MEDWELL PUBLICATIONS SINCE 2000
ASIAN JOURNAL OF
INFORMATION TECHNOLOGY
Journal Home Page: https://medwelljournals.com/journalhome.php?jid=1682-3915

Hybrid Optimized Multi Sink Network for Optimal Data Aggregation in Wireless Sensor Network Using Genetic Algorithm and Invasive Weed Optimization

Satyanarayana Mummana and Kuda Nageswara Rao
Department of CS&SE, Andhra University, 530003 Andhra Pradesh (AP), Visakhapatnam, India

Key words: Wireless Sensor Network (WSN), clustering, Cluster Head (CH) selection, Low Energy Adaptive Clustering Hierarchy (LEACH), Genetic Algorithm (GA), Invasive Weed Optimization (IWO)

Corresponding Author:
Satyanarayana Mummana
Department of CS&SE, Andhra University, 530003 Andhra Pradesh (AP), Visakhapatnam, India

Page No.: 193-202
Volume: 18, Issue 8, 2019
ISSN: 1682-3915
Asian Journal of Information Technology

Abstract: Wireless Sensor Network (WSN) is a network that is formed using many sensor nodes which are positioned inside an application based environment for monitoring the physical entities within a target area. A primary challenge in the organizing of such networks is its efficacy of energy. Clustering is found to be an efficient technique that can prolong the WSN lifetime. It includes the grouping of the sensor nodes into various clusters and also the electing of the Cluster Heads (CH) for the clusters. The CHs will collect data from their respective cluster's and their nodes to forward all aggregated data to the Base Station (BS). The CH selection is a Non-deterministic Polynomial (NP)-hard problem. This study proposes a very energy efficient CH algorithm for selection that has been based on the Genetic Algorithm (GA) and the Invasive Weed Optimization (IWO) algorithm. The sensor network and its lifetime will be extended using the power based clustering protocols. The operations in clustering will be optimized using the optimizations of swarm and their Evolutionary Algorithms (EA). Here, there is an improved IWO based upon the hybrid genetic (GA-IWO) has been presented. In this new algorithm, the inertial weight is adaptively adjusted for improving the speed of convergence and the weeds are multipled by means of selection and hybridization of the GA. This import of such hybrid genes will improve the t performance of weeds and will further reduce the likelihood of getting into the local optima. The results of the experiment prove that the method proposed can achieve better performance compared to the other methods.

INTRODUCTION

Wireless Sensor Network (WSN) is a type of energy constrained network and is formed using a number of the sensor nodes that will be powered by the batteries and the recharging or the replacement processes are quite difficult. The sensor nodes get used for the monitoring of the physical or the environmental conditions. In recent times, the development in technology like temperature, sound or motion and in the Micro Electronic Mechanical System (MEMS) and all the wireless communication based technologies are enabling invention of the tiny, the low power and the low cost along with the multi-functional smart sensor nodes found in the WSN. This type of a transmission of that of a finite information is supported using finite energy and the most critical factor will be the efficiency of energy that can prolong the lifetime of the network and that can balance the consumption of energy (Rana *et al*., 2015).

Aside from the limitations of the devices, the WSN also suffers from several other challenges like the multi node-to-node transmission, the data redundancy and its high unreliability as the sensors will be subject to some failure and physical damage. Such limitations as well as problems have presented several challenges in both the design and the deployment of the WSNs. There are many recent researches that have been carried out which attempt in solving both the design as well as the issues in implementation and some researchers have attempted to address this problem of limited energy using energy efficient solutions that can prolong the lifetime of the network. Such solutions also include the protocols of data routing and this will focus on the unreliable data transmission which can provide the solutions for increasing the reliability as well as the data availability like the multi-sink and the multi-path protocols (Bakr and Lilien, 2014).

Routing transmits the data within the network among nodes and their routing protocols for preforming routing. They choose an efficient path for reaching the target node and the network layer is found to be responsible for the implementing of routing of incoming data. Most source nodes will not be able to reach the destination owing to the range of transmission. Then intermediate sensor nodes that will forward the packets. As mentioned before the WSN will also have certain constraints like supply of energy, bandwidth and son on. Earlier, the routing protocols were designed for the WSN which were Low Energy Adaptive Clustering Hierarchy (LEACH), the Directed Diffusion (DD), the Adaptive Threshold-sensitive Energy Efficient sensor Network (APTEEN) and the Stateless Protocol for End-to-End Delay (SPEED). Such protocols generally focused on the consumption of energy and the protocol design that was tailored using the application scenario which is the backbone of this network. This WSN routing based protocols were classified on the basis of the mode of such functions, the network structure and finally the styles of participation for the sensor nodes (Anwar *et al*., 2015).

In case of a typical WSNs, the sensor nodes will be resource constrained and will avoid the overwhelming amounts of the network traffic and with an extensive work that is on the schemes of data aggregation that will eliminate the redundant data that can enhance the energy lifetime in the WSN. Data aggregation is that process of the several sensors that will collect the result of detection from another sensor and this data that is collected will have to be processed using a sensor for the reduction of transmission even before they have been transmitted to the sink of the BS. WSN contain 3 types of nodes the simple regular sensor nodes, the aggregator node and the querier (Dagar and Mahajan, 2013). The regular sensor nodes will tend to sense data packet from its environment to send it to aggregator nodes and will aggregate the data packet by means of using a certain aggregation function like the sum, the average, the count or the max min and will then send the aggregates result to its upper aggregator node and sometimes to the querier node and using data aggregation the robustness as well as the accuracy of the information that has been obtained by the whole network can be made. One more advantage of this is that it can reduce the load of traffic and also conserve the sensors and their energy (Maraiya *et al*., 2011).

