Shashi BHUSHAN[1], Svitlana G. ANTOSHCHUK[2], Pavlo O. TESLENKO[2], Vasyl V. KUZMENKO[3], Waldemar WÓJCIK[4], Orken MAMYRBAEV[5], Ulzhalgas ZHUNISSOVA[6]

School of Computer and Information Science (1),Odesa National Polytechnic University (2), Community Higher Education Establishment, Kherson Academy of Life-Long Education (3), Lublin University of Technology (4), Institute Information and Computational Technologies CS MES RK (5), Astana Medical University (6)

doi:10.15199/48.2019.04.24

Hybrid approach to energy efficient clustering for heterogeneous wireless sensor network using biogeography based optimization and k-means

Abstract. *The paper presents the proposed protocol a hybrid approach is applied for clustering of sensor networks combining BBO and K-means algorithm. The performance of the protocol is compared with SEP, IHCR and ERP in terms of stability period, network life time, residual energy and throughput. The simulation results show that the proposed protocol named as KBBO has improved the performance of these parameters significantly.*

Streszczenie. *W pracy przedstawiono protokół, w którym stosuje się podejście hybrydowe do grupowania sieci czujników łączących algorytm BBO i K-średnich. Jego wydajność jest porównywana z SEP, IHCR i ERP pod względem okresu stabilności, żywotności sieci, energii resztkowej I przepustowości. Wyniki symulacji pokazują, że prezentowany protokół nazwany KBBO znacznie poprawił wydajność tych parametrów. (**Hybrydowe podejście do efektywnego energetycznie tworzenia klastrów dla heterogenicznej sieci czujników bezprzewodowych z wykorzystaniem optymalizacji opartej na biogeografii i k-średnich**).*

Keywords: clustering, network life time, stability period, optimization
Słowa kluczowe: grupowanie, żywotność sieci, okres stabilności, optymalizacja

Introduction

A WSN consists of a large number of sensor nodes which can sense physical properties of environment such as humidity, temperature, sound, etc. , collect sensed data and transmit it to the base station through a wireless link. Sensor nodes are characterized by limited energy, low processing capability, low communication range and low memory capacity. Sensor nodes have become more smarter and cheaper in recent days due to development in Micro-Electro-Mechanical Systems (MEMS) [1]. Due to availability of smarter and cheaper sensors, WSN has received wide acceptability with potential application in a large domains: environment monitoring, disaster warning system, health system and military application [2]. Unlike a traditional network, the major research challenge in WSN is how to maximize the network lifetime by reducing energy consumption of the sensor nodes. Clustering is a key technique to reduce energy consumption and extend the lifetime of the network. Besides reduction in energy consumption and extending network life time , clustering technique has many other advantages: scalability, latency reduction, collision avoidance, less overhead and load balancing [3] .Selection of cluster heads (CHs) in energy efficient clustering mechanisms depends upon several factors: residual energy of a node, initial energy, average energy of a network and energy consumption of the node [4]. Clustering is also a key technique for implementing energy efficient routing in WSN. LEACH [5], SEP [6], HEED [7] and PEGASIS [8] are some prominent clustering based routing protocols. The probabilistic models are used to select CHs in these protocols which may result in random selection of CHs irrespective of distribution of nodes and the residual energy of the network. In recent days a number of meta-heuristic techniques such as GA [9], PSO [10] and BBO [11] have been applied in design of energy efficient clustering based routing protocols for WSN. IHCR [12] and ERP [13] are two important protocols. Heuristic approach to find a solution to optimal CHs in WSN is not easy, because this is an NP hard problem [14]. A total time complexity to find the optimal solution increases exponentially with the increase in a network size. BBO has been found to be a powerful search and optimization technique because it combines both exploration and exploitation features based on migration [15]. BBO like GA is an evolutionary algorithm applied to solve global optimization problems ,In the proposed work, BBO based clustering technique is used to search for appropriate cluster heads such that the resulting clusters are optimal. To have good quality habitats, K-means is applied to seed the population of BBO. K-means [16-18] is one of the most popular data clustering techniques which partitions the data into K-clusters such that similar data should be in the same cluster. The main contribution of the paper is the following:

- Population initialization through K-means to select good quality features in a habitat,
- Clustering through BBO.

