International Journal of Engineering & Technology, 7 (4) (2018) 5917-5921

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.22671
Research paper

An energy aware data aggregation in wireless sensor network using connected dominant set

K. Santhoshkumar [1] *, P. Suganthi [2]

[1] *Tiruvallur Arts and Science College Kurinjipadi-607302, TamilNadu, India*
[2] *Namakkal Kavignar Ramalingam Government Arts College for Women Namakkal-637001, TamilNadu, India*
**Corresponding author E-mail: santhosh.phd16@gmail.com*

Abstract

The technological progress in various domains has enabled the formulation of Wireless Sensor Networks (WSN). These areas include highly integrated digital circuits, low energy consumption processors, wireless communications among the others. For de-creasing the energy consumed in data communication, some of the superfluous data is eliminated during data aggregation. One of the most famous protocols based on hierarchy and which factors routing in WSNs is Low Energy Adaptive Clustering Hierarchy (LEACH). Among various improvisations suggested for LEACH, the stable election protocol is one of them .For improving the performance of the WSNs, the virtual backbone of the network is formulated by A Connected Dominating Set (CDS) of a graph. This work factors the formulation of clusters based on CDS. This technique has been contrasted with LEACH.

Keywords: *Wireless Sensor Networks (WSN); Low Energy Adaptive Clustering Hierarchy (LEACH); Connected Dominating Set (CDS).*

1. Introduction

The physical environment can be studied and controlled using the sensors. There are several nodes that are contained in WSNs. These sensor nodes are densely disseminated and they have to evaluate various network parameters and send the collated data to the sink using hop-to-hop communication. After sensing this data, the sink sends the same to the users. In several domains like biomedicine and healthcare applications, the WSN application has improved significantly over the years, including defense and habitat monitoring areas [1].
The methodical aggregation of the data sensed from several sensors to be ultimately transmitted to BS for processing is referred to as data gathering [2]. The sensor node is ineffective in sending the data directly to the base station as they are power constrained. The neighboring sensors also generate data. This is not only highly correlated but is also redundant. Additionally, very huge amounts of data are generated in large networks for processing by BS. Thus, there are techniques involved for aggregating the data into information of high quality at the sensors or the intermediate nodes for decreasing the amount of packets that are sent to the BS which can save bandwidth as well as energy. This task can be achieved using data aggregation. The process by which data from various sensors is collated and sent to the base station so that there is an elimination of redundant transmission and provision of only the useful information is known as data aggregation. Data from many sensors is collected and fused at the intermediate nodes. These will send the data further to the sink/base station (BS).
Hierarchical routing is a concept where, a cluster is created and a chosen sensor node within each cluster known as the cluster head (CH) is given extra privileges. This scheme decreases the number of transmissions to the BS. It thus helps in conserving energy as data is effectively combined in every cluster. The Low Energy Adaptive Clustering Hierarchy (LEACH) was the pioneering hierarchical protocol to be incorporated [3]. The basic tenet of LEACH is formation of clusters of sensor nodes on the basis of the robustness of the received signal and usage of the CHs to route the data to the BS. Thus, only CH participates in data transmission rather than all sensor nodes. This helps in conserving energy. Based on this tenet, there are several hierarchical routing protocols that emerged. The objective is to providing a survey of the protocols based on LEACH.
An important factor to be taken into account is the management of power in WSN which is a sought after field of research. This problem focuses on extending the lifespan of WSN and also on improving the reliability of the link .The most pressing challenges of the WSN is extending the network lifetime and management of the network. For decreasing the overheads, controlling the network topology and also extending the network lifetime, a strategy called Connected Dominating Set (CDS) has been found to be very effective. This helps achieve all stated targets [4] as the bandwidth efficacy is improved by the CDS backbone. Turning off some of the superfluous nodes, the CDS backbone ensures that the nodes in the network are connected and removes the useless transmission links. This leads to data being effectively delivered .Some of the nodes in virtual backbone WSNs have been chosen as the backbone node/dominator node in the formulation process of CDS.
A subset of nodes which can perform special tasks and service the non-backbone nodes is referred to as a backbone. The task to be done determines the construction of the backbone. In case of WSNs, the backbone/CDS comprises active sensors with the remaining ones in sleep state. In order for the backbone nodes to interact for performing special tasks, the network backbone needs to be connected. This connected backbone can facilitate effective routing as well as broadcasting. Dominator or backbone nodes refer to the nodes in the CDS. The remaining are referred to as the non-backbone/dominate nodes. The CDS enables ease of routing and rapid adjustment to network topology transformations. In order to decrease the traffic during data communication, a mini-