Clustering is a very important technique that is energy efficient wherein the sensor nodes get organized into some groups called the clusters. These regular nodes that are in the cluster will be termed as the cluster members and there is a CH that is selected from among them. There have been two types of such traffic that are clustered within the WSN: the data transmission inside a cluster that is defined as the intra-cluster traffic as well as the data transmission among clusters that has been termed as the inter-cluster traffic. This cluster member will sense all the real world parameters and will transmit these sensed value to the CH. A CH receives as well as aggregates the data for removing the redundant data and will transmit all of the aggregated data directly to the CH or through an intermediate CHs (Yang *et al*., 2013). These cluster members will not send any data to the BS directly and only to the CH which in turn forwards it to the BS. The main advantages of such clustering are: the energy consumption will be reduced by means of improving the bandwidth utilization, a reduced overhead, an increased connectivity with a stabilized network topology along with a decreased delay with effective load balancing along with a reduction in size of its routing table. This CH that is near the BS will consume more energy and will drain out the energy faster than the CH that is farther from the BS. The CH will be close to the BS and is loaded with some heavy traffic owing to the intra-cluster traffic which is from its cluster members with data aggregation, along with an inter-cluster traffic which may be from the other CHs for the relaying BS data (Arjunan and Sujatha, 2017).

The problem of clustering in the WSN may be defined as Given a set of the G of the n nodes and with a sink node (or the BS), there is a randomly positioned one in the monitoring area. The issue will be to find the nodes of the CH to minimize the energy for all the non-CH nodes the regular (Matos *et al*., 2012). This will be a Non-deterministic Polynomial (NP)-hard combinatorial optimization problem, making it difficult to develop an algorithm for solving this. But for finding the optimal clustering is an NP-hard problem and so, the efficient heuristics will be required for finding a proper clustering and this includes the engineering domains for tackling the hard problems to show is competitiveness or superiority using meta-heuristics like the GA, the Ant Colony Optimization (ACO), the Particle Swarm Optimization (PSO) and so on. The GA and the IWO are efficient algorithms that are nature inspired which may be an ideal choice here that can escape from its local optima and have some quick convergence (Rao and Banka, 2017).

Literature review: Arumugam and Ponnuchamy (2015) had introduced another Energy Efficient LEACH (EE-LEACH) protocol that was used for data gathering. It had offered an energy-efficient routing that was used in WSN and was based on an effective ensemble of data with an optimal clustering. For this system, there was a CH that was elected for all the clusters for minimizing the dissipation of all the sensor nodes and also optimized its resource utilization. Such energy efficient routing was obtained by all the nodes that had a maximal residual energy. So, these highest residual nodes were chosen for forwarding the data to its BS. The results of the experiment proved that this proposed EE-LEACH was able to improve the lifetime of the network.

Corn and Bruce (2017) had further proposed the LEACH-Centered Cluster Head (the LEACH-CCH) to be the clustering algorithm and aimed at improving the lifetime of the WSN for the mobile sensor nodes. The LEACH-CCH was a modification of a LEACH algorithm that had been developed for all the stationary networks. There is an analysis of the consumption of energy for a LEACH algorithm that was presented for identifying the data transmission which is energy expensive for the node through its entire lifetime and the LEACH-CCH will bring down the energy that is expended at the time of an expensive data transmission. By means of predicting the future position of such sensor nodes and by restructuring the clusters in accordance to this there is an overall improvement that is seen for the network lifetime on being compared to LEACH.

Kumar and Kumar (2016) had proposed another energy efficient clustering mechanism that was based upon the Artificial Bee Colony (ABC) algorithm as well as the factional calculus which was to maximize an network energy along with the life time of the nodes by means of optimally choosing the cluster-head. This type of a hybrid optimization algorithm that is called, the multi-objective Fractional ABC (FABC) has been developed for controlling the rate of convergence for the ABC that has some newly designed function taken to be the objectives like the consumption of energy, the distance travelled and the delay in minimizing the objective. The results proved that this proposed FABC maximized energy along with the node life time compared to the currently existing protocols.

Rao *et al*. (2017) further proposed another energy efficient CH selection based algorithm that was further based upon the PSO known as the PSO-ECHS. This algorithm has been developed using an efficient scheme that had particle encoding as well as fitness function. For the purpose of energy efficiency of this PSO approach, various other parameters were considered by the researcher like the intra-cluster distance, the sink distance and finally the residual energy of these sensor nodes. The researcher further presented a cluster formation within which the non-CH sensor nodes tend to join their CHs that were based upon the derived weight function. This algorithm had also been tested very extensively on the different scenarios of the WSNs with a varying set of sensor nodes and Chs. In many such protocols of routing, the selection of the CH will be based on a random probability based equation. Bhatia *et al*. (2016) had proposed a Genetic Algorithm Distance Aware (GADA) LEACH routing that made use of that of an evolutionary GA that was for the purpose of improving the selection of the CH in the legacy of the LEACH routing protocol within the sensor networks. This concept of that of a relay node will be introduced and can act as the intermediary that is between the CH and the BS. The results of simulation had got support that this algorithm was efficient.