Performance comparison with high performance heuristic and meta-heuristic techniques.

Literature survey

Many heuristics have been proposed for cluster based routing for WSN. LEACH is the most prominent protocol which does load balancing by dynamically rotating CHs among sensor nodes in order to save energy. However the main drawback is that a selection of a CH is done probabilistically. As a result a node with very low energy may be selected as a CH which may not last for a long time. Therefore a large number of protocols have been proposed to improve the performance of LEACH such as SEP, HEED and PEGASIS and its variants [19-21].

Besides heuristics, several meta-heuristic techniques have been used for design of energy efficient clustering and routing. In the area of GA, Martin et, al [12] have used GA for energy efficient cluster based routing to extend the network lifetime and minimize the energy consumption of the network. Its cluster head selection process is based on multi-objective parameters: sum of all distances from sensor nodes to the CH and from the CH to the Base station, sum of distances from sensors node to CH , distances from CH to the sink, standard deviation in cluster distance, transfer energy and number of transmissions respectively. Bara'a et al, [13] have further enhanced the clustering parameters of IHCR by adding three parameters: intra-cluster distance, inter cluster distance and a number of

CHs. Unlike the previous two algorithms where GA has been used for CH selection, in [22] GA has been used for load balancing of CHs. Its fitness function is based on standard deviation of the gateway load to achieve even distribution of the load per cluster. In [23] two algorithms, one for routing and an another for clustering algorithm have been proposed using PSO. The main objectives for routing is to minimize the maximum transmission distance between two nodes in the routing path and maximum hop count of the gateways whereas clustering path of the algorithm achieves load balancing by minimizing distance between sensor nodes and their corresponding gateways. In [24-25] BBO based clustering has been proposed. Metrics like clustering density and clustering head dispersion are used to select cluster heads. Similarly in [26-27], BBO based clustering and multi-hop routing algorithms are proposed with different fitness functions. Multi-objective parameters which include residual energy of a CH, intra-cluster distance and distance between CH and the base station. The formulation of fitness function for routing algorithm is based on residual energy, Euclidean distance and node degree. Unlike the above mentioned protocols where population initialization is random, the proposed protocol follows the deterministic approach to population initialization. The entire population is seeded with K-means algorithm to have good quality genes. BBO will be used further for clustering.

BBO [11,28] is a population based optimization algorithm inspired by immigration and emigration of species(animals, birds, fish) between habitats (islands) in search of good living condition (rainfall, temperature and others). Each candidate solution also called a habitat with a habitat suitability index (HSI). Variables of an individual solution represents features of habitability are called suitability index variables (SIVs). High HSI represents a good solution. The poor solution represents habitat with low HSI. The low HSI receives new features from high HSI. There are three kinds of operators in BBO namely, migration, elitism and mutation. Migration is used for information sharing. Elitism is used to keep best solution for the next generation. The objective of mutation is to increase exploration among the population. One unique feature of BBO is that the original population is not discarded after each generation unlike GA. Rather, it improves its solution in each iteration by changing its features (SIVs). Like PSO, BBO also shares information among solutions but directly. It can directly share its SIVs (features) with other solutions. However PSO does sharing of information among solutions indirectly through a velocity variable [29-30].

Solution

In the current work a new hybrid evolutionary algorithm named as KBBO which combines K-means and BBO is proposed to solve clustering problem in WSN. A combination of deterministic and metaheuristic approaches for population initialization and clustering respectively provide a good BBO based clustering solution. The proposed protocol goes through several rounds. Each round has two phases: (i) setup phase and (ii) steady state phase . In the setup phase ,sensor nodes are partitioned into clusters. The setup phase of the algorithm is same as that of LEACH but based on BBO algorithm. In the second phase, each non-CH node uses its TDMA schedule to transmit its data to the base station through its respective CHs in multi-hop manner [31-33].

a) Habitat representation and initialization

In BBO a habitat represents a set of sensor nodes elected as CHs, normal nodes and dead nodes. Therefore a size of a habitat is a number of sensor nodes in a network. A binary representation is used to distinguish nodes. CH is represented as 1,a non-CH node as 0 and a dead node. The habitat in the population is seeded with K-means algorithm.