mum CDS has to be formulated. For decreasing the redundant as well an unnecessary overhead in communication, a lot of attention is being given to topology based on Connected Dominating Set (CDS). The main task of communicating only with the dominators is restricted by using such CDS. A CDS graph associated with the wireless network is normally utilized as the network backbone [6]. A Dominating Set (DS) D of a graph G is a subset of V, and each node in V\D is adjacent to at least one node in D. A CDS C of a graph G is a DS of G, and the sub-graph G[C] induced by C is connected.

For energy aware data aggregation, this clustering approach based on CDS has been suggested, in WSNs. The related literature in the work has been explained in the second section. The details on the used algorithm are given in the third section. The results are discussed in the 4th section and the conclusion of the work is given in the 5th section.

2. Literature survey

The data aggregation schemes in Flat& Hierarchical WSNs has been presented by Sangolgi&Zakir [2] . These schemes are compared based on parameters such as data precision, latency and lifetime of the network .The cluster based network which makes use of the LEACH protocol delivering ten times the data compared to the least energy transmission routing is mainly focused on data aggregation. This in turn results in improving the lifespan of the system and provides a reliable transmission of data with a power consumption improved by a factor of 8 in comparison with direct connection. The challenges and issues in LEACH are dealt with in this work& the proposed protocol LEACH Access Point(LEACH-AP).

A stream aggregation prototype has been suggested by Zechinelli-Martini et al [7]. This makes use of the resources of the sensors including the implementation as well as experimental validation suited to the monitoring of environment. In place of sampling and then instantaneously sending the new samples to the BS, the samples are stored temporarily in its RAM's history. After aggregating, the data is forwarded in any of these scenarios: on receiving a query, when the history is full and when the stored data size is up to the optimal size of the packet. As the sent packet can attain larger sizes for a regular wireless sensor network, the packet error level can be held less than some user selected threshold , by using FEC scheme which is packet level adaptive scheme .Outcomes have demonstrated that this scheme performs better compared to the standard schemes like sample-and-transmit. This can enhance the lifetime of the sensors by at least fifty percent and simultaneously provide better throughput in terms of the rate of loss of network packets. For achieving better outcomes, the data aggregation of the source may be combined along with the other schemes for network aggregation as well.

A recent evaluation of the LEACH-based hierarchical routing protocols has been conducted by Hani&Ijjeh [3]. Particularly, for every protocol, the network lifetime as well as the power consumed has been shown. Additionally, these protocols have been contrasted in terms of CH selection condition, drawbacks, assumptions and benefits that include the improvisations over LEACH.

For wireless sensor networks, an Optimized Region Based Efficient Data(AORED) routing protocol has been suggested by Faheem et al [8]. For formulating a virtual backbone comprising the network nodes that communicate, CDS has been incorporated in AORED. It has been proven by experimental studies that the suggested routing protocol outperforms the popular DEEC and SEP protocols. Associated with AORED are enhanced number of transmissions, enhanced CHs and fewer packets sent to BS.

There is a considerable influence of the CDS of a graph on the effective designing of the routing protocols in the WSNs. All the nodes that have been analyzed in the extensive study of this problem have similar ranges for transmission and also taken into account is how big the dominating sets are. By means of constructing the Steiner tree, the modelling of a network as a unidirected graph is done in Zhang et al [9]. This has led to the CDs that are smaller in size. There is also constancy in the ratio of performance of these algorithms. It has also shown an association in the size between the maximal autonomous sets and a CDS; the limitation on the highest number of autonomous nodes in the neighborhood has also been shown. In order to verify the suggested schemes, the theoretical analysis as well as outcomes of simulation has also been given.