Using an optimum clustering in the WSNs which is an NP-hard problem, currently the bio-inspired metaheuristic approaches have been found to be extremely popular to solve them. Adnan *et al*. (2016) further presented another centralized energy-aware clustering algorithm that was in the WSN with a novel bio mimic Cuckoo Search (CS) algorithm. For this the cost function was defined using a goal of maximizing the lifetime of the network and for minimizing all the intra-cluster distance. This performance for the algorithm was evaluated using some well-known centralized as well as decentralized protocols and the results proved that this solution was able to enhance the lifetime of the network over that of the comparatives.

Zhou *et al*. (2017) had presented another new method for prolonging this network lifetime that was based on an improved Particle Swarm Optimization (PSO) algorithm, being a method of optimization. This proposed protocol had results that were better in terms of the distributed sensors along with a clustering system that was well-balanced. The results of simulation proved that this proposed protocol performs better than the other comparative protocols in the different scenarios.

Shankar *et al*. (2016) had proposed another hybrid of the HSA and the PSO algorithm for the energy efficient selection of the CH. This algorithm exhibited a high search efficiency in the Harmony Search Algorithm (HSA) and also the dynamic capability of that of the PSO which can improve the sensor nodes and their lifetime. The hybrid algorithm was further evaluated using the alive nodes, the number of the dead nodes, the throughput and its residual energy. This proposed hybrid HSA-PSO algorithm has proved that it had an improvement in that of the residual energy as well as the throughput by about 83.89% and about 29.00%, respectively, than that of a PSO algorithm. Clustering networks that are for minimizing the total distance will be an NP-hard problem. Prasad *et al*. (2017) had suggested the hybrid Differential Evolution using a Multi Objective Bee Swam Optimization (MOBSO-DE) for an efficient method of clustering. The process of CH selection has been based upon the communication energy and residual energy along with the energy constraint metrics. The simulation proved that there is a new MOBSO-DE method that had outperformed the LEACH as well as the MOBSO for its packet delivery ratio along with the network lifetime.

Rana *et al*. (2015) further discussed this performance of the Improved IWO algorithm (IIWO) along with its application for the coverage optimization within the WSNs. Firstly, based on the premise where the connectivity among all the nodes had been guaranteed it can establish a mathematical model for achieving a coverage of the objective area with the WSNs. The IWO algorithm had been used for that of an optimal deployment using a strong search performance. There is also a cubic mapping chaotic operator that had been introduced for enhancing the local search ability as well as the robustness with its gauss mutation operator that had been used for keeping the population as well as its diversity. The results proved that this proposed algorithm was able to show some fast speed of convergence, proper robustness and a strong data mining ability. So, it also had the problem solving ability in the WSNs.

MATERIALS AND METHODS

In case of the WSN data aggregation which is an effective method to save all of the limited resources. The prime goal of such data aggregation algorithms was to gather as well as aggregate the data in a manner that was energy efficient to ensure that the lifetime of the network has been enhanced. Flooding was used to disseminate the codes and update the changes. Here every node will broadcast a flooding packet and the costs are significant only if the keys of the neighbourhood are sued. The LEACH protocol, the GA, the IWO and the hybrid GA-IWO based methods have been discussed here.

Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol: The LEACH, proposed by Heinzelman has been a pioneering clustering approach to routing in the WSNs that has the basic objective of choosing the sensor nodes as their CH by means of rotation to ensure that there is a high dissipation of energy in communication using the BS (Liu, 2012). The operation of that of the LEACH will be broken into lots of various rounds in which each round will be separated into two different phases which are the set-up phase and a steady-state phase. In the former the clusters will be organised and in the latter it is delivered to its BS. In the set-up phase, every node will decide if the CH has to be used for the current node. The decision will be based upon the percentage that is suggested for this network and the actual number of times the node was able to be a CH. The decision was made by the node becoming the CH for its current round in case the number is lower than the one in the following threshold as in Eq. 1:

$$T(n) = \begin{cases} \dfrac{P}{1\text{-}P\left(r \mod \dfrac{1}{P}\right)}, & \text{if } n \in G \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

In which the P will be the desired percentage for the CH where the r will be the current round and G will be the set of the nodes that are not elected in the final 1/P rounds. Whenever, the node has been elected a CH it will broadcast the message of advertisement to that of the other nodes and based on the strength of the signal that is received, the other nodes will decide the cluster that it will join. At the time of a stead state phase and for its even distribution of sensor node energy (Fu *et al*., 2013), the sensor nodes will transmit the data to the CH and this CH will compress the data that arrives from the nodes which will be part of the respective cluster by sending either the aggregated or the fused packet to the BS. Aside from this, the LEACH makes use of a Time Division Multiple Access (the TDMA) or the Code-Division Multiple Access (the CDMA) Medium Access Control (the MAC) for reducing the inter-cluster as well as the intra-cluster collisions. Once a particular time is passed there will be the determined priori, this network will go back into its set-up phase again and will enter one more round of CH election.

