Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 57 (2015) 1026 – 1035

Procedia Computer Science

3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

Analyzing the Performance of Centralized Clustering Techniques for Realistic Wireless Sensor Network Topologies

Gaurang Raval*, Madhuri Bhavsar, Nitin Patel

Institute of Technology, Nirma University, SG Highway, Ahmedabad 382481, India

Abstract

Clustering techniques in wireless sensor networks enables energy efficient coordination among the densely deployed nodes for data delivery till the base station. Many clustering protocols have been suggested in the recent past. The topology of the nodes, mostly seen in the literature, is of random type. This paper analyzes the performance aspects of various centralized clustering techniques for wireless sensor networks. LEACH-Centralized, KMeans-CP, FCM-CP and HSA-CP protocols have been compared with respect to clustering and data delivery process for various realistic topologies. The simulations were performed for these protocols and performance of the protocols has been critically analyzed. HSA-CP clustering method performs better compared to other techniques for almost each topology examined in the paper.

Peer-review under responsibility of organizing committee of the 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

Keywords: Harmony Search; FCM; Clustering; Wireless Sensor Network;

1. Introduction

Wireless sensor networks are very useful in emergency operations such as search/rescue/alert during natural disasters like earth quake, forest fire, flood etc. For example, sensors are deployed over a wild fire in a forest, from an air-plane. They collectively produce a temperature map of the area or determine the perimeter of areas with high temperature that can be accessed from the outside. The major factors that favor WSNs for such tasks are self configuration of the system with minimal overhead, independent of fixed or centralized infrastructure, the nature of the terrain of such applications, the freedom and flexibility of mobility, and the unavailability of conventional communication infrastructure. Wireless sensor nodes are often deployed for monitoring of vehicles, animals, machines, medical purposes, environment studies, structural health etc.

* Corresponding author. Tel.: +91-79-30642527 ; fax: +0-000-000-0000.
E-mail address: gaurang.raval@nirmauni.ac.in

1877-0509
Peer-review under responsibility of organizing committee of the 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)
doi:10.1016/j.procs.2015.07.516

An important constraint, to be taken into consideration, in wireless sensor networks is that the nodes usually run on batteries. This means that energy expenditure must be carefully controlled or else the node life and consequently network life would be quite short. Most of the energy expenditure comes from the transmission and receiving of data packets. Energy can be conserved if data transmission and receiving is made more efficient, or data transfers are reduced. For effective data transfer from the field to the base station, many routing protocols have been suggested. The responsibilities of a routing protocol include exchanging the route information, finding a feasible path to a destination based on criteria such as hop length, minimum power required, lifetime of wireless link, gathering information about power and utilizing minimum bandwidth. Mohammad et al.[1] have described many challenges which affect the routing protocol design. Major categories of routing mechanism are data centric, hierarchical and location based, as mentioned in the papers[2 3]. In flat or data centric protocols, all nodes are assigned equal roles. In hierarchical, however, nodes will play different roles in the network; while in location-based routing, sensor nodes are addressed by means of their locations. Always there are design trade-offs between energy and communication overhead savings in every routing paradigm.

One of the key techniques to control the energy drain from the network and effectively manage the coordination among the sensor nodes is clustering. D J Dechene et al.[4] classify them into heuristic, weighted Schemes, hierarchical schemes and grid based approaches. They analyze performance with respect to energy and QoS of these protocols. As sensor network deployments for critical applications are generally large scale in nature, specific routing techniques are required to handle such large networks. Hierarchical routing technique helps network achieve scalable and efficient operating conditions so that it can last longer. Nodes having higher energy than the average network energy may be used to aggregate data and send it to BS, while low energy nodes can sense the environment in the vicinity and report information to the nodes which are responsible to collect data. Hierarchical routing protocols acting as defacto standard are LEACH[5] and LEACH-C[6]. Several other protocols also exist, like PEGASIS[7], TEEN[8], APTEEN[9] and HEED[10], which extends LEACH protocol. These protocols consider network lifetime parameter for improvement by enhancing the data transmission efficiency, but the cluster structures are not efficiently created. Few other clustering protocols also exist like K-Means[11 12] based clustering, FCM[15 16 17] based clustering, HSA[20 21 22] based clustering.

