A Multi-Armed Bandit Problem-Based Clustering AlgorithmforWireless Sensor Network

Mahmoud Parvin & Mohammad Reza Meybodi

Computer Engineering Department, Islamic Azad University of Qazvin, Qazvin, Iran
E-mail: mahmoud.parvin@yahoo.com, meybodi@ce.aut.ac.ir

Abstract -Wireless Sensor Networks consist of tiny sensors which have limitations such as energy. These networks are utilized in a variety of applications which aids in boosting the network lifetime hence great amount of researches have been devoted to them. One of the key problems to achieve this goal is for the network to deploy clustering the nodes, thus the energy consumption of the nodes of the network is uniformed and optimal. In this paper, we have presented a Multi-Armed Bandit Problem-Based Clustering Algorithm, called MABCA, for Wireless Sensor Networks, which endeavours to recognize the Cluster Head in several phases, based on a variety of parameters like the amount of the energy of the node, the number of the neighbors etc. and organizes the clusters in a way that the energy consumed in the nodes would be balanced and optimal. To determine the effectiveness, the proposed algorithm was simulated using ns2 simulator and the results have been compared to that of other paramount clustering methods.

Keywords: Clustering, multi-armed bandit, network lifetime, Wireless sensor networks.

I. INTRODUCTION

Sensor networks are an especial type of computer networks that due to their high potential and vast applications have had a considerable growth. Such networks consist of thousands of tiny sensor nodes which are high potential, cost effective, limited processing and computing resources, low power which enables them to sense special attributes like damp, temperature, pressure etc. from the environment and transmit it to the Sink.

So, on the whole, these networks are noticeable in two ways:

- sensing especial parameters from the environment and
- communicating so as to pass their received packets to the Sink.

Sensor networks have many applications some of which include intelligent surveillance on expressways and areas which are difficult to reach, environmental monitoring and target tracking. These applications entail sensor networks to be deployed as wireless [2].

Basically these networks have static nodes or nodes with a limited movement and a Sink to which any node transmits its data either directly (single-hop) or indirectly (multi-hop). In a direct transmission, any sensor transmits its information to the sink directly hence consuming lots of energy for any transmission due to long distance between the sensors and the sink. On the contrary, those designs which shorten the communication distances could extend the duration of network lifetime. As a result, in these networks multi-hop communications are more fruitful and cost effective than that of single-hop.

This point is worth noting, however, that even in multi-hop communications the most energy of the nodes is dissipated as any sensor communicates with its neighbors which would in turn lead to faster energy depletion in sensor nodes. To solve this problem, we can group nodes into clusters.

An example of this application is found in hierarchical routing in which nodes are divided into logical clusters. In any cluster one node is the cluster head and the other nodes are considered as non-cluster head yet a member of the cluster.

The members of the cluster gather, from the environment, the expected information based on their application and then transmit the information to the cluster head. The cluster head after gathering this information transmits it to the Sink.

Clustering in sensor networks as an effective method of managing the nodes of the network enjoys huge advantages some of which are as follows: scalability, local routing, reducing the size of forwarding table for each node, gathering the data in cluster head which leads to the reduction of energy consumption of the nodes as well as reducing the overload of maintaining the network topology [1,2].

Owing to its advantages in sensor networks, clustering has been the subject of extensive research hence various clustering algorithms have been introduced for it.In this paper a Multi-Armed Bandit-based Cluster Algorithm head called MABCA is presented. The proposed algorithm is a distributed algorithm which operates independently of the size and the structure of the sensor network.

The remainder of the paper is organized as follows. Section II presents the survey of related works. Section IIIMulti- Armed Bandit Problem is explained briefly. In section IV the proposed Algorithm is discussed. Simulation results are explained in section V and section VI deals with conclusion.

II. RELATED WORKS

In this section, we will briefly overview some of the existing clustering algorithms for wireless sensor networks.