The nodes in WSN have limited energy. They are deployed arbitrarily. Due to the absence of a physical infrastructure, a virtual backbone is employed so that the communication is efficacious. A Connected Dominating Set (CDS) is constructed and a virtual backbone is formulated. In the recent past, topology control is the chief research issue in the efficient and the rapid formulation of the CDS. The benefits and the drawbacks of the CDS have been presented by Mohanty & Mandal [10]. They have also presented an in-depth survey of the CDS construction algorithms. The chapter has been concluded with some of the existing issues as well as proposal of some of the interesting ones. For formulating a virtual backbone in WSNs, the CDS is a popularly adopted approach. This can lead to the forwarding of traffic by the virtual backbone. The other nodes can switch off radio and thus the power can be conserved. Additionally, the smaller the CDS, the fewer will be interference associated. However, this happens to be a NP-hard problem. Thus, the focus of most researchers would be how the approximate algorithms are derived. A new algorithm is suggested by Zhang et al [11]. This is based on the induced tree of the crossed cube (ITCC). This strives to obtain a Maximal Independent Set (MIS), which depends on construction of an induced tree of the crossed cube network, and then to connect the MIS nodes to form a CDS. Based on a novel parameter, the priority of an autonomous tree has been found. This is the node's degree in a graph's square. There has been a presented proof that a CDS is generated using the ITCC and a lower approximation ratio. It has also been proven further that a Fibonacci sequence constitutes the cardinality of the induced tree. This establishes a threshold for the amount of the dominating sets. It has been shown via the simulations that the smallest size of the CDS is provided by the algorithm in comparison with the others. For expanding the lifespan of the WSN, a distributed heuristic was suggested by Ramalakshmi& Radhakrishnan [12]. This presented a stable MPR based CDS which was also efficient in terms of power. It took into account the energy as well as the velocity of the nodes .It has been shown by means of simulation outcomes that not only does the proposed scheme conserve power but it also enhances the network lifetime by about 25% compared to the others.

3. Methodology

This section details on Low Energy Adaptive Clustering Hierarchy (LEACH) and CDS.

3.1. Low energy adaptive clustering hierarchy (LEACH)

A popular clustering heuristic in WSN is the LEACH. It has the ability to extend the network lifespan. It is adaptive, auto-structured clustering algorithm. Iterations are used. The assumptions in LEACH are as follows: The distance between sensor nodes and base station is huge. Similarity exists among all sensor nodes. The energy source is constrained. The environment can be sensed at a constant rate. The sensors not only communicate amongst each other but also with the BS. The tenet of LEACH is organizing the nodes into clusters so that the power is disseminated amongst the nodes. Every node has a special node called Cluster Head (CH) chosen through election.

Each round in LEACH comprises of two stages like setup phase and steady phase. Formation of clusters is in the former stage and the transfer of data is in the latter. Initially in the set up phase, an arbitrary number between 0 and 1 is picked up by every node. Then, the threshold formula is computed. If the chosen random number is lower than this value, the chosen node is the cluster head. Threshold formula is given as below:

$$T(n) = \begin{cases} \dfrac{P}{1 - P*(r \bmod \frac{1}{p})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