Most popular routing protocol LEACH uses cluster based routing to control the energy consumption. LEACH employs distributed cluster formation, whereas LEACHC is a centralized version of LEACH where base station assists for clustering process. Most of the work seen in literature generally applies random topology for analyzing the performance of the protocols. In this paper, several realistic topologies have been tested against centralized routing protocols. The performance of various centralized clustering protocols have been analyzed in terms of clustering effect and data reported at the base station. In section 2, various clustering techniques are discussed in detail. In section 3 energy model has been described. Section 4 presents the realistic deployment strategies, as suggested by Tiago et al[23]. Section 4 presents simulation results and analysis of these methods. Section 5 concludes the paper.

2. Related Work

LEACH[5] is popular protocol among researchers, used for hierarchical routing. LEACH attempts to minimize energy dissipation. It is based on a simple clustering mechanism by which energy can be conserved since cluster heads are selected for data transmission instead of other nodes. LEACH reduces the energy usage by applying simple clustering scheme. It employs cluster-heads for data transmission to the BS station. LEACH uses two phases during the algorithm progress. In setup-phase, cluster-heads are created in distributed way with probability. Later on clusterheads attract the members to form clusters. Then after steady state-phase enables data collection from member nodes to cluster-heads and then to base station. Tyagi et al.[13] present detailed review of LEACH and its variants. LEACH Centralized (LEACH-C)[6], has similar process like LEACH. In set-up phase of LEACH-C, all nodes send their energy and location to base station. After finding the average network energy, BS creates cluster-heads (CH) using the simulated annealing algorithm[14]. The candidate

members are only who are having energy above the average network energy. BS announces the cluster information to members and then steady state phase starts.
Sasikumar et al.[12] have proposed K-Means based clustering. K-Means Algorithm generates randomly 'k' centroids and finds the Euclidian distance of every node with each centroid. It finds the centroid locations in the cluster iteratively, until no change is found in the position of the centroid location. Tan et al.[11] have proposed BPK-Means clustering algorithm. They create balanced clusters by fixing number of members to be strictly same in every cluster. FCMCP[17] applies Fuzzy C-Means clustering algorithm for centralized clustering. FCMCP assigns degree of membership to the members in an overlapping fashion. Initially, the degree of membership is assigned randomly. As the algorithm progresses, clusters are formed iteratively. The algorithm converges with termination conditions, least overlapping membership of members with neighboring clusters and the fuzziness coefficient of the membership. Once the clusters are formed, member node near the centroid is declared cluster-head depending on members' available energy.

Geem et al.[19] proposed Harmony Search Algorithm (HSA). Hoang et al. [20] (HSA-CP) applied HSA for clustering process in WSN. HSA-CP creates clusters by minimizing intra cluster distance and residual average energy, the same parameters used by LEACHC, KMeans and FCM-CP. They define Harmony Memory as matrix HM (eq. 1) having randomly selected member elements in every row. Each vector in the matrix HM candidate cluster-head set. Vectors in HM may be declared as cluster-head set at last or some mutated vector becomes cluster-head set. To create final cluster-head set, the harmony memory is manipulated iteratively. Two parameters control the progress of HSA, Harmony Memory Consideration Rate(HMCR) and Pitch Adjustment Rate (PAR).

$$HM = \begin{vmatrix} V_1^1 & V_1^2 & \cdots & V_1^k \\ V_2^1 & V_2^2 & \cdots & V_2^k \\ \cdots & \cdots & \cdots & \cdots \\ V_{HMS}^1 & V_{HMS}^2 & \cdots & V_{HMS}^k \end{vmatrix} \begin{vmatrix} F^1 \\ F^2 \\ \cdots \\ F^{HMS} \end{vmatrix} \quad (1)$$

HM is generated randomly first, then a harmony memory is randomly picked from HM. This harmony is improvised using algorithm 1. If the new harmony is from HM itself, then it is pitch adjusted with PAR. Member elements in the new harmony may be replaced with neighboring members in the in the Harmony Memory with higher residual energy. Otherwise, the new harmony will be left unaltered with the probability 1-PAR. Preferred HMCR is 0.9 and PAR is 0.8 as seen in literature. U(0; 1) is random uniform function used for probabilistic number generations.