One of the first clustering algorithms introduced for sensor networks is the low energy adaptive clustering hierarchy (LEACH) algorithm[18]. The operation of LEACH is separated into two phases: the setup phase and the steady state phase. During the setup phase, a predetermined fraction of nodes p, elect themselves as cluster heads by comparing a chosen random number with a predefined threshold. After the cluster heads have been elected, they broadcast an advertisement message to the rest of the nodes in the network that they are the new cluster heads. Upon receiving this advertisement, all the non-cluster head nodes decide on the cluster to which they want to belong, based on the signal strength of the advertisement. The non-cluster head nodes inform the appropriate cluster heads that they will be members of the cluster[5].

There are algorithms which elect cluster heads based on a certain criterion such as number of neighbor nodes, or remaining energy of the nodes. In the algorithm presented in [3], each node waits for a random duration. After this period, if no message is received from a certain cluster head, the node states itself as a new cluster head. In some other methods, such as the one in [20], cluster heads are specified prior to network deployment. These specified nodes are placed in certain positions and other nodes are scattered around them. After the network deployment, these nodes try to form clusters having balanced traffic load. This means that fewer nodes are assigned to the clusters closer to the sink node and more nodes are assigned to the clusters far from the sink. In the sensor network considered in [4], two classes of nodes are available; sensor nodes and aggregator nodes. Aggregator nodes are placed in certain position and play the role of cluster heads. In this paper three different algorithms are presented for scattering sensor nodes around the aggregators[5].

In [8], sink node is assumed to be a mobile node which queries data from different part of the network based on its distance from different nodes. For this reason, it is required to have a dynamic clustering algorithm which can adapt itself to the changing position of the mobile sink, considering the overall energy consumption in the network.

A number of different routing algorithms for sensor networks are surveyed in [21]. One major class of these algorithms is the hierarchical routing. In hierarchical routing algorithms, at first a clustering scheme is applied to form a kind of hierarchy and afterward, the process of routing is divided into inter and intra-cluster phases. Some examples of these algorithms are "Fixed size cluster routing", TEEN4, and APTEEN5. AIMRP6 [18] assumes that the sink node is placed in the center of the network. In the setup phase, sink initiates a Tire message to the nodes in the radius of $a$around itself. These nodes constitute the first level. These nodes rebroadcast the Tier message to the nodes in the radius of $a$around themselves. This process continues until all nodes in the network find their level in the hierarchy to the sink node[5].

In [7] an algorithm called Hybrid Energy-Efficient Approach (HEED) has been given. This algorithm has four primary goals: (i) prolonging network lifetime, (ii) terminating the clustering process within a constant number of iterations/steps, (iii) minimizing control overhead (to be linear in the number of nodes), and (iv) producing well-distributed cluster heads and compact clusters. Before a node starts executing HEED, it sets its probability of becoming a cluster head,*CHprob*based on its residual energy. Using this probability, each node decides to become a tentative cluster head or not. Tentative cluster heads, declare themselves to their neighbors using a cluster-head message, and double their *CHprob*probability. Once *CHprob*in a node reaches 1, that node becomes a final cluster head. Nodes which are not cluster head (tentative or final), upon receiving a cluster-head message, join to the declared cluster. Nodes which do not receive ant cluster-head message after a while, decide to become cluster head[5].

In [20] an enhancement on HEED algorithm in which, only nodes with residual energy higher than a specified threshold can become cluster head is presented. Also at the end of the algorithm, when some nodes do not join to any cluster yet, unlike HEED in which all these nodes become cluster head, HEED algorithm is executed again for electing cluster heads among them.

III. MULTI-ARMED BANDIT (MAB)

The classical Multi-Armed Bandit Problem is a random, probabilistic decision-making problem which in any phase of its function there is K possible actions or functions to select. Selecting the function j makes observing the results possible. These observations could be considered as reward for the expected function and provide useful information about adopting a function in the future. An MAB is defined as a process of Multi-Armed Bandit involving a controller and some arms [10][11].

