Improvment Multiplicity of Routs in Directed Diffusion by Learning Automata

New Approach in Directed Diffusion

Ali Sayyad
Msc. Information Technology, Network Eng. Computer & Elec. Department
Islamic Azad University of Qazvin
Qazvin,Iran
later_opt@yahoo.com

Ali Ahmadi
Msc. Information Technology, Network Eng. Computer & Elec. Department
Islamic Azad University of Qazvin
Qazvin,Iran
Ali_Ahmadi@gmail.com

Mohammad Shojafar
Msc. Computer Eng. Computer & Elec. Department
Islamic Azad University of Qazvin
Qazvin,Iran
Shojafar@qiau.ac.ir

Mohammad Reza Meybodi
PHD in Computer science. Computer & IT. Department
Amirkabir University,Tehran,Iran
mmeybodi@aut.ac.ir

***Abstract*—one of the important and challenging matters in sensor network is energy of life span of nodes in the network. Directed routing algorithm is one of propounded methods in sensor networks which are a data-oriented algorithm. This algorithm focuses on saving energy within life span of network nodes. One of problems of directed diffusion method is existence of multiple routes. Now, consider that some sinks from the same origin request the same data who's Data Volume is very much. Directed routing algorithm establishes one route toward targeted route for each query. The problem of this algorithm is multiplicity of routes for the same data. Therefore, if we can establish a route which has the most common feature with regards to nodes which forms the route, we have prevented wasting energy.**

In this paper, it is tried to remove problem of multiplicity of routes for the same data by learning automata. We named this algorithm RDDLA. RDDLA decrease overhead and energy in the network Considerable against with some others methods.

Keywords-component; Wireless Sensor Network; Directed Diffusion; Learning Automata; Multiplicity Routing

I. INTRODUCTION

Current progress in technology of short radio range and micro electro mechanical system leads in production intelligent sensors which are generally calculated by capability of data sense, wireless connection in device is strongly limited with regards to equipped resources. These tiny sensors can be considered for gathering information and they can be distributed to central controlling units named sink in different regions. For transferring sensed data in the network to sink, many wireless steps may be taken through a collection of small and limited sensor nodes with regards to energy. This characteristic is important for diffusing data in wireless sensor network.

Within previous years, different algorithms and protocols were suggested with goal of achieving more efficiency and reliability for diffusing data in wireless sensor networks. For example, flooding diffusion is the most reliable method for sending data from sensor to a sink. This protocol doesn't impose any extra cost for keeping topology and discovery of the route and its implementation is very simple and easy. But, main problem of this protocol is creating too much overhead in the network due to sending repeated messages. This problem specially leads to inefficiency in energy consumption. A new data-oriented method which was different from traditional routing was created for wireless sensor networks in order to solve this problem. The main goal for generating this method is because of concentrating of sensor network on the data generated with sensors and it doesn't focus on sensed nodes. Data-oriented protocols are based on query and are dependent on naming (for instance mentioning data with binary of trait-value) of desirable data. Therefore, too much transfer of many packs is prevented and it causes energy efficiency.

Direct diffusion[1] is an instance of one of data-oriented protocols because of its characteristics for saving energy, such as, data query based on query (named interest) by sinks, collecting and saving data by sensor and mechanism of gradient and route enhancement. Gradient concept usually means direction toward those neighboring by which sink is accessible.

In wireless sensor network, most data packs are sent toward sink from a sensor complex. Therefore, here, each sensor node's important task is to create and keep gradient value in each node. Totally gradient value management is done by primary and frequent primitive flooding diffusion of a series of a controlling pack (like interest packs in direct diffusion [1]) from a sink. Please note that this frequent flooding diffusion all over the network leads in too much overhead (regarding band width and energy consumption) in sensor networks. Moreover, when network topology change because of failure of senor nodes, wireless connections and value of some gradients will become unreliable, thus, there is need to frequent flooding diffusion.

But the point which we are going to discuss in this Paper is eliminating one of deficiencies of this algorithm. Existence

978-1-4244-8845-2/10/ $ 26.00 © 2010 IEEE

of multiple routes can be mentioned as one of problems of directed diffusion method. Imagine it is to send the same data to several sinks. Therefore a separate rout is made for each sink. Multiple routes means that there are some distinct routes from origin toward sink and each route engages especial nodes of that route. The point here is that although the same data is being transferred to several sinks but it uses several routes. Therefore if we can share a route between node of origin and several sinks, it causes repetition of data panel only from one route. So instead of engaging many nodes in the route, fewer nodes participate in transfer and this fact results in saving lots of energy in the network. In means only chosen nodes are engaged in one route till the time that their energy reaches to a threshold and changing route occurs. We want to make a shared route which is optimal one. For reaching this, we use an intelligent method named learning automata. Learning automata chooses the best route in each period of time in order to have least energy consumption in the network.