2.1. Improvise New Harmony

Algorithm 1 **The procedure of improvising a new harmony**

for j = 1 to k **do**
  //Harmony Memory Consideration
  if U(0,1) ≤HMCR **then**
    Choose a harmony Vi from HM randomly, i ∈ [1..HMS]
    $V'_j = V^j_i$
    //Pitch Ajustment
    if $U(0, 1) \leq PAR$ **then**
      V'_j = Neighbouring element of V'_j in HM
    end if
  else
    V'_j = Pick random element from pool of all nodes //random selection
  end if
end for

The protocols discussed in previous section have been simulated or tested with random uniform topologies as per the literature review. The sensor nodes deployment can be optimized only when there are favourable conditions like a plain surface, open grounds or inside buildings. In these circumstances nodes may be placed in such a way that the number of required nodes remain under control. But when the area to be monitored is unknown like mountains or forest area or a disaster situation, simple topologies may not work. The area to be monitored may be completely devastated. Deployments in this situations should be fast and unorganized by nature, different kind of deploying should be adopted here. Tiago et al. 23 have proposed a tool (GenSen) for generating realistic wireless sensor network topologies. Several topologies have been created for 50 sensor nodes to 200 sensor nodes, with BS location varied from center to 50,175 in the field. The types of deployment

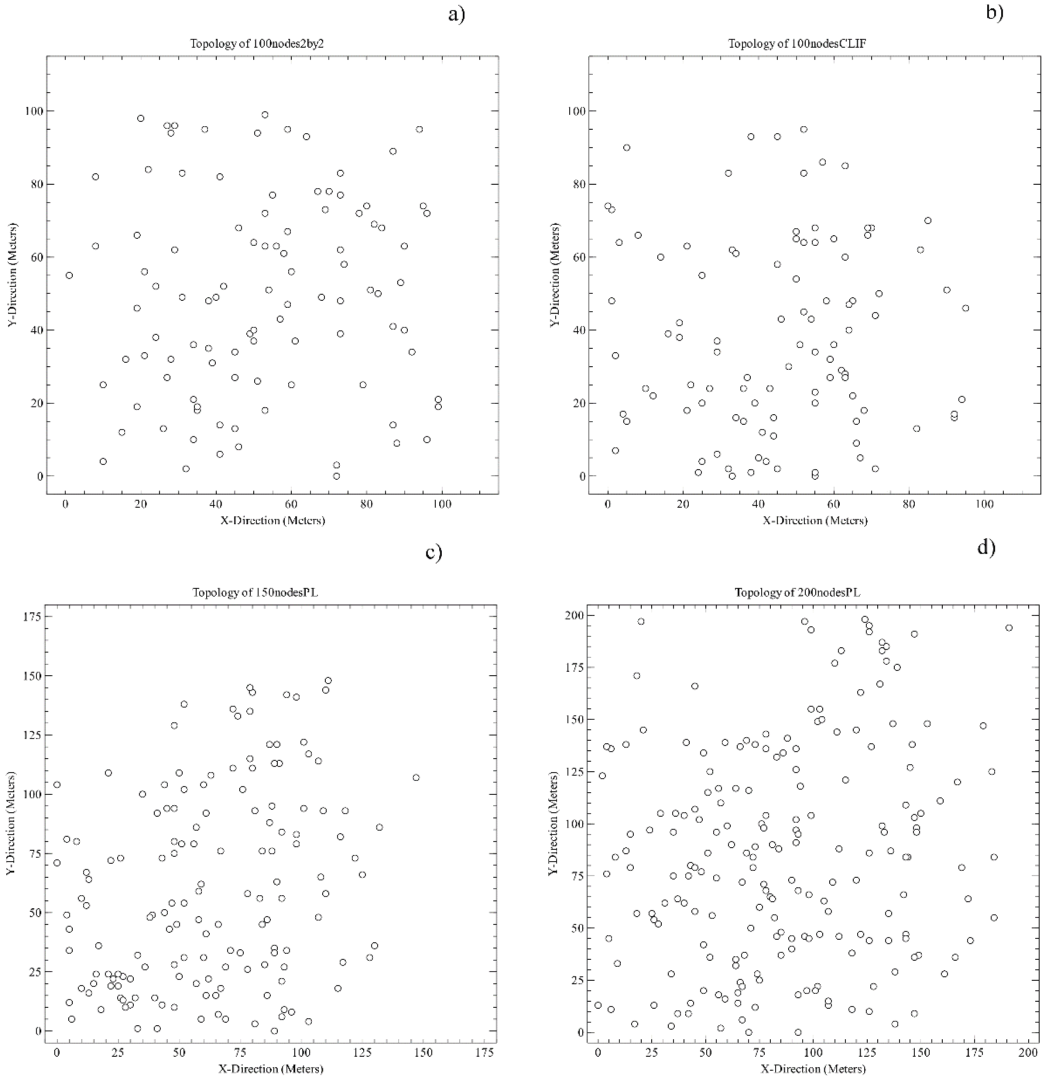

Fig. 1. Topology Plot a) 100 Nodes 2 by 2; b) 100 Nodes Cliff ; c) 150 Nodes Propellant d) 200 Nodes Propellant .

are as follows:

- One-by-One: Nodes are dropped one by one separately. The deployment is done with the estimation between the nodes being thrown individually. Node position is actually not used during this type of deployment.
- Two-by-two: This strategy differs from the previous one only in number. Here, two nodes are dropped together in the field.
- Three-by-three: This method is similar to one by one and two by two, but the number of nodes are three here..
- Cliff: In this method, nodes are dropped from higher altitude. The area is far below from the point from where the nodes are to be dropped. It may be compared with nodes dropped from a flying object high in the air.
- Propellant: Sensors are dropped in the area to be monitored with the help of a propellant. All nodes are dropped at the same time. The propellant drops the sensors nodes in the center of the area to be monitored.

Figure 1 shows various realistic topologies for variety of deployments like 100, 150 and 200 nodes for 2 by 2, Cliff and Propellant topologies. It may be observed that part b, c and d in figure 1 show the concentrated deployment in particular region of the deployment field, with the Cliff and Propellant topologies.