MAB controller in any instant of time t based on an idea named Gittins Index adopts an arm for action and does it. In this way active arm state changes according to its performance and other inactive arms remain frozen. Each armi, i = 1, 2,..., k, is described as

(1){ ($S_i(t)$), $R(S_i(t))$ }

where $S_i(t)$ is the state of i arm of the machine at time t and $R(S_i(t))$ is its reward in case it is chosen to perform in $S_i(t)$ state.
Since the controller can operate on exactly one arm at eachinstant time, the control action U(t) takes values in $\{e_1,e_2,\ldots,e_k\}$
, where ej = (0, 0,...,0, 1, 0,...,0,0)is a unit k-vector with 1 at the j^{th} position. So the modifications in the states of machines is determined as follows:

$$S_i(t+1) = \begin{cases} f & U_i(t) = 1, \quad (2) \\ S_i(t) & U_i(t) = 0, \end{cases}$$

armi generates a reward only when it is operated by the controller and in this way modifies the arm status. Let $R_i(t)$ denote the reward generated by armi at time t; then

$$R_i(S_i(t)) \begin{cases} & \text{if} U_i(t) = 1, (3) \\ 0 & \text{if } U_i(t) = 0, \end{cases}$$

A Scheduling policy is one in the form of $\pi=(\pi_1,\pi_2,\ldots)$ and consist of some decision making rules that in any instant of time t sets a value to a controlling word U(t) in the set $\{ e_1,e_2, \ldots ,e_k \}$ according to

$$U(t) = \pi_t(S_1(t), \ldots, S_k(t), U(t-1)) \quad (4)$$

The goal of MAB Problem is to represent an optimal Scheduling policy π in order to maximize the obtained reward. This theorem is formulated as follows :

$$J^{\pi} = E\left[\sum_{t=0}^{\infty}\beta^t \sum_{i=1}^{k} R(S_i(t),U_i(t))\middle|S(0)\right] \quad (5)$$

whereβ called the discount factorwhere $0 \leq\beta\leq 1$ and is used to give importance to the obtained rewards in early phase. Applying the Bellman's equation which shows the relationship between the value of the state of S and the immediate occurred state, the expected discountreward is calculated as following:

$$V_{s_i}(s_i(0)) = \max_{\tau>0} \frac{E\left[\sum_{t=0}^{\tau-1}\beta^t R_i(S_i(t))\middle|S_i(0)\right]}{E\left[\sum_{t=0}^{\tau-1}\beta^t \mid S_i(0)\right]} \quad (6)$$

Here $V_{Si}(S_i(0))$ is the maximum expected discounted reward per unit of expected discounted time that can be obtained at the first stage of a forward induction policy that continues the arm i with initial state $S_i(0)$. The corresponding stopping time $\tau_i(S_i(0))$is the first time at which the expected discounted reward per unit of expected discounted time equals $V_{Si}(S_i(0))$.

$$V_{s_j}(s_j(0)) = \max_i V_{s_i}(s_i(0)) \quad (7)$$

Consequently, at t = 0 an optimal forward induction policy continues an arm j such that:

Arm j is continued until $\tau j(sj(0)) - 1$. This constitutes the first stage of an optimal forward induction policy, and can be summarized as follows:

Step1: Determine $V_{Si}(S_i(0))$, i = 1, 2. . . k.

Step2 : Select an arm j such that

$$j = \arg\max_i V_{s_i}(s_i(0)) \quad (8)$$

Continue operating arm j until the minimum time that achieves the maximum in the right hand side of the above calculation[11].

IV. PROPOSED PROTOCOL

In this section we present the proposed clustering algorithm which is based on Multi-Armed Bandit Problem. In this algorithm it is assumed that all the sensor nodes are the same and each can take either the state of cluster head or non-cluster head. Since in any sensor network the optimal utilization of the energy and its reasonable performancein the node connectivity and the transmission of the data of every sensor node to the sink is taken into consideration, so criteria like the amount of the energy and the connectivity of the network is tested in choosing the status of the node, that is, being a cluster head or a member of a cluster; so that the more the level of energy of a node is, the more probability it will have to be chosen as cluster head. The why of the matter is that the energy consumption in the node with the status of cluster head is more than that of ordinary, so the choice of a cluster head with a higher level of energy reduces the probability of dying out of the node which in turn reduces the probability of ending the network lifetime. Also every node needs to be a cluster head to transmit the data to the Sink or act as a member of a cluster and in this way transmit its data to the cluster head and then cluster head transmits that data to the Sink.