Genetic Algorithm (GA): The GAs are those efficient stochastic optimization based search procedures which mimic all the adaptive processes of evolution in the natural systems and they are applied successfully to various NP hard problems like the design of the multiprocessor, the task scheduling, the optimization and the Travelling Salesman Problem (the TSP). The GAs will be useful in those problems that have some large and irregular search space in which there is a need for a global minimum. The methods that are traditional and gradient based will encounter problems in which the search space will be multimodal and stacked inside the local maxima. The GAs will suffer less here owing to the premature convergence problem (Njini and Ekabua, 2014).

The GA being an iterative approach that includes trial and error, aims at finding a global optimum and this is nature's equivalent of the evolution in time in which there are several members that are created. It can simulate the process of evolution and create one more pool of chromosomes in which each one will represent its typical solution with the intention of solving and also taking the steps as: Creating of a random population of the N chromosomes (the candidate solutions use in the population). The evaluation of the fitness function the f (x) of each of the chromosome x within the population. The generation of a new population by means of repeating these steps till such time there is a new population which reaches a population N:

Algorithm 1; Genetic algorithm:

Step 1: Choose two parent chromosomes which are from the population by giving preference to fitter chromosomes (the high f (x) values). Normally the fittest chromosome is copied to its next generation (this is "elitism")

Step 2: Using a given probability of crossover, cross over all the parent chromosomes in forming two offspring. (In case there is no crossover that is performed, the offspring will be a copy of such parents)

Step 3: Using a mutation probability, swap randomly two offspring genes

Step 4: Copying the new population over the existing one

Step 5: Copying all newly generated population over the existing one

Step 6: In case the loop termination condition has been satisfied, the best solution is returned to its current population

Step 7: Else repeat from step 2

Normally, the process goes on for a certain number of generations or until there is a standard deviation of that of the fitness which converges to zero (while the standard deviation begins to converge, chromosomes will get fitter and would have arrived at an ideal solution). Making an assumption that its initial population is large, the fitness being well defined it has arrived at a proper solution. Making an assumption as to the initial population being large enough the fitness being well defined it can reach a proper solution.

The GA will be categorized to be a global search based heuristic algorithm wherein there is some optimal solution which is estimated by means of generating some individuals. This consists of the processes like the focused fitness functions. The fundamental aspects of the GA have been explained (Norouzi *et al*., 2011).

Initialization: In the initial stage, the GA will start with one primary population that includes some random chromosomes that contain genes having a sequence of either 0s or 1s. in the subsequent stem, the biases of the algorithm and its population is produced using two methods: the steady-state GA and the generational GA. In the former either one or two of the members of the population will be replaced and in case of the generational GA that replaces all such individuals in every generation. The next one is adopted to ensure that the GA will keep all specified and qualified individuals from its current generation and will copy them into the new one.

Fitness: A fitness function has been defined for the GA to be a process of scoring for every chromosome based on its qualifications and the value will be a survival trait and more of its reproduction. Such a fitness function will be dependent on the problem and it may also be impossible to define. The individuals in nature will be authorized for being able to pass on to a subsequent generation that can determine the individuals and their fate.

Selection: At the time of each successive generation, there is a new population that is by means of choosing the members of the present generation for mating on the basis of its fitness. The fitter individuals will be chosen leading to a selection which is preferential and most functions will have an element that is designed stochastically for choosing a few less fit individuals for maintaining the diversity of population (Norouzi and Zaim, 2014). Among the various methods of selection, the Roulette-Wheel has been selected for distinguishing the suitable individuals with the probability as in Eq. 2:

$$P_i = \frac{F_i}{\sum_{j=i}^{n} F_i} \tag{2}$$

In which the F_i and the "n" denote the fitness chromosome along with the population size and based on the Roulette-Wheel, every individual will be assigned one value falling between 0 and 1.

Crossover: The main step which produces a new generation is the reproduction or the crossover process. There may be a simulation of the process of sexual reproduction for inheriting the traits that are transferred naturally. For generating new children, the process of crossover will choose the individuals as the parents based on the process of breeding.

The process continues until its desired size is obtained and generally the operations of crossover are developed for various aims. A simple method is the single-point where there is a random point for dividing the contribution as two parents. There are two children from one single set of parents and the bit sequence of that of the offspring will duplicate the one parent's bit based sequence till the crossover point. After this the bit sequence of that of the other parent is replicated to be the next part of its children.

Invasive Weed Optimization (IWO) algorithm: A numerical stochastic optimization algorithm that has been inspired from the weed colonization was introduced first by Mehrabian and Lucas in the year 2006. The weeds invade fields by means of dispersing seeds and they occupy the spaces and grow into some flowering weeds that are dispersed randomly to grow into weeds. This technique has proved to outperform the optimizers like the PSO, the FA and the Biogeography-Based Optimization (the BBO) and can also handle some other new problems of optimization in the WSN. Below are the properties that have been considered for the colonizing behaviour of weeds in the IWO (Saravanan *et al*., 2014):

- One finite number of the seeds being spread out in a search area
- Each seed within the search area will grow into one complete flowering plant producing seeds that are based on their fitness
- These produced seeds will be randomly distributed in the search area for growing the new plants

The process will so, continue till such time the maximum plant number is reached and the survival of these plants are dependent on their fitness. The process will continue till such time the maximum iteration number is reached and the fittest of the plants will be the near optimal solution. This process has been explained in detail below.