In KBBO, the fitness function F is defined as follows:

$$(1)\quad F=\frac{\text{intra-cluster distance}}{\text{inter-cluster distance}}=\frac{\sum_{i}^{CH_k}\sum_{\forall S_j\in CH_i} d\left(S_j,CH_i\right)}{\sum_{i=1}^{CH_k}\sum_{\forall CH_i\neq CH_j} d\left(CH_i,CH_j\right)},$$

where S_j is a sensor node CH_i, CH_j are two different cluster heads and d denotes the Euclidean distance.

In order to achieve optimal clustering in WSN, intra-cluster distance must be minimized whereas inter-cluster distance must be maximized.

b) BBO based clustering algorithm

GA based clustering algorithm in KGA has (i) setup phase which is a clustering formation phase and (ii) steady phase in which intra-cluster and inter-cluster communication takes place. The setup phase of the algorithm is same as that of LEACH but based on BBO algorithm.

c) Simulation results: energy model performance metrics and network model

The following are the assumptions with respect to WSN in the proposed protocol:

- Sensors are randomly deployed throughout the sensing area,
- Sensor nodes are heterogeneous in terms of initial energy,
- Energy consumption of the sensor node during data transmission depends upon the (i) distance between a sender and a receiver and (ii) the size of data,
- A node calculates distance to other nodes and CHs based on Euclidean distance,
- Sensor nodes communicates with a sink node deployed in the middle through their respective CHs,
- Nodes are stationary after deployment,
- Communication links are symmetric.

In this work a simple radio model [5] is used to model the energy dissipation. The free space model (d^2 power loss) is employed in case the distance between the transmitter and receiver is less than a threshold value d_0, otherwise multipath fading channel model is (d^4 power loss) is employed. The energy consumption of l bit message between two nodes at a d distance is given in (2-3) and (4).

$$(2)\quad E_{Tx}\left(l,d\right)=E_{elec}*1+\varepsilon_{fs}*1*d^2, d\le d_0,$$

$$(3)\quad E_{Tx}\left(l,d\right)=E_{elec}*1+\varepsilon_{mp}*1*d^4, d>d_0,$$

$$(4)\quad E_{Rx}\left(l,d\right)=E_{elec}*1.$$

where $E_{Tx}(l, d)$ is l bit transmission energy consumption at a distance of d ,$E_{Rx}(l, d)$ is energy consumption in receiving l bits data. E_{elecis} per bit energy consumption of transmitter and receiver circuits.ε_{fs} and ε_{mp} refer to energy consumption factor of amplification for the free space and multipath fading models respectively.

The performance of KBBO has been compared with SEP, HCR and ERP on five metrics: network life time (number of alive nodes Vs number of rounds), residual energy,stability period and throughput for 10% and 20% advanced nodes. The results have been further validated quantitatively (through tables) The network and energy models are used as defined in LEACH.

- Network Lifetime: It is defined as the number of rounds when all the nodes exhaust their energy and eventually they die.
- Stability Period: It refers to the time interval from the start of the network operation(first round) until the first node dies.
- Residual Energy: It measures the remaining energy of the network every round which is calculated by subtracting

the energy consumed by the nodes from the total energy per round.

- Number of alive nodes per round. It refers to a number of advanced nodes or normal nodes whose residual energy values is greater than zero.
- Throughput: It is a number of packets received at the sink node from the CHs.

d) Simulation Settings

The protocol has been implemented using Intel Core i3 processor with 2GB RAM with 100 sensor nodes which are randomly deployed in the area of 100×100 m^2. The location of the sink is in the middle of the deployment area. The parameters for simulation settings are shown in Table 1.