Proposed algorithm, MABCA, by applying the local information of nodes endeavours to determine suitable clusters and clustering the network so that the number of clusters head as well as their distribution in the network environment is more judicious than that of LEACH which operates based on probabilities.

In this algorithm every node has a multi-armed bandit machine which endeavours to select the cluster head in a way that the energy consumption of the nodes is balanced and optimal hence the network lifetime is boosted.The bandit machine for each node includes *k* arm in relation to *k* neighbor and one arm for the node itself.In each phase of clustering the network, the bandit machine by choosing one arm amongst its k+1 arm determines the role of that node as a cluster head or a member of a specific cluster head.The selection of each of the nodes as cluster head depends on the Gittins Index obtained for its corresponding arm.

According to what was said, after the arrangement of the nodes of the network, clustering in proposed algorithm is implemented when two messages as H_1 and H_2 are flooded from each node. In the first phase each node through the emission of the message H_1 which includes the identifier and its amount of energy conveys to all of its neighbors in the transmission range its information. By so doing, every node upon receiving the H_1 message transmitted from its neighbors saves them in its clustering table. This table includes k record which contains the neighbors' information and is in relation to k arm of bandit machine of the expected node. Employing the received information, the Multi-Armed Bandit Machine of every node can, based on an index which is in relation to the amount of the energy of the neighbor nodes and the node itself, decide whether that very node should be chosen as cluster head or one of the neighbors with the topmost energy should be chosen as cluster head. Applying this policy, the energy consumption in the nodes of the network will be balanced.

In the proposed algorithm, in order for optimizing the energy consumption in the network, the second phase starts with the flooding of the message H_2 from each node. H_2 message encompasses the average energy of the neighbor nodes and also the number of the neighbor nodes.

Every node upon receiving the H_2 message from its neighbors updates its clustering table and puts the received H_2 message fields which include the number of the neighbors of each immediate node and their average energy, into their suitable corresponding record in the clustering table.

One of the criteria which have been taken into account in this algorithm, for a node to be chosen as cluster head, is the number of its neighbors. This criterion is accomplished using H_2 message so that the node which is in a position with the more neighbors has the more probability to be chosen as cluster head because in this way we can reduce the number of clusters head which consumed lots of energy.

The average energy of the neighbors which is achieved using the H_2 message is considered as a criterion in order to increase the probability of the nodes which are placed in areas with the topmost energy to be chosen as cluster head. This criterion can make the consumption of the energy of the nodes of the network more balanced hence reduce the probability of generating areas with the low energy amount and the early die out of the nodes of the network.Considering the above mentioned points, we can calculatethe Gittins index as the following calculation:

$$\gamma_i = c\frac{e_i}{e_s} + (1-c)(r_i) \quad (9)$$

In the above calculation e_i is the amount of energy of the corresponding neighbor node with i arm and e_s is the sum of the energy of all of the neighbors of that intended node which is saved into the clustering table through the reception of H_1 message and $r_i$value is calculated based on the received values from H_2 messages for any of the neighbors. Following calculation determines this value:

$$r_i = h\frac{a_i}{a_s} + (1-h)\left(\frac{n_i}{n_s}\right) \quad (10)$$

In this calculation α_i is the amount of average energy of the neighbors of the neighbor node corresponding to i arm and $\alpha_s$is the sum of the average energy of the neighbors of the all neighbor nodes of the intended node. Also n_i is the amount of the neighbors of the neighbor node corresponding to *i* arm and n_s is the average number of the neighbors of all neighbor nodes of the intended node. C and h coefficients are deployed in order to moderate the effect of each of the criteria used for indexing in the algorithm.

After indexing, the intended node chooses the cluster head based on the selected arm among the arms of its bandit machine which is the arm with the maximum Gittinsindex. If the arm corresponding to the node itself is chosen then the intended node will take the state of cluster head, otherwise the node corresponding to the selected arm is chosen as cluster head and the intended node introduces itself as a member of that node cluster. If the neighbor node which is chosen by the intended node as its cluster head has the cluster head status conveys a reception message to the intended node and saves its information into its cluster table as a member of the cluster; but in case the neighbor node selected as cluster head hasn't introduced itself as cluster head, the rejection message is issued and the intended node selects another arm with the higher Gittins index as a cluster head.