Table 1. Radio Parameters

Operation	Symbol	Energy dissipated
Energy consumed in electronics circuit for transmitting and receiving	E_{elec}	50nJ/bit
Energy consumed by amplifier to transmit at shorter distance i.e. if $d_{toBS} < d0$	ϵ_{fs}	10*pJ*/*bit/m*^{2}
Energy consumed by Amplifier to transmit at longer distance i.e. if $d_{toBS} \geq d0$	ϵ_{mp}	0.0013*pJ/bit/m*^{4}
Energy consumed during data aggregation	E_{DA}	5nJ/bit

Table 2. Simulation Parameters

Network Parameter	Value
Node Distribution	(0,0) to (200,200)
BS location	Center and (50,175)
No. of nodes	50,100,150,200
Initial Node Energy	2J
Simulation Time	3600s
Desired no, of cluster-heads	5%
Bandwidth of the channel	1 Mbps
Packet header size	25 Bytes
Message Size	500 Bytes

3. Energy Model

The radio energy model describes the radio characteristics, including energy dissipation in the transmit and receive modes. Transmitter dissipates energy to run the radio electronics and power amplifier. Receiver dissipates energy to run the radio electronics. The energy model mostly seen in literature, is given by eq. 2. Energy required to transmit L bits at a distance d is,

$$E_{TX}(L,d) = \begin{cases} L\left(E_{elec} + \epsilon_{fs}\, d^2\right) & d < d_0 \\ L(E_{elec} + \epsilon_{mp}\, d^4) & d \geq d_0 \end{cases} \quad (2)$$

Here d_0 is the deciding factor whether to use free space propagation model or multi-path radio propagation model. ϵ_{fs} and ϵ_{mp} are the amplification components depending on the propagation model in use. Energy required to receive L bits is,

$$E_{RX}(L) = LE_{elec} \quad (3)$$

The radio parameters used during simulation are mentioned in table 1.

4. Simulation Results and Analysis

The protocols discussed in section 2, LEACHC, KMeans-CP, FCM-CP, HSA-CP were implemented in ns-2 with TCL and C++. The network simulation parameters are shown in table 2. The simulations were carried out for 50 nodes to 200 nodes, with different BS locations. The protocols were simulated against the realistic topologies discussed in section 3. Figure 2 shows the snapshot of cluster formation process. Each protocol elects five percentage nodes of the alive nodes in the network. It demonstrates the clustering effect captured at the same round for all four protocols for 200 nodes with Propellant topology. It may be seen that LEACHC assign very few members to some clusters, whereas KMeansCP elects cluster-heads concentrated in a particular region in the area to be monitored. FCMCP and HSACP creates better clusters in comparison to LEACHC and KMeans, but few cluster-heads are quite close to each other, so there is very less cluster separation.