Initializing of a population: The population of the initial solutions will be distributed out in the k-dimensional problem space that has random positions.

Reproduction: Every member of population (the plants) will be permitted to produce seeds based on the lowest as well as the highest fitness in a manner in which the weed that has the worst fitness will produce the minimal number of seeds and the weed that is best fitted to follow the seed production. In case of the minimization-type objectives, a formula of weeds for producing seeds will be Eq. 3:

$$\text{Number of seeds} = \frac{f_{max}-f}{f_{max}-f_{min}}(S_{max}-S_{min})+S_{min} \qquad (3)$$

In which f denotes the current weed and its fitness, f_{max} and f_{min} the maximum and its least fitness and S_{max} and S_{min}, respectively representing a maximum as well as a least value of the weed.

Spatial dispersal: The seeds that are generated once there is a reproduction that is distributed over a d dimensional space that is by the distributed random numbers that has a zero mean and a varying variance. This will further ensure that the seed generated is produced closer to the parent weed and the result of this will be that there is a local search for each such plant. At the same time the standard deviation (σ) for this random function will be decreased at the time of the consequent iterations (Sharma and Kumar, 2018).

In case the $\sigma_{initial}$ and σ_{final} denote its initial as well as the final standard deviation and the m will be its non-linear modulation index and the itr_{max} and itr will be the maximum number of such iteration and its current iteration where for a certain iteration the standard deviation (σ_{itr}) will be calculated in accordance to Eq. 4:

$$\sigma_{itr} = \frac{(itr_{max}-itr)}{itr_{max}}(\sigma_{initial}-\sigma_{final})+\sigma_{final} \qquad (4)$$

Competitive exclusion: In case a plant does not produce any progeny, then it can vanish or else take over earth. It is evident that the plant that is fitter will reproduce more seeds than that of a plant that is less fit. Owing to this, once there are some iterations there is a maximum level that is achieved by the plants in this colony. Once, it reaches there is a maximum level that is achieved using the plants and once this is reached (P_{max}), there is a procedure that is activated for the purpose of excluding plants that have deprived fitness within the generation.

The procedure of elimination will work as: where there is a maximum weed production and all seeds have identified the position within the search area, they tend to get ranked along with their parents. This mechanism of elimination will be made by means of eliminating the lower and the fitted weeds for achieving the population within this colony. This way, the plants that have a better level of fitness will be permitted to replicate and the population control mechanism for controlling offspring will be applied to the final part of its run.

Proposed GA-IWO algorithm: Here, for this research, making use of the best among the GA features and the IWO a hybrid algorithm for the purpose of solving clustering of the NP-hard problem. The GA, being an adaptive and a computational process that is modelled based upon the mechanics of the genetics systems has got some good features that are as: it will operate using the codes of variables. Searches for the optimal points from the variable points and the search has been directed by means of paying off the information without any need for the derivative of such information. There are several scholars that find GAs to have some better ability in finding optimal solutions compared to certain heuristic algorithms like the Tabu Search (TS) algorithm, the Simulated Annealing (SA) and the ant algorithm in certain situations. It does have some drawbacks like discarding the previous knowledge as soon as there is a change in population which will be independent of the fitness of the parents of the offspring. Therefore, there was an attempt to hybrid the GA using the IWO for enhancing the performance. The IWO, being a population-based EA, also has certain interesting specifications like the creation of offspring that is based on fitness of the parents and the increasing size (Atabaki *et al*., 2017).

It has been accepted generally that any of the GA based attempts to have a solution to the problems will have these basic components: the chromosome representation, the production of initial population, the evaluation of function rating and their solutions in relation to fitness, process of selection and their genetic operators (population size, evaluation and function rating) with the values of their parameters (the crossover, mutation and the reproduction). Mutation is that operation of the GA methods that are used for exploring all solutions and for applying the tools to the new areas and to avoid the drop in their local optimum solutions. In nature, the evolution is normally determined by its natural selection and also the encoding of its genetic information which is done in such a manner that is can admit the asexual reproduction that can result in the offspring which are genetically identical. This improved IWO based upon the hybrid GA that is a combination of the crossover and the mutation of the GA and using the cross factor that will come out of a solution based set which is similar in its natural evolution and also the same than its previous population generation that can adapt to this environment and the search of its global optimal solution (Yin *et al*., 2012).