This trend continues until the intended node finds for itself a cluster head and introduces itself as one of its members. In case none of the neighbors act as cluster head for the intended node, the intended node chooses and introduces itself as a cluster head.

V. SIMULATION CONCLUSIONS

The proposed algorithm in this paper is simulated using ns_2simulator, and its effectiveness has been evaluated. It has been assessed based on criteria like the number of the clusters and data packets which could serve as an influential criterion for the network lifetime and the results were compared to that of clustering protocols like LEACH, HEED and ExtendedHEED.

To fulfill this goal, the results are simulated in several regions which have areas of 1000*1000meters having 50, 100, 200, 300, 400 and 500 nodes respectively. The Sink and the sensor nodes are assumed to be fixed during the simulation and the radius of the transmission of sender nodes are fixed during simulation and equals 25. The initial energy of all nodes is the same and equals 100 Joules.

As it was mentioned, two of the criteria including the number of the clusters and the network lifetime were used for assessment. So to arrive at the expected results, two experiments are carried out. Here go the results:

Considering the high energy consumption in the cluster head and the direct influence of their number on the reduction of the energy of the network thus reducing the network lifetime, we can use the number of the cluster as a criterion to assess the clustering algorithm. The proposed algorithm by considering the energy of each node and its position in the network endeavours to choose the number of clusters minimally and with a provident approach.

The Figure 1 demonstrates the number of the clusters generated by any of the algorithms like LEACH, HEED, ExtendedHEED and MABCA in this experiment. The results demonstrate the minimum of the number of the clusters generated in the proposed algorithm as compared to that of the other 3 algorithms.

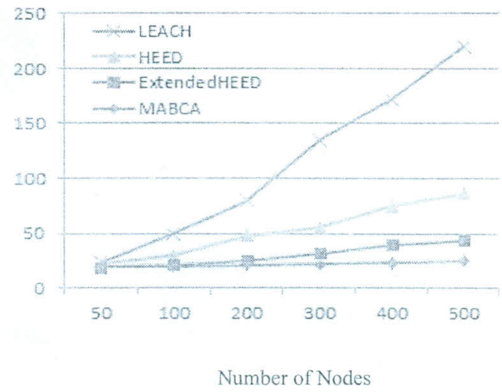

Figure 1: number of the Clustersfor four algorithms i.e. LEACH, HEED, ExtendedHEED, andMABCA.

In the second experiment the network lifetime is considered as a criterion for assessment. The reason for selecting this criterion is the important role the network lifetime plays in evaluating the quality of service for a sensor network.

As it is demonstrated in Figure 2, in the proposed algorithm the network lifetime is more than that of 3 other algorithms like LEACH, HEED and ExtendedHEED.

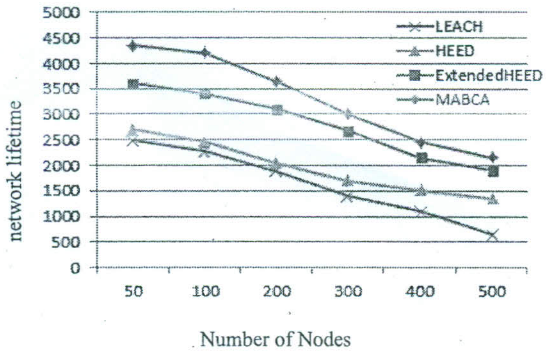

Figure 2 :network lifetime for four algorithms i.e. LEACH, HEED, ExtendedHEED, andMABCA.

CONCLUSION

In this paper a multi-armed bandit-based clustering algorithm for sensor networks called MABCA was presented. The proposed algorithm based on the idea of balancing and optimizing the energy consumption in the nodes of the network through the selection of the minimum number of cluster head and also selecting the cluster head on the basis of the suitable position of these nodes in the network, endeavours to boost the network lifetime and increase its qualityof service. Simulation results imply the superiority of the proposed algorithm over algorithms like LEACH, HEED and ExtendedHEED from the point of view of reducing the number of clusters and extending the network lifetime.