Table 1. Parameters Setting for Simulation

Description	Parameters	Value
Network Size	M × M	100 x 100 $meter^2$
Location of Sink	BS	(50,50)
Number of Nodes	n	100
Initial Energy of Nodes	E_0	0.5 J
Proportion of advanced nodes	m_0	0.1
Energy factor for advanced nodes	α	1
Energy dissipated per bit	E_{elec}	50 nJ/bit
Transmit amplifier if $d_{BS} \leq d_0$	E_f	10pJ/bit/m^2
Transmit amplifier if $d_{BS} \geq d_0$	E_{mp}	0.0013pJ/bit/m^4
Data aggregation energy by CH	E_{DA}	5 nJ/bit/message
Size of Data Packet		4000 bits

It is observed from table 2 that the stability protocol of KBBO is better than SEP, HCR for 20% advanced nodes except ERP.

Table 2. Stability Period

	SEP	HCR	ERP	KBBO
m=0.1	864	914	1012	888
m=0.2	999	920	1079	1034

Network lifetime can be shown by capturing the number of alive nodes at each round till every node in the network dies. Figure 1 and Figure 2 highlight the network lifetime of each considered algorithm over 10% and 20 % of node heterogeneity respectively. In each of the scenario, KBBO performs better than the rest of the protocols.

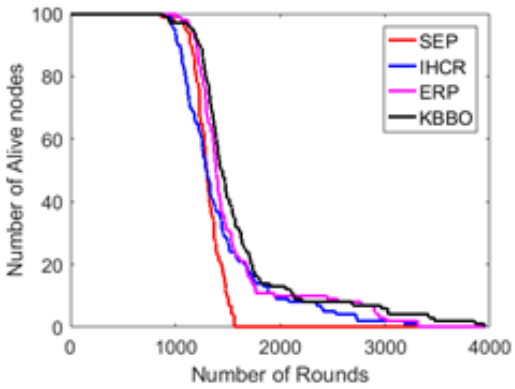

Fig. 1. Number of Alive nodes Vs Number of rounds(10% advanced nodes)

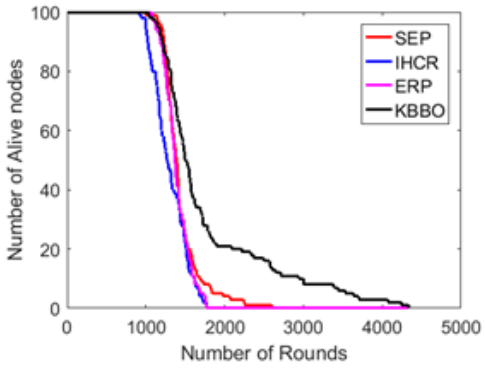

Fig. 2. Number of Alive nodes Vs Number of rounds (20% advanced nodes)

e) Throughput

Figures 3 and 4 demonstrate the number of data packets sent to the base station by CH nodes per round with 10 % and 20 % node heterogeneity respectively. From these figures it can be observed that in KBBO cluster heads sends more packets to the base station compared to the other protocols.

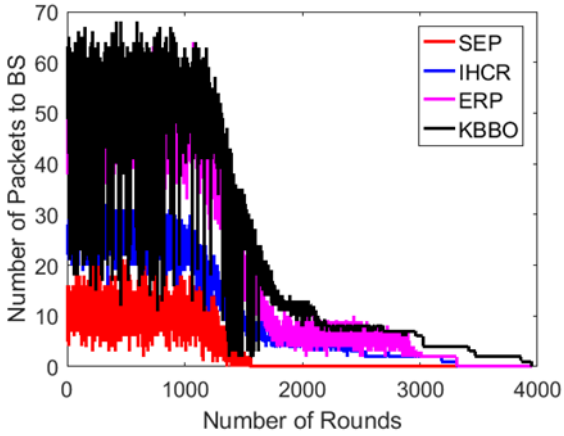

Fig. 3. Throughput (10% advanced nodes)

