE Transactions on Computers*, 2005, 54(11), pp. 1409–0421, doi: 10.1109/TC.2005.172.

[13] TSANG-YI WANG., HAN Y., VARSHNEY P., PO-NING CHEN. Distributed fault-tolerant classification in wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 2005, 23(4), pp. 724–734, doi: 10.1109/JSAC.2005.843541.

[14] TSANG-YI WANG., HAN Y., BIAO CHEN., VARSHNEYA P. Combined decision fusion and channel coding scheme for distributed fault-tolerant classification in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 2006, 5(7), pp. 1695–1705, doi: 10.1109/TWC.2006.1673081.

[15] YU M., MOKHTAR H., MERABTI M. Fault management in wireless sensor networks. *IEEE Wireless Communications,* 2007, 14(6), pp. 13–19, doi: 10.4236/ijcns.2008.11008.

[16] CHATZIGIANNAKIS V., PAPAVASSILIOU S. Diagnosing Anomalies and Identifying Faulty Nodes in Sensor Networks. *IEEE Sensors Journal*, 2007, 7(5), pp. 637–645, doi: 10.1109/JSEN.2007.894147.

[17] RAJASEGARAR S., LECKIE C., BEZDEK J., PALANISWAMI M. Centered Hyperspherical and Hyperellipsoidal One-Class Support Vector Machines for Anomaly Detection in Sensor Networks. *IEEE Transactions on Information Forensics and Security*, 2010, 5(3), pp. 518–533, doi: 10.1016/j.aasri.2013.10.052.

[18] XIAOFENG HAN., XIANG CAO., LLOYD E., CHIEN-CHUNG SHEN. Fault-Tolerant Relay Node Placement in Heterogeneous Wireless Sensor Networks. *IEEE Transactions on Mobile Computing,* 2010, 9(5), pp. 643–656, doi: 10.1109/TMC.2009.161.

[19] FENG Y., TANG S., DAI G. Fault tolerant data aggregation scheduling with local information in wireless sensor networks. *Tsinghua Science and Technology*, 2011, 16(5), pp. 451–463, doi: 10.1016/s1007-0214(9)70065-7.

[20] CHAN S., WU H., TSUI K. Robust Recursive Eigen decomposition and Subspace-Based Algorithms with Application to Fault Detection in Wireless Sensor Networks. *IEEE Transactions on Instrumentation and Measurement*, 2012, 61(6), pp. 1703–1718, doi: 10.1109/TIM.2012.2186654.

[21] AZHARUDDIN M., JANAA P. A distributed algorithm for energy efficient and fault tolerant routing in wireless sensor networks. *Wireless Networks*, 2014, 21(1), pp. 251–267, doi: 10.1007/s11276-014-0782-2.

[22] CHEN L., LU R., CAO Z. PDAFT: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications. *Peer-to-Peer Networking and Applications*, 2014, 8(6), pp. 1122–1132, doi: 10.1109/jsen.2019.2895769.

[23] CHEN J., MA H., ZHAO D. Private data aggregation with integrity assurance and fault tolerance for mobile crowd-sensing. *Wireless Networks*, 2015, 23(1), pp. 131–144, doi: 10.1007/s11276-015-1120-z.

[24] WU Z., LU K., WANG X., CHI W. Topology-aware network fault influence domain analysis. Computers and Electrical Engineering, 2017, 57, pp. 266–280, doi: 10.1007/s11704-014-3503-1.

[25] CHERAGHLOU M., KHADEM-ZADEH A., HAGHPARAST M. Increasing Lifetime and Fault Tolerance Capability in Wireless Sensor Networks by Providing a Novel Management Framework. *Wireless Personal Communications*, 2016, 92(2), pp. 603–622, doi: 10.1007/s11277-016-3559-3.

[26] XU Z., SHI S., GU X., WANG Z., ZHU S. An Admission Control Method Based on QoS Constraint of BSN Traffic Aggregation. *International Journal of Distributed Sensor Networks*, 2016, 12(2), p. 5812658, doi: 10.1155/2016/5812658.

[27] ADIL MAHDI O., ABDUL WAHAB A., IDRIS M., ABU ZNAID A., AL-MAYOUF Y., KHAN S. *WDARS:* A Weighted Data Aggregation Routing Strategy with Minimum Link Cost in Event-Driven WSNs. *Journal of Sensors*, 2016, pp. 1–12, doi: 10.1155/2016/3428730.

[28] HENNA S. Energy Efficient Fault Tolerant Coverage. *Wireless Sensor Networks. Journal of Sensors*, 2017, pp. 1–11, doi: 10.1155/2017/7090782.

[29] GUAN Z., SI G. Achieving privacy-preserving big data aggregation with fault tolerance in smart grid. *Digital Communications and Networks*, https://www.sciencedirect.com/science/journal/23528648/3/4, November 2017, 3(4), doi: 10.1016/j.dcan.2017.08.005.

[30] ALMEIDA P., BAQUERO C., FARACH-COLTON M., JESUS P., MOSTEIRO M. Fault-tolerant aggregation: Flow-Updating meets Mass-Distribution. *Distributed Computing*, 2016, 30(4), pp. 281–291, doi: 10.1007/s00446-016-0288-5.

[31] KAMALESH S., GANESH KUMAR P. Data aggregation in wireless sensor network using SVM-based failure detection and loss recovery. *Journal of Experimental & Theoretical Artificial Intelligence*, 2016, 29(1), pp. 133–147, doi: 10.1080/0952813X.2015.1132262.

[32] MIRJALILI S., MIRJALILI S., LEWIS A. *Grey Wolf Optimizer. Advances in Engineering Software*, 2014, 69, pp. 46-61, doi: 10.1016/j.advengsoft.2013.12.007.

[33] MIRJALILI S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Systems*, 2015, 89, pp. 228–249, doi: 10.1016/j.knosys.2015.07.006.

[34] RAJA BASHA A., YAASHUWANTH C. Optimal Partial Aggregation Based Energy Delay Compromise Technique for Wireless Sensor Network. *IETE Journal of Research*, Taylor and Francis Publications, 2018, doi: 10.1080/03772063.2018.1464966.

[35] RAJA BASHA A., YAASHUWANTH C. Double Secure Optimal Partial Aggregation Using Trust Inference and Hybrid Syncryption Algorithm for Wireless Sensor Networks, Journal of Computational and Theoretical Nanoscience, 2018, 15(2), pp. 423–436(14), doi: 10.1016/j.dcan.2017.08.005.

[36] ADAM R.B. An Energy and Delay Aware Optimal Data Aggregation using Three Fold Algorithm for WSN. International Journal of Business Information Systems, *Inderscience Publications,* 2020, doi: 10.1504/IJBIS.2020.10014475.