In this paper, we have used learning automata which is considered as an intelligent method in order to amend inefficiency of direct diffusion towards dynamism and works done regarding wireless sensor networks have been briefly explained in the next parts. In the third part, there is a simple introduction to the structure of standard learning automata. Region directed diffusion algorithm has been discussed in detail by learning automata or RDDLA. Dramatization of RDDLA with some previous algorithm have been discussed in detail in part five and there is comparison based on considered quality parameters. Finally, in the last parts there are conclusion and future works.

II. Related Works

Recently some researchers have been done on routing in wireless sensor networks. One of the famous protocols in this field is directed diffusion.

The most basic method is directed diffusion method which has been presented by Fabio Silva. This method uses three phases in its routing. Interest phase, discovery phase and choosing the best route phase. Two-Phase attraction method is a very suitable method for some of applications but in opposite it acts very weakly for some other applications. Especially, those applications in which there are many sources and receivers and receivers are such connected to each other that traffic data volume will much increased.

For solving this problem, push diffusion method [2] is brought up in which role of information diffusers and data users are changed and it causes data sources to actively look for users. Advantage of push diffusion is distraction of control data in it in order to find receives. The problem of this method is that it doesn't act well if the numbers of sources are many and numbers of receivers OR central nodes are few.

One Phase Pull method [3] was presented for improving this problem. This idea is based on requester who omits one of Phase of two Phase Pull methods. The application of this method when number of sources is many and number of receivers is few is done.

Another improvement of directed diffusion is the method based on diameter of hexagon [4]. In this method, sensor nodes are ordered as a beehive with a fixed topology and main information is sent and received on the diameter of hexagon. In fact, available hexagons are considered as basic routs of data. In order to distribute energy consumption among sensors in this method, basic routs are periodically changed. For more decrease in energy consumption, we activate sensors when sending data and deactivate them when they are idol. It has excellent efficiency in comparison with basic diffusion when responding to an interest with regards to energy consumption rate and delay.

Another improvement is directed diffusion energy productivity by use of Passive clustering [5]. With regards to huge cost that data flooding diffusion imposes on energy resources, a method was suggested for preventing energy waste which is named passive clustering. Basic job of clusters is optimizing of flooding message exchanges which prevents it's too much overhead. Normal clustering is done by use of flooding diffusion of controlling messages when there is not any traffic in the network, therefore, it results are in energy waste for keeping cluster structure. In fact, passive clustering is a suitable mechanism for increasing efficiency of the whole network. Simulation results show that directed diffusion along with passive clustering has better performance than basic diffusion with regards to transfer rate, network density increase and delay.

Another improvement which has been realized in directed diffusion is E-Span method and LPT [6] for data density in wireless sensor networks. When an event is observed in a special zone in sensor networks, data is sent toward central node or destination by available sources in that zone. Data of these sources are usually accumulated on the route. Data density decreases communication of overhead cost and increases energy productivity.

In suggested structure, available sources in event zone form a fixed tree structure named LPT in order to ease data density. This method is used for increasing life span of sources which frequently transfer data report. In E-Span, sources which have the most remained energy are chosen as root and nodes of another source choose their parent node from their neighbors based on remained energy information and distance from root, then in LPT those nodes which have more remained energy are chosen as accumulator parent.

This method includes self-amendment method by which the tree is reconstructed whenever a node stops or a failed connection is discovered. Simulation results show that whenever LPT algorithm is used for data density in directed diffusion, source's life span is considerably increased. Another improvement is light directed diffusion [7].

For Decreasing network overhead, number of exchanged packs for periodical diffusion of interests and delivery of discovered data to central node whose result is energy consumption reduce because of decreasing exchanged controlling message and as a result increase of network life span.

Light directed diffusion [8] is based on this idea that it produces a spars logical topology by simple local regulations and then directed diffusion is executed on this topology.

III. LEARNING AUTOMATA

Learning Automata are adaptive decision-making devices operating on unknown random environments. A Learning Automaton has a finite set of actions and each action has a certain probability (unknown to the automaton) of getting rewarded by the environment of the automaton. The aim is to learn to choose the optimal action (i.e. the action with the highest probability of being rewarded) through repeated interaction on the system. If the learning algorithm is chosen properly, then the iterative process of interacting on the environment can be made to result in selection of the optimal action.

Fig. 1 illustrates how a stochastic automaton works in feedback connection with a random environment. Learning Automata can be classified into two main families: fixed structure learning automata and Variable Structure Learning Automata (VSLA). In the following, the variable structure learning automata which will be used in this Situation is described [9].