The data reporting has been analyzed and shown in tables 3 to 10. The tables present various statistical parameters regarding data reported from individual nodes for all four protocols. The term mean(μ) represents average data reported from every node. StdDev (σ) measures the variation from the mean data reported from the network. The parameter Cv(σ/μ), co-efficient of variation, relates the two terms mean and StdDev. The parameter Max-Min, shows the difference between maximum data reported from a node and minimum data reported from a node for the given network deployment during the simulation.

It may be observed in each table that the data reported is comparatively less when BS location is 50,175. For the deployments with less number of nodes, HSACP performs better compared to other protocols. The standard deviation is reasonably less in each such case, as shown in tables 3,4,5 and 6. However, when the network is dense with large number of nodes, FCMCP and HSACP performs equally good in terms of data delivery as shown in tables 7 and 8. But for the same 150 number of nodes when the topology is different, HSACP shows less variation in average data reported from every node from the network. Evidence for the same is available in tables 9 and 10. Standard deviation and co-efficient of variation, both are less for HSACP compared to all other protocols.

Table 3.BS Location Center – 50 nodes 3 by 3

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	1206.78	214.98	17.81	849
KmeansCp	1158.94	239.15	20.64	829
FCMCP	1146.64	285.16	24.87	863
HSACP	1213.44	202.27	16.67	729

Table 4.BS Location (50,175) – 50 nodes 3 by 3

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	1056.42	163.30	15.46	678
KmeansCp	1054.32	173.36	16.44	702
FCMCP	1041.84	208.87	20.05	703
HSACP	1061.08	129.04	12.16	545