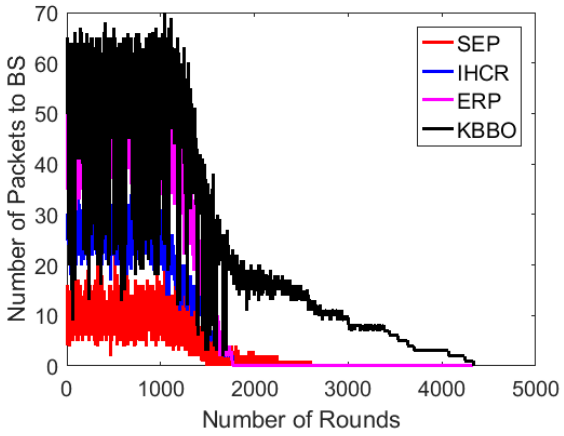

Fig.4. Throughput (20% advanced nodes)

Conclusions

Many clustering based routing techniques have been proposed in the literature in order to save energy and extend the network lifetime. LEACH is the most prominent protocol but it has own limitation. Selection of CHs is done probabilistically and the network lifetime is not maximized significantly. Selection of optimal CHs in WSN is an optimization problem. Meta-heuristic techniques offer effective alternatives to solve optimization problems. In this paper , a new protocol KBBO is proposed based on BBO.

BBO is the latest evolutionary technique to solve optimization problem. The proposed technique is explained through population initialization, setup phase and transmission phase algorithms. The performance of KBBO has been rigorously tested and compared with SEP,IHCR and ERP on several metrics: network life time (number of alive nodes Vs number of rounds), residual energy,stability period and throughput for 10% and 20% advanced nodes. The results have been further validated quantitatively (through tables) . The result shows the superiority of the protocol over several parameters.

***Authors**: Shashi Bhushan, School of Computer and Information Science, Indira Gandhi National Open University, New Delhi, India, e-mail: s.bhushan2k5@gmail.com, Professor Svitlsna G. Antoshchuk, Institute of Computer System, Odesa National Polytechnic University, Shevchenko avenue, 1, Odessa, 65044, Ukraine, e-mail: asgonpu@gmail.com, Docent Pavlo O. Teslenko, Institute of Computer System, Odesa National Polytechnic University, Shevchenko avenue, 1, Odesa, 65044, Ukraine, e-mail:p_a_t@ukr.net; Prof. Waldemar Wójcik, Lublin University of Technology, Institute of Electronics and Information Technology, Nadbystrzycka 38A, 20-618 Lublin, Poland, e-mail: waldemar.wojcik@pollub.pl; M.Sc. Orken Mamyrbaev, Institute Information and Computational Technologies CS MES RK, email: morkenj@mail.ru; M.Sc. Ulzhalgas Zhunissova, Astana Medical University, Kazakhstan, email: ulzhalgas@list.ru.*