r- current iteration.
p- required percentage of CHs.
n- Set of nodes which are not CHs in the last G iterations that can become CHs.
The heuristic states that the nodes which become CHs have a five percent likelihood under usual scenario. A node which is cluster head in round 0 cannot again become a CH in the subsequent iteration. After 1/p-1 rounds, the threshold value will beT(n)=1. Then, the nodes can again become cluster heads. Every CH after being selected will send a promotional information to the remaining nodes using CSMA MAC protocol. This is followed by every node choosing a CH which depends on the promotion's Received Signal Strength Indication (RSSI). All the cluster heads must turn on the receivers at that point. Then, as per how many nodes are comprised in a cluster, every CH formulates a TDMA schedule upon the cluster formation .Every node broadcasts its data to the CH during the assigned transmission time in TDMA.
Every sensor node can sense as well as send the data to its CH based on the schedule of the TDMA, during the steady state phase. The data received by the CHs is aggregated before it is sent to the BS. Post sometime that is decided beforehand, the network commences iteration by returning to the set up phase and the steady state phase. A simple radio network delineating the dissemination of energy by means of electronic devices, transmitter, power amplifier and receiver is assumed by the LEACH protocol.
There are many drawbacks of LEACH [11]. The data is aggregated at the CHs. This decreases the traffic across the network. The routing from the nodes to CH is single hop. This leads to energy conservation. Also the WSN lifetime is enhanced. Despite all these benefits, drawbacks also exist. It is assumed in the LEACH that the initial energy of all nodes is the same. However, practically, that is not the case. It is also assumed that the nodes are static. Hence, the same cannot be applied in case of mobile nodes. Additionally, issues arise due to the failure of CHs and also presence of several Base stations. Dynamic clustering presents additional overhead. This lead to the emergence of protocols for improving LEACH.

3.2. Connected dominating set (CDS)

Dominating Set, Given a graph G = (V,E), a Dominating Set (DS) of G is a subset C ⊂ V such that each node either belongs to C or is adjacent to at least one node in C. This means that a Dominating Set of a graph G = (V,E) is a set of nodes V′ such that ∀(v,w) ∈ E, v ∈V′ or w ∈ V′. A set of vertices covering all the edges is given by vertex cover. However, a vertices set that covers all vertices is referred to as dominating set (c). Only if there is a route between any 2 nodes in the set comprising only these nodes, then a Dominant Set is said to be connected and called as Connected Dominating Set (CDS) . A Connected Dominating Set of G = (V,E) is a Dominating Set of G such that the subgraph of G induced by the nodes in this set is connected. The nodes in a CDS are called the dominators. Apart from dominators, the other nodes are referred to as dominates. The number of dominators and the CDS' sizes are equal. A dominator dominates every dominate. In the CDS C, the nodes in C can communicate with any other node in the same set without using nodes in V – C. A dominating set is a maximal independent set. Nodes that are connected in MIS forms a CDS. Yet, every dominating set is not an MIS as nodes in the dominating set may be beside one another.
CDSs have been used as the basic structure for the MAC, multicast/broadcast, location-based routing, energy conservation, resource discovery protocol and so on. Some of the advantages of CDS based routing protocols are:

- Used in routing, sending data and avoiding collisions.
- Decreased searching of routes and the process of routing to the subnetwork induced from the CDS. Routing information can be maintained only by dominators.
- CDS improvises the efficacy of the multicast routing [14].
- The message overheads that are concerned with the routing updates are decreased. An underlying architecture is formed by CDS. This has been used by protocols like media access coordination, unicast, multicast/broadcast, and location-based routing, energy conservation, and topology control.
- CDS can be used as forwarding nodes for decreasing the energy consumed in WSNs.
- Serving as database servers, the virtual backbone in the dominating set can send the link quality information for the selection of route in multimedia traffic.

4. Results and discussion

Table 1 shows the simulation parameters. Table 2 to 6 shows the Number of clusters formed, average end to end delay (sec), average packet loss rate, lifetime computation and remaining energy computation for LEACH and CDS-WSN respectively.