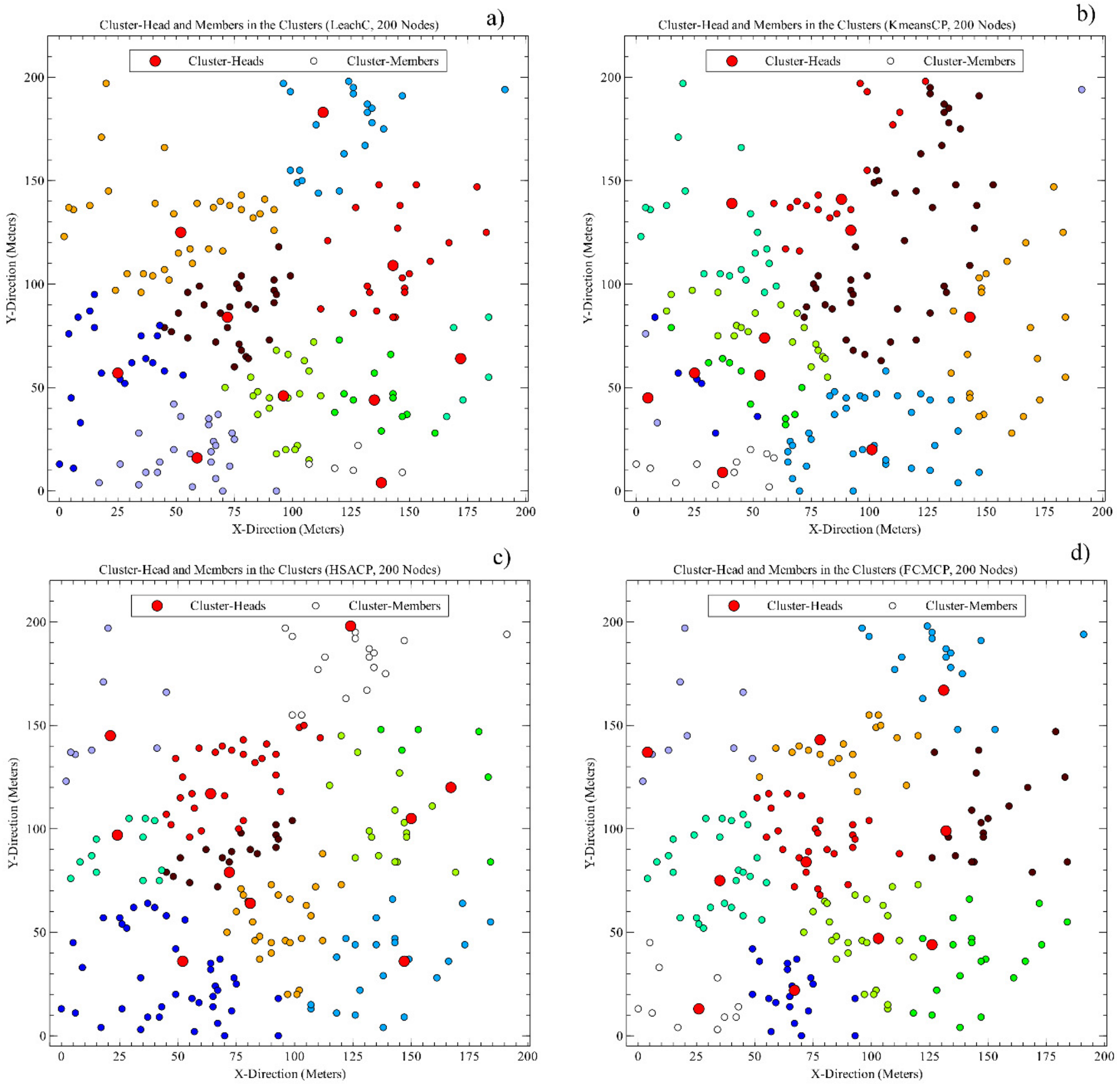

Fig. 2. Clustering for Propellant Topology with 200 nodes a) LEACHC; b) KMeansCP ; c) HSACP; d) FCMCP

Table 5.BS Location Center – 50 nodes Propellant Topology

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	1197.68	251.46	21.00	1115
KmeansCp	1164.24	285.69	24.54	917
FCMCP	1183.74	248.80	21.02	724
HSACP	1196.76	210.53	17.59	852

Table 6.BS Location (50-175) – 50 nodes Propellant Topology

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	1020.58	176.72	17.32	764
KmeansCp	1012.22	211.40	20.88	994
FCMCP	1022.38	159.36	15.59	750
HSACP	999.06	137.38	13.75	617

Table 7.BS Location Center – 150 nodes 3 by 3

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	258.54	135.21	52.30	505
KmeansCp	244.23	132.80	54.37	522
FCMCP	248.97	125.44	50.38	431
HSACP	252.89	126.80	50.14	444

Table 8.BS Location Center (50-175) – 150 nodes 3 by 3

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	519.69	83.79	16.12	388
KmeansCp	503.63	94.38	18.74	484
FCMCP	521.18	72.64	13.94	284
HSACP	521.02	75.03	14.40	308

5. Conclusion

Various clustering protocols like LEACHC, KMeansCP, FCMCP and HSACP were compared especially for realistic topology deployments. HSACP and FCMCP creates similar types of clustering effects. However, in long run, data delivery analysis reveals that HSACP outperforms the other three protocols when it comes to data delivery especially for peculiar topologies like Cliff or Propellant. HSACP may report good amount of data, if applied to unknown or strange deployment conditions.

Table 9.BS Location (50-175) – 150 Propellant Topology

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	514.81	87.73	17.04	436
KmeansCp	493.19	86.36	17.51	392
FCMCP	504.26	81.49	16.16	345
HSACP	505.41	74.96	14.83	305

Table 10.BS Location Center – 150 nodes Propellant Topology

Protocol	Mean(μ)	StdDev(σ)	Cv(σ/μ)%	Max-Min(Data)
LeachC	279.73	152.04	54.35	550
KmeansCp	285.23	153.92	53.97	528
FCMCP	264.17	148.54	56.23	513
HSACP	283.99	149.24	52.55	518

References

[1.] Mohammad Ilyas and Imad Mahgoub. *Handbook of sensor networks: compact wireless and wired sensing systems*, Book, CRC Press, 2005: Chapter 1 & 6.

[2] Kemal Akkaya, Mohamed Younis. *A survey on routing protocols for wireless sensor networks*, in the Elsevier Ad Hoc Network Journal, 3.3 (July, 2005): 325-349.

[3] Jamal N. Al-Karaki,Ahmed E. Kamal,. *Routing Techniques In Wireless Sensor Networks: A Survey*. IEEE Wireless Communications, December 2004.