In case of the cross factor method, half of the particles are selected that have a fitness value that is higher and can directly move into its next generation. This will at the same time make use of the fitness and the good of the first half of the positon of the particle along with that of the speed vector that can replace the fitness of the lower half of such particles and can keep the rest of the vector that corresponds to the individual extreme without any change. In case of the cross mechanism, the h half after particles of the as for the cross factor random and its combination pairing, there are the same operations of crossover that can produce the offspring and compare this to the father generation, the half particle having a fitness value that is better to get into its subsequent generation. So, such a cross will be able to increase the particles and their diversity by jumping out of its local optimum and at the same time can increase the speed of convergence. But most of the variants of the IWO will not avoid radically the IWO and their disadvantages. For this research, a clustering strategy was introduced and is deployed even before the reproduction for dispersing the solution to ensure that there are new individuals that can locate the regions that will be able to avoid the over-explored as well as the premature convergence. This clustering algorithm will divide the individuals into various regions in accordance to the distance among each individual as well as the cluster centres. After this, a certain number of individuals that are found to be the fittest and is based upon the cluster scale is chosen and the seed number of every selected weed that is chosen is distributed. For this method, the standard deviation value will be based on its statistical information that is calculated using the fittest individuals in every cluster. This way the standard deviation is accurate and this is for the actual value of its problem. The consumption of calculation can go up to a small extent and there is no increase in the consumption of time. Based on the results of the experiment, there was a significant improvementin the performance (Ren *et al*., 2016). This proposed GA-IWO algorithm has been described as:

Algorithm 2; Flowchart for proposed GA-IWO algorithm:

1. Generate random plants of N_0 individuals from
   the set of feasible solutions
2. i =: 1
3. do

a. Compute maximum and minimum fitness in the colony
b. For each individual w ε W
i. Compute the number of seeds for w
   corresponding to its fitness
ii. Randomly select the seeds from the feasible solutions
   around the parent plant (w) in a neighborhood with
   normal distribution, the seed number
iii. Add the generated seeds to the solution set, W
iv. For that parent plant whose seeds number
    is limited to zero, select corresponding number
    of generated seeds to do hybrid operation
    Seed(x) = Px Parent(x) + (1.0 - p) x Parent(x)
    P is random value between 0 and 1
Add the generated seeds to the solution set, again
c. If total number exceeds p max
i. Sort the population N in descending order of their fitness
ii. Truncate population of weeds with smaller fitness until N = Pmax
d. i = i+1
4. Repeat 3 until the maximum number of iterations

The flowchart for proposed GA-IWO algorithm as shown in Algorithm 1.

RESULTS AND DISCUSSION

In this study, the flooding and GA-IWO methods are used. The number of clusters formed, average end to end delay, average packet loss rate and average number of nodes alive as shown in Fig. 1-5 and Table 1-4. Experiments are carried using number of sinks such as 1-6 and 6 and number of rounds such as 0-800, respectively.