Table 1: Simulation Parameters

Transmission range of node	100m
Number of nodes	100 to 600
size of network	2000 * 2000 m
Location of base station	centre of network
Data aggregation energy cost	50 pj/bit j

Table 2: Number of Clusters Formed for CDS-WSN

Number of nodes	LEACH	CDS - WSN
100	12	13
200	18	20
300	30	31
400	34	34
500	36	36
600	40	40

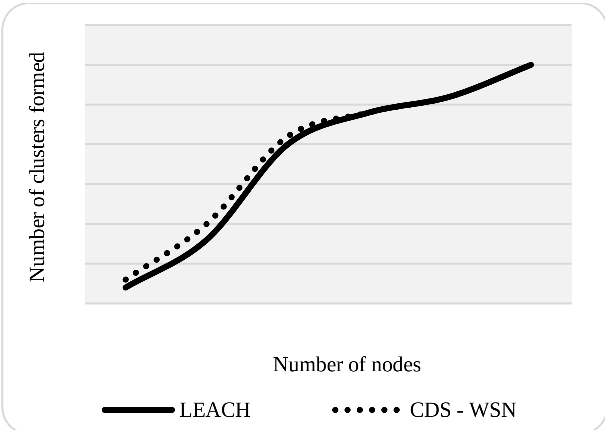

Fig. 1: Number of Clusters Formed for CDS-WSN.

From table 2 and figure 1 it is observed that the number of clusters formed for CDS-WSN performs better by 8%, by 10.53%, by 3.28%, no change, no change and no change than LEACH at number of nodes 100, 200, 300, 400, 500 and 600 respectively.

Table 3: Average End to End Delay (sec) for CDS-WSN

Number of nodes	LEACH	CDS - WSN
100	0.001624	0.001616
200	0.001716	0.001989
300	0.01693	0.018007
400	0.026305	0.022011
500	0.054397	0.052581
600	0.061074	0.056286

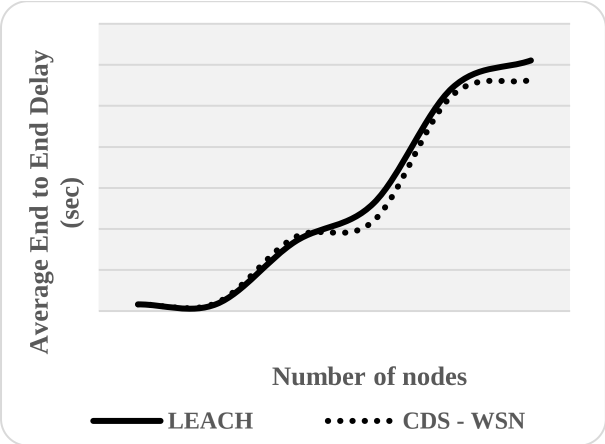

Fig. 2: Average End to End Delay (sec)for CDS-WSN.

From table 3 and figure 2 it is observed that the average end to end delay (sec) for CDS-WSN performs better by lowering delay by 0.49%, by 14.74%, by 6.17%, by 17.78%, by 3.4% and by 8.16% than LEACH at number of nodes 100, 200, 300, 400, 500 and 600 respectively.

Table 4: Average Packet Loss Rate (%) for CDS-WSN

Number of nodes	LEACH	CDS - WSN
100	10.09536	9.150242
200	15.542937	13.765609
300	16.066463	15.443441
400	22.324248	20.771756
500	30.138109	28.094322
600	41.979386	39.944671

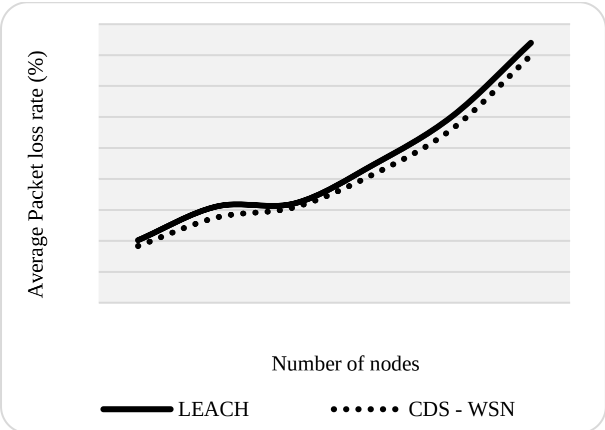

Fig. 3: Average Packet Loss Rate (%) for CDS-WSN.