[4] D. J. Dechene, A. El Jardali, M. Luccini, and A. Sauer, *A Survey of Clustering Algorithms for Wireless Sensor Networks* ,Computer Communications, Butterworth-Heinemann Newton, MA USA, October 2007.

[5] W. B. Heinzelman, *Application-specific protocol architectures for wireless networks*, Ph.D. dissertation, Mass. Inst. Technol., 2000.

[6] W. B. Heinzelman, Anantha P. Chandrakasan,Hari Balakrishnan, *An Application-Specific Protocol Architecture for Wireless Microsensor Networks*, 2002.

[7] Lindsey, S., Raghavendra, C.*: PEGASIS: power-efficient gathering in sensor information systems.* IEEE Aerospace Conf. Proc.,3 (2002)

[8] Manjeshwar, A., Agarwal, D.P.: *TEEN: a routing protocol for enhanced efficiency in wireless sensor networks*. Proc. 15th Int. Parallel and Distributed Processing Symp (2001): 2009-2015.

[9] Manjeshwar, A., Agarwal, D.P.: *APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks*. Proc. Int. Parallel and Distributed Processing Symp.(2002): 195-202.

[10] Younis, O., Fahmy, S.: *HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,* IEEE Trans. Mob Comput., 3 (2004): 366-379.

[11] Liansheng Tan, Yanlin Gorg, Gong Chen, *A Balanced Parallel Clustering Protocol for Wireless Sensor Networks Using K-means techniques* , IEEE Second International Conference on Sensor Technologies and Applications, 2008.

[12] P. Sasikumar, Sibaram Khara, *K-Means Clustering In Wireless Sensor Networks*, IEEE Fourth International Conference on Computational Intelligence and Communication Networks, 2012.

[13] Sudhanshu Tyagi, NeerajKumar, *A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks*, Journal of Network and Computer Applications, 2012.

[14] T. Murata and H. Ishibuchi, *Performance evaluation of genetic algorithms for flowshop scheduling problems*, Proc. 1st IEEE Conf. Evolutionary Computation, 2 (June 1994): 812-817.

[15] Raghuvanshi, Ajay Singh and Tiwari, Sudarshan and Tripathi, Rajeev and Kishor, Nand, *Optimal number of clusters in wireless sensor networks: a FCM approach*, IJSNet, 2012

[16] Medhat, Fady and Ramadan, Rabie A and Talkhan, Ihab, LM, *Smart Clustering for Multimodal WSNs*, IEEE Seventh International Conference on Broadband, Wireless Computing, Communication and Applications, 2012
[17] Duc Chinh Hoang; Kumar, R.; Panda, S.K., *Realisation of a cluster-based protocol using fuzzy C-means algorithm for wireless sensor networks*, Wireless Sensor Systems, IET , 3.3 (September 2013): 163-171.
[18] Hoang, DC and Kumar, R and Panda, SK, *Fuzzy C-means clustering protocol for wireless sensor networks*, IEEE International Symposium on Industrial Electronics, 2010
[19] Geem, Zong Woo, Joong Hoon Kim, and G. V. Loganathan. *A new heuristic optimization algorithm: harmony search,* Simulation 76.2 (2001): 60-68.
[20] Hoang, DC and Kumar, R and Panda, SK , *A Robust Harmony Search Algorithm based Clustering Protocol for Wireless Sensor Networks*, IEEE International Conference on Communications Workshops (ICC), 2010
[21] Osama Mohd Alia, Zyad Shaaban, Ahmad Basheer, Alaa Al-Ajouri and Ahmed Alsswey, *Musicians'-Inspired Clustering Protocol for efficient Energy Wireless Sensor Networks*, IEEE International Conference Communications and Networking (ComNet), 2014
[22] Hoang, DC and Kumar, R and Yadav, P and Panda, SK ,*Real-Time Implementation of a Harmony Search Algorithm-Based Clustering Protocol for Energy-Efficient Wireless Sensor Networks*, IEEE Transactions on Industrial Informatics, 10.1, February 2014
[23] Camilo, Tiago and Silva, Jorge and Rodrigues, Andre and Boavida, Fernando, *GENSEN: A Topology Generator for Real Wireless Sensor Networks Deployment*, Software Technologies for Embedded and Ubiquitous Systems, Lecture Notes in Computer Science, 2007:436-445.