REFERENCES

[1] J. Yick, b. Mukherjee and d. Ghosal, *"Wireless Sensor Network Survey"*, Journal of Elsevier on Computer Networks, Vol. 52, 2008.

[2] F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, *"A Survey on Sensor Networks"*, Of IEEE Communications Magazine, Vol. 40, 2002.

[3] M. Younis, M. Youssef and K. Arisha,*"Energy-Awaremanagement in cluster-based sensor Networks"*. TheInternational Journal on Computer Networks, Vol. 43, No.5, December 2003, 649-668.

[4] I. S. Misra, S. Dolui and A. Das, "*Enhanced-Efficientadaptive clustering protocol for distributed sensornetworks*". ICON 2005.

[5] M. Esnaashari and M. R. Meybodi, "A novel clustering algorithm for wireless sensor networks using Irregular Cellular Learning Automata", Proceedings of theInternational Symposium on Telecommunications,2008.

[6] S. Madiraju, C. Mallanda, R. Kannan, A. Durresi and S. S. Iyengar, "*EBRP: Energy Band based Routing Protocol for Wireless Sensor Networks*", Department of Computer Science, Louisiana State University 298 Coates Hall LA 70803, USA,2004.

[7] O. Younis and S. Fahmy, "*Distributed Clustering inAd-hoc Sensor Networks: A Hybrid, Energy-EfficientApproach*", In Proc. of IEEE INFOCOM, March2004.

[8] M. Lotfinezhad and B. Liang, "*Energy Efficient Clustering in Sensor Networks with Mobile Agents*",In Proc. of the IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, Louisiana, March 2005.

[9] E. Frostig and G. Weiss, "*Four proofs of Gittins' multi-armed bandit theorem. Technical report*", The University of Haifa, Mount Carmel, 31905, USA, November 1999.

[10] J. C. Gittins, *"Multi-Armed Bandit Allocation Indices"*, John Wiley and Sons, New York, NY, 1989.

[11] A. Mahajan and D. Teneketzis, "*Multi-Armed Bandit Problems*", Foundations and Applications of Sensor Management, pp.121-151 ,2008.

[12] V. Krishnamurthy, "*Emission Management For Low Probability Intercept Sensors In Network Centric Warfare*", of IEEE Transactions on Aerospace and Electronic Systems, Vol. 41, 2005.

[13] C. Yunxia, Z. Qing, V. Krishnamurthy and D. Djonin, "*Transmission Scheduling for Optimizing Sensor Network Lifetime: A Stochastic Shortest Path Approach*", Of IEEE Transactions on Signal Processing, Vol. 55, 2007.

[14] J. Mersereau, P. Rusmevichientong and J. N. Tsitsiklis, "*A Structured Multiarmed Bandit Problem and the Greedy Policy*", In Proc. Of 47th IEEE Conference on Decision and Control, Mexico, 2008.

[15] V. Phan, Y. Park, H. H. Choi, J. Cho and J. G. Kim, "*An Energy-Efficient Transmission Strategy For Wireless Sensor Networks*", Of IEEE Transactions On Consumer Electronics, Vol. 56 , 2010

[16] *http://www.isi.edu/nsnam/ns/.*

[17] S. Bandyopadhyay and E. Coyle, "*An Energy-EfficientHierarchical Clustering Algorithm for Wireless SensorNetworks*", Proceedings of IEEE INFOCOM, 2003.

[18] A.A. Abbasi and M. Younis, "*A Survey on Clustering Algorithms for Wireless Sensor Networks*", In Proc.of ComputerCommunication 30, 2007.

[19] S. Banerjee and S. Khuller, "*A Clustering Scheme for Hierarchical Control in Multi-hop Wireless Networks*", Proceedings of IEEE INFOCOM, 2001.

[20] H. Huang, J. Wu, "*A Probabilistic Clustering Algorithm in Wireless Sensor Networks*", In Proc. of IEEE 62nd Semiannual Vehicular Technology Conference (VTC), Sept. 2005.

[21] K. Akkaya, M. Younis, "*A survey on routing protocols for wireless sensor networks*", *Elsevier AdHoc Network Journal, pp. 325-349,* 2005

◈◈◈