From table 4 and figure 3 it is observed that the average packet loss rate (%) for CDS-WSN performs better by lowering packet loss rate by 9.82%, by 12.13%, by 3.95%, by 7.2%, by 7.02% and by 4.97% than LEACH at number of nodes 100, 200, 300, 400, 500 and 600 respectively.

Table 5: Lifetime Computation for CDS-WSN

Number of rounds	Percentage of nodes alive_LEACH	Percentage of nodes alive_ CDS - WSN
0	100	100
100	100	100
200	89	91
300	68	84
400	73	76
500	22	54
600	3	18
700	0	0

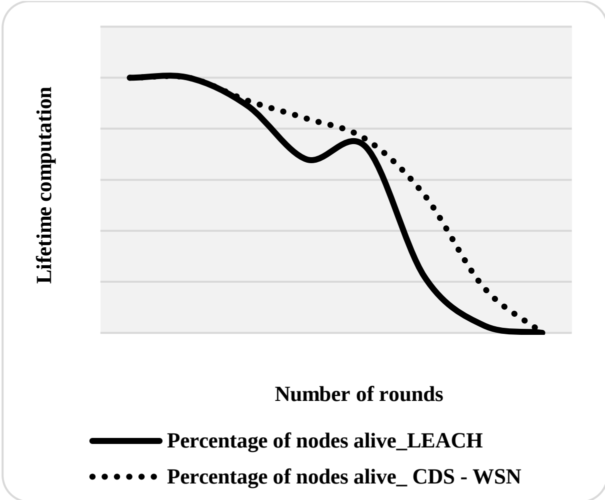

Fig. 4: Lifetime Computation for CDS-WSN.

From table 5 and figure 4 it is observed that the lifetime computation for CDS-WSN performs better by lowering delay by 2.22%, by 21.05%, by 4.03%, by 84.21%, and by 142.86% than LEACH at number of rounds 200, 300, 400, 500 and 600 respectively.

Table 6: Remaining Energy Computation for CDS-WSN

Number of rounds	Average remaining energy of nodes(J)_LEACH	Average remaining energy of nodes(J)_CDS - WSN
0	0.5	0.5
100	0.43	0.46
200	0.23	0.34
300	0.18	0.29
400	0.19	0.27
500	0.11	0.18
600	0	0.1
700	0	0

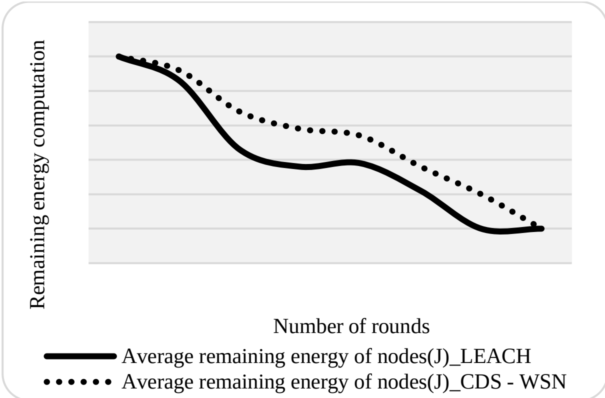

Fig. 5: Remaining Energy Computation for CDS-WSN.

From table 6 and figure 5 it is observed that the Remaining energy computation for CDS-WSN performs better by lowering delay by %, by %, by %, by %, by % and by % than LEACH at number of rounds 200, 300, 400, 500 and 600 respectively.

5. Conclusion

There are several issues of WSNs in terms of restricted battery power, less memory and low speed of processing. In all applications, it is necessary to design an energy effective data processing as well as communication scheme that can assure precision of the data sent. One of the best schemes for decreasing the WSN energy consumption is aggregating data. The routing overheads can be decreased using a Connected Dominating Set (CDS) that serves as a virtual backbone for WSNs. Results show that the number of clusters formed for CDS-WSN performs better by 8%, by 10.53%, by 3.28%, no change, no change and no change than LEACH at number of nodes 100, 200, 300, 400, 500 and 600 respectively.