REFERENCES

[1] Yick J., Biswanath M., Dipak G., Wireless sensor network survey, *Computer networks* 52 (2008), No. 12, 2292-2330
[2] Akyildiz I.F., et al., A survey on sensor networks, *IEEE Communications magazine* 40 (2002), No. 8, 102-114
[3] Liu X. A survey on clustering routing protocols in wireless sensor networks, *Sensors* 12 (2012), No. 8, 11113-11153
[4] Katiyar V, Narottam C., Surender S., Clustering algorithms for heterogeneous wireless sensor network: Asurvey, *International Journal of Applied Engineering Research* 1 (2010), No. 2, 273
[5] Heinzelman W.B., Chandrakasan A.P., Balakrishnan H., An application-specific protocol architecture for wireless microsensor networks, *IEEE Transactions on wireless communications* 1 (2002), No. 4, 660-70
[6] Georgios S., Matta I., Bestavros A., SEP: A stable election protocol for clustered heterogeneous wireless sensor networks, *Boston University Computer Science Department*, 2004
[7] Ossama Y., Fahmy S., HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks, *IEEE Transactions on mobile computing*, 3 (2004), No. 4, 366-379
[8] Lindsey S., Cauligi S.R., PEGASIS: Power-efficient gathering in sensor information systems, *Aerospace conference proceedings 2002 IEEE,* 3 (2002)
[9] Michalewicz Z. Genetic algorithms + data structures = evolution programs, *Springer*, (2009)
[10] Kennedy J., Particle swarm optimization, *Encyclopedia of machine learning Springer* , (2011), 760-766
[11] Simon D., Biogeography-based optimization*, IEEE transactions on evolutionary computation*, 12 (2008), No. 6, 702-713
[12] Matin A.W., Sajid H., Intelligent hierarchical cluster-based routing, *Life,* 7 (2006), 8
[13] Attea B.A., Khalil E.A., A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks, *Applied Soft Computing* 12 (2012), No. 7,1950-1957
[14] Basagni S. et al., A generalized clustering algorithm for peer-to-peer networks, *Workshop on Algorithmic Aspects of Communication*, (1997)
[15] Ma H. et al., Biogeography – Based Optimization: A 10 year Review, *IEEE transaction on emerging topics in computational intelligence*,9 (2017), No.5
[16] Han J., Kamber M., Data Mining: Concepts and Techniques, *Morgan Kaufman Publishers*, 2 (2006)
[17] Krak Yu.V., Barmak A.V., Baraban E.M., Usage of NURBS-approximation for construction of spatial model of human face, *Journal of Automation and Information Sciences*, 43 (2011), No. 2, 71-81
[18] Kirichenko, M.F., Krak, Yu.V., Polishchuk, A.A. Pseudo inverse and projective matrices in problems of synthesis of functional transformers, *Kibernetika i Sistemnyj Analiz*, 40 (2004), No. 3, 116-129
[19] Krak Yu.V., Dynamics of manipulation robots: Numerical-analytical method of formation and investigation of computational complexity, *Journal of Automation and Information Sciences*, 31 (1999), No. 1-3, 121-128
[20] Wójcik W., Kotyra A., Golec T. et al., Vision based monitoring of coal flames, *Przegląd Elektrotechniczny*, 87 (2008), n.3, 241-243
[21] Vassilenko, Valtchev S., Teixeira J.P., Pavlov S., Energy harvesting: an interesting topic for education programs in engineering specialities, *Internet, Education, Science*, (2016) 149-156
[22] Kuila P., Suneet K.G., Prasanta K.J., A novel evolutionary approach for load balanced clustering problem for wireless sensor networks, *Swarm and Evolutionary Computation,* 12 (2013), 48-56
[23] Kuila P., Prasanta K. J. Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach, *Engineering Applications of Artificial Intelligence,* 33 (2014), 127-140
[24] Al-Maitah M., Timchenko L.I., Kokriatskaia N.I. et al., Parallel-hierarchical network as the model of neurocomputing, *Proceedings of SPIE* 10808, (2018)
[25] Pal R., Himashu Mittal A.P., Mukesh S., BEECP: Biogeography optimization-based energy efficient clustering protocol for HWSNs, *Contemporary Computing (IC3),* (2016)
[26] Lalwani P., Haider B., Chiranjeev K., BERA: a biogeography-based energy saving routing architecture for wireless sensor networks, *Soft Computing*, (2016), 1-17
[27] Vyatkin S.I., Romanyuk S.A., Pavlov S.V., et al., Using lights in a volume-oriented rendering**"**, *Proc. SPIE* 10445, (2017)
[28] Vyatkin S.I., Romanyuk A.N., Gotra Z.Y, et al., Offsetting, relations, and blending with perturbation functions*, Proc. SPIE* 10445, (2017)
[29] Vyatkin, S.I., Romanyuk, A.N., Pavlov, S.V., et al., Fast ray casting of function-based surfaces, *Przegląd Elektrotechniczny*, 93 (2017), No. 5, 83 – 86.
[30] Timchenko L.I., Pavlov S.V., Kokryatskaya N.I., et al., Bio-inspired approach to multistage image processing, *Proc. SPIE* 10445, (2017)
[31] Mosorov V., Panskyi T., Biedron S., Testing for revealing of data structure based on the hybrid approach, Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOŚ, 7 (2017), No. 2, 119-122
[32] Romanyuk O.N., Pavlov S.V., Melnyk O.V., Romanyuk S.O., Smolarz A., et al., Method of anti-aliasing with the use of the new pixel model, *Proc. SPIE* 9816, (2015)
[33] Romanyuk S.O., Pavlov S.V., Melnyk O.V., New method to control color intensity for antialiasing, *Control and Communications (SIBCON)*, (2015)