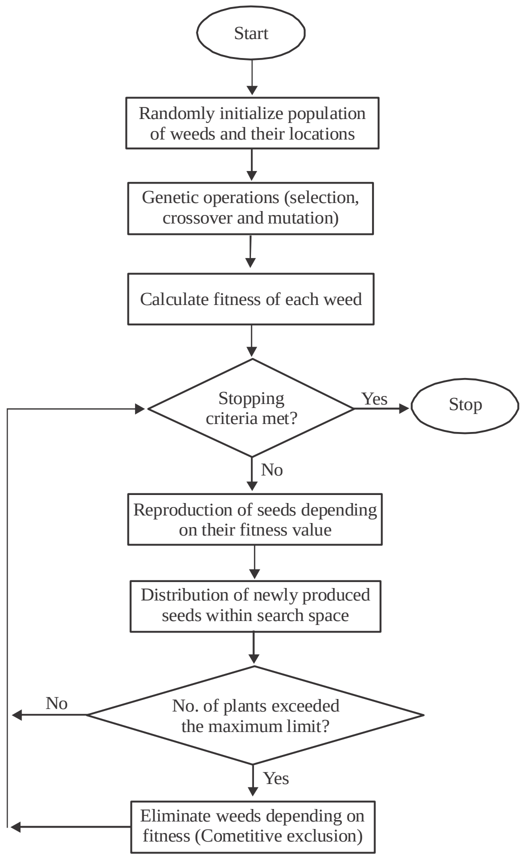

Fig. 1: Flowchart for proposed GA-IWO algorithm

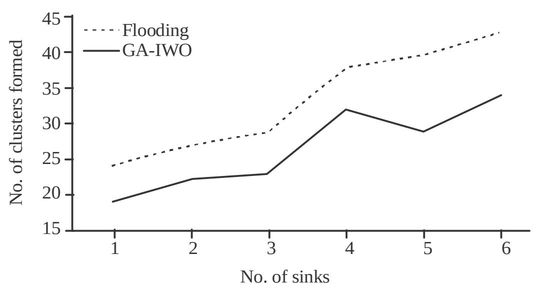

Fig. 2: Number of clusters formed

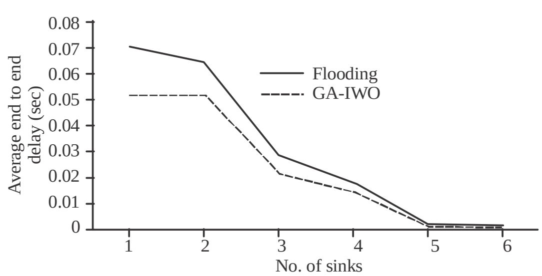

Fig. 3: Average end to end delay

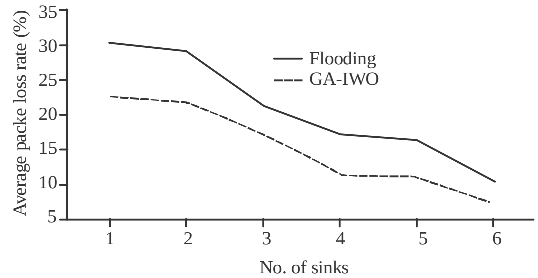

Fig. 4: Average packet loss rate

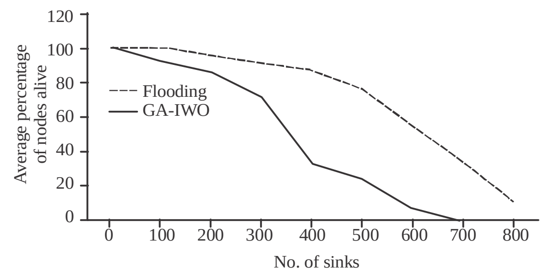

Fig. 5: Average number of nodes alive

Table 1: Number of clusters formed

No. of sinks	Flooding	GA-IWO
1	19	24
2	22	27
3	23	29
4	32	38
5	29	40
6	34	43

Table 2: Average end to end delay

No. of sinks	Flooding	GA-IWO
1	0.07032	0.05207
2	0.06445	0.05168
3	0.02906	0.02186
4	0.01782	0.01427
5	0.0018	0.00136
6	0.00173	0.0014

Table 3: Average packet loss rate

No. of sinks	Flooding	GA-IWO
1	30.31	22.5
2	29	21.56
3	21.27	17.18
4	17.16	11.26
5	16.12	10.98
6	10.24	7.24

Table 4: Average number of nodes alive

No. of rounds	Flooding	GA-IWO
0	100	100
100	92	100
200	85	95
300	71	91
400	33	87
500	24	76
600	7	54
700	0	35
800	0	11

From Fig. 2, it can be observed that the GA-IWO has higher number of clusters formed by 23.25% for 1 number of sinks by 20.4% for 2 number of sinks by 23.07% for 3 number of sinks by 17.14% for 4 number of sinks by 31.88% for 5 number of sinks and by 23.37% for 6 number of sinks when compared with flooding.

From Fig. 3, it can be observed that the GA-IWO has lower average end to end delay by 29.82% for 1 number of sinks by 21.99% for 2 number of sinks by 28.27% for 3 number of sinks by 22.12% for 4 number of sinks by 27.84% for 5 number of sinks and by 21.08% for 6 number of sinks when compared with flooding.

From Fig. 4, it can be observed that the GA-IWO has lower average packet loss rate by 29.57% for 1 number of sinks by 29.43% for 2 number of sinks by 21.27% for 3 number of sinks by 41.52% for 4 number of sinks by 37.93% for 5 number of sinks and by 34.32% for 6 number of sinks when compared with flooding.

From Fig. 5, it can be observed that the GA-IWO has higher average number of nodes alive by 8.33% for 100 number of rounds by 11.11% for 200 number of rounds by 24.69% for 300 number of rounds by 90% for 400 number of rounds by 104% for 500 number of rounds and by 154.09% for 600 number of rounds when compared with flooding.

CONCLUSION

Data gathering is that common at the same time critical operation in several applications of the WSN and the aggregation of data with the hierarchical mechanism are the techniques that are used widely. The mechanisms of clustering will be found to be effective for the management of this type of a high population of the nodes and can also help in the reduction of the energy consumption of the nodes. For this research a proposed CH selection algorithm that is based on the GA, the IWO and the GA-IWO which is energy efficient has been proposed. The GAs are found to be efficient procedures of search with stochastic optimization that are applied. The IWO is that continuous and stochastic numerical algorithm which has been inspired from the weed colonization. The IWO has been showing some successful results in practical applications and a simple as well as modified version of this integrated algorithm that exploits all the features in the IWO and the EA. The novelty feature of this research (the GA-IWO) will be the use of the diversity which is for activating its crossover operator. This crossover operator will be devised for improving the capability of the global capability and also for enhancing the capacity of escaping made from the local minimum. The results have proved that the GA-IWO has a much higher number of such clusters that are formed by about 23.25% for the 1 number of sinks by about 20.4% for the 2 number of sinks by about 23.07% for the 3 number of sinks by about 17.14% for the 4 number of sinks by about 31.88% for the 5 number of sinks and finally by about 23.37% for the 6 number of sinks on being compared with that of flooding.

REFERENCES

Adnan, M.A., M.A. Razzaque, M.A. Abedin, S.S. Reza and M.R. Hussein, 2016. A Novel Cuckoo Search Based Clustering Algorithm for Wireless Sensor Networks. In: Advanced Computer and Communication Engineering Technology, Sulaiman, H., M. Othman, Y. Rahim and N. Pee, (Eds.). Springer, Berlin, Germany, ISBN:978-3-319-24582-9, pp: 621-634.

Anwar, R.W., M. Bakhtiari, A. Zainal and K.N. Qureshi, 2015. A survey of wireless sensor network security and routing techniques. Res. J. Appl. Sci. Eng. Technol., 9: 1016-1026.

Arjunan, S. and P. Sujatha, 2017. A survey on unequal clustering protocols in wireless sensor networks. J. King Saud Univ. Comput. Inf. Sci., 1: 1-14.