References

[1] Tang, Q., Luo, Y. S., Xie, M. Z., & Li, P. (2016). Connected dominating set construction algorithm for wireless networks based on connected subset. *J. Commun*, *11*(1), 50-57. https://doi.org/10.12720/jcm.11.1.50-57.

[2] Sangolgi, N. B., &Zakir, S. K. A. (2013). Energy aware data aggregation technique in WSN. *International Journal of Scientific and Research Publications*, *376*.

[3] Hani, R. M. B., &Ijjeh, A. A. (2013). A survey on LEACH-based energy aware protocols for wireless sensor networks. *Journal of Communications*, *8*(3), 192-206. https://doi.org/10.12720/jcm.8.3.192-206.

[4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," Computer Networks, vol. 38, no. 4, pp. 393–422, March 2002. https://doi.org/10.1016/S1389-1286(01)00302-4.

[5] M. Rai, Sh. Verma, and Sh. Tapaswi, "A Power Aware Minimum Connected Dominating Set for Wireless Sensor Networks," Journal of networks, Vol. 4, no. 6, August 2009. https://doi.org/10.4304/jnw.4.6.511-519.

[6] Yu, J., Wang, N., Wang, G., & Yu, D. (2013). Connected dominating sets in wireless ad hoc and sensor networks–A comprehensive survey. *Computer Communications*, *36*(2), 121-134. https://doi.org/10.1016/j.comcom.2012.10.005.

[7] Zechinelli-Martini, J. L., Bucciol, P., & Vargas-Solar, G. (2011, February). Energy aware data aggregation in wireless sensor networks. In *Wireless Communication, Vehicular Technology, Information Theory and Aerospace &Electronics Systems Technology (Wireless VITAE), 2011 2nd International Conference on* (pp. 1-5). IEEE. https://doi.org/10.1109/WIRELESSVITAE.2011.5940864.

[8] Faheem, H., Ilyas, N., ulMuneer, S., &Tanvir, S. (2016). Connected Dominating Set based Optimized Routing Protocol for Wireless Sensor Networks. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, *7*(11), 322-331. https://doi.org/10.14569/IJACSA.2016.071142.

[9] Zhang, L., Li, C., Mao, Y., Xu, L., & Wang, H. (2009, September). Connected dominating sets in wireless sensor networks. In *Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on* (pp. 1-4). IEEE. https://doi.org/10.1109/WICOM.2009.5301402.

[10] Mohanty, J. P., & Mandal, C. (2017). Connected Dominating Set in Wireless Sensor Network. In *Handbook of Research on Advanced Wireless Sensor Network Applications, Protocols, and Architectures* (pp. 62-85). IGI Global. https://doi.org/10.4018/978-1-5225-0486-3.ch003.

[11] Zhang, J., Zhou, S. M., Xu, L., Wu, W., & Ye, X. (2015). An efficient connected dominating set algorithm in WSNs based on the induced tree of the crossed cube. *International Journal of Applied Mathematics and Computer Science*, *25*(2), 295-309. https://doi.org/10.1515/amcs-2015-0023.

[12] Ramalakshmi, R., & Radhakrishnan, S. (2012, January). Energy efficient stable connected dominating set construction in mobile ad hoc networks. In *International Conference on Computer Science and Information Technology* (pp. 63-72). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-27299-8_7.

[13] B. A. Sabarish, M. S. M. Guru, M. A. Dhivya, K. S. Naveen, and S. Vaishnavi, "A survey on clustering protocols in wireless sensor networks," International Journal of Advances in Computing and Information technology, vol. 1, no. 2, 2012.

[14] Yu, J., Wang, N., Wang, G., & Yu, D. (2013). Connected dominating sets in wireless ad hoc and sensor networks – A comprehensive survey. Computer Communications, 36(2), 121-134. https://doi.org/10.1016/j.comcom.2012.10.005.