A VSLA is a quintuple $(\alpha, \beta, p, T(\alpha, \beta, p))$, where a, β, p are an action set with s actions, an environment response set and the probability set p containing s probabilities, each being the probability of performing every action in the current Internal automaton state, respectively. If the response of the environment takes binary values learning automata model is P-model and if it takes finite output set with more than two elements that take values in the interval [0, 1], such a model is referred to as Q-model, and when the output of the environment is a continuous variable in the interval [0, 1], it is refer to as S-model. The function of T is the reinforcement algorithm, which modifies the action probability vector p with respect to the performed action and received response. Assume β € [0, 1]. A general linear schema for updating action probabilities can be represented as follows. Let action i be performed then (1), (2) shows as follows [10]:

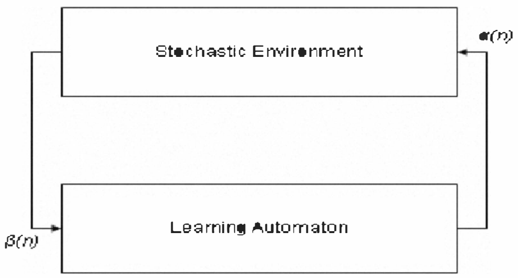


Figure 1. Learning Automata and Environment

$$p_i(n+1) = p_i(n) + a[1 - p_i(n)]$$
$$p_j(n+1) = (1-a)p_j(n) \qquad \forall j \quad j \neq i \qquad \text{Desirable (1)}$$

$$p_i(n+1) = (1-b)p_i(n)$$
$$p_j(n+1) = (b/r-1) + (1-b)p_j(n) \quad \forall j \quad j \neq i \qquad \text{Non Desirable (2)}$$

Where a and b are reward and penalty parameters. When a=b, the automaton is called LRP. If b=0 the automaton is called LRI and if 0<b<<a<1, the automaton is called LReP [11].

IV. REGION DIRECTED DIFUSION LEARNING AUTOMATA(RDDLA)

This Approach is used in cases in which there is one source and more than one receiver or Sinks in the network.

We named this algorithm RDDLA and R stands for Region. In Basic Directed Diffusion, a Path from Source toward Sinks is made by directed diffusion (Fig. 2). But, in RDDLA, First, whole network divided in to various regions by Booly algorithm [12] that illustrated in Fig. 3. In Booly algorithm, those nodes which are very close to each other are placed in the same region. Supervisor is the node whose energy is more than other nodes in the region.

This segmentation has two outcomes. The First outcome, removing nodes overlap, in fact, The nodes are in a region don't simultaneously show reaction to a received pack and only supervisor node as the representative of all, does the transfer action, this action, results preventing decreasing in nodes' energy (the goal, is to prevent wasted energy). The Second outcome, we use region supervisor to increase speed of transferring data.

The goal of this method is finding optimal method based on current Path for transferring data to Sinks. In this Approach, an intermediate node considered which information can be sent to sink via this node. Here, intermediate node is the one which is close to Sinks and has often more energy for sending and delivery of packs to Sinks. Role of Learning Automata is finding the best intermediate node. Region of pack delivery is the region in which packs are sent from Source to Sinks. In fact, this node considered as a virtual intermediate sinks. After finding virtual Sinks, packs can pass from a Path on which intermediate node (virtual Sinks) exists and then virtual Sinks can transfer the pack to real sinks.

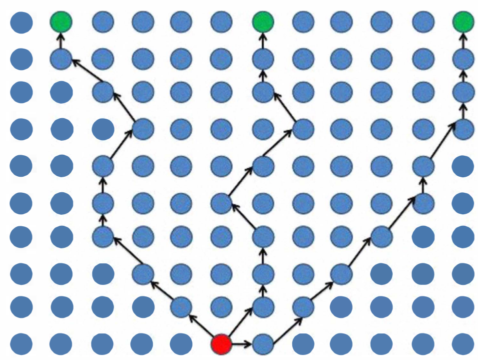


Figure 2. Three paths and three sink with one source illustrated

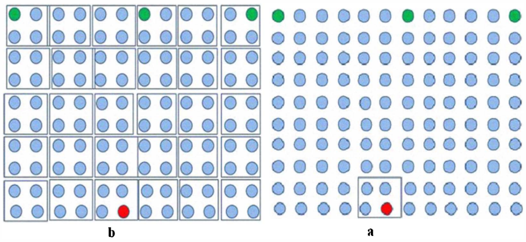


Figure 3. a: before rejoining and b: after rejoining

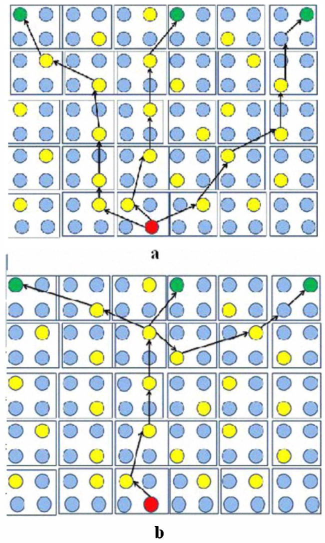


Figure 4. a: illustrated paths before using RDDLA, b: illustrated paths after using RDDLA

In suggested algorithm, the packs for a data have a code or unique ID which is produced by source. ID is used for determining data type and each node can transfer different data. Each region has an HOP. HOP is number of steps of each region of source. Here, attach fields of ID and HOP to the sent pack and each region sends it to its neighboring regions. This algorithm uses Stochastic Learning Automata. Here, we have learning automata whose reward rate is