Arumugam, G.S. and T. Ponnuchamy, 2015. EE-LEACH development of energy-efficient LEACH protocol for data gathering in WSN. EURASIP. J. Wirel. Commun. Netw., 2015: 1-9.

Atabaki, M.S., M. Mohammadi and B. Naderi, 2017. Hybrid genetic algorithm and invasive weed optimization via priority based encoding for location-allocation decisions in a three-stage supply chain. Asia Pac. J. Oper. Res., 34: 1-44.

Bakr, B.A. and L.T. Lilien, 2014. Comparison by simulation of energy consumption and WSN lifetime for LEACH and LEACH-SM. Procedia Comput. Sci., 34: 180-187.

Bhatia, T., S. Kansal, S. Goel and A.K. Verma, 2016. A genetic algorithm based distance-aware routing protocol for wireless sensor networks. Comput. Electr. Eng., 56: 441-455.

Corn, J. and J.W. Bruce, 2017. Clustering algorithm for improved network lifetime of mobile wireless sensor networks. Proceedings of the 2017 International Conference on Computing, Networking and Communications (ICNC), January 26-29, 2017, IEEE, Santa Clara, California, USA., ISBN:978-1-5090-4589-1, pp: 1063-1067.

Dagar, M. and S. Mahajan, 2013. Data aggregation in wireless sensor network: A survey. Intl. J. Inf. Comput. Technol., 3: 167-174.

Fu, C., Z. Jiang, W.E.I. Wei and A. Wei, 2013. An energy balanced algorithm of LEACH protocol in WSN. Intl. J. Comput. Sci. Issues, 10: 354-359.

Kumar, R. and D. Kumar, 2016. Multi-objective fractional artificial bee colony algorithm to energy aware routing protocol in wireless sensor network. Wirel. Networks, 22: 1461-1474.

Liu, X.X., 2012. A survey on clustering routing protocols in wireless sensor networks. Sensors, 12: 11113-11153.

Maraiya, K., K. Kant and N. Gupta, 2011. Wireless sensor network: A review on data aggregation. Int. J. Sci. Eng. Res., 2: 1-6.

Matos, V.D.O, J.E.C. Arroyo, A.G. Dos Santos and L.B. Goncalves, 2012. An energy-efficient clustering algorithm for wireless sensor networks. Intl. J. Comput. Sci. Network Secur., 12: 6-15.

Njini, I. and O.O. Ekabua, 2014. Genetic algorithm based energy efficient optimization strategies in wireless sensor networks: A survey. Adv. Comput. Sci. Intl. J., 3: 1-9.

Norouzi, A. and A.H. Zaim, 2014. Genetic algorithm application in optimization of wireless sensor networks. Sci. World J., 2014: 1-15.

Norouzi, A., F.S. Babamir and A.H. Zaim, 2011. A New Clustering Protocol for Wireless Sensor Networks Using Genetic Algorithm Approach. Wireless Sens. Netw., 3: 362-370.

Prasad, D.R., P.V. Naganjaneyulu and K.S. Prasad, 2017. A hybrid swarm optimization for energy efficient clustering in multi-hop wireless sensor network. Wirel. Pers. Commun., 94: 2459-2471.

Rana, S., A.N. Bahar, N. Islam and J. Islam, 2015. Fuzzy based energy efficient multiple cluster head selection routing protocol for wireless sensor networks. Intl. J. Comput. Network Inf. Secur., 4: 54-61.

Rao, P.S. and H. Banka, 2017. Energy efficient clustering algorithms for wireless sensor networks: Novel chemical reaction optimization approach. Wirel. Networks, 23: 433-452.

Rao, P.S., P.K. Jana and H. Banka, 2017. A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks. Wirel. Networks, 23: 2005-2020.

Ren, Z., S. Huang, C. Sun and Y. Liang, 2016. Improved invasive weed optimization based on clustering strategy. Proceedings of the 2016 International Conference on Chinese Control and Decision (CCDC), May 28-30, 2016, IEEE, Yinchuan, China, ISBN:978-1-4673-9715-5, pp: 4810-4815.

Saravanan, B., E.R. Vasudevan and D.P. Kothari, 2014. Unit commitment problem solution using invasive weed optimization algorithm. Intl. J. Elect. Power Energy Syst., 55: 21-28.

Shankar, T., S. Shanmugavel and A. Rajesh, 2016. Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks. Swarm Evol. Comput., 30: 1-10.

Sharma, G. and A. Kumar, 2018. Fuzzy logic based 3D localization in wireless sensor networks using invasive weed and bacterial foraging optimization. Telecommun. Syst., 67: 149-162.

Yang, Y., C. Lai and L. Wang, 2013. An energy-efficient clustering algorithm for wireless sensor networks. Proceedings of the 2013 10th IEEE International Conference on Control and Automation (ICCA), June 12-14, 2013, IEEE, Hangzhou, China, ISBN:978-1-4673-4707-5, pp: 1382-1386.

Yin, Z., M. Wen and C. Ye, 2012. Improved invasive weed optimization based on hybrid genetic algorithm. J. Comput. Inform. Syst., 8: 3437-3444.

Zhou, Y., N. Wang and W. Xiang, 2017. Clustering hierarchy protocol in wireless sensor networks using an improved PSO algorithm. IEEE. Access., 5: 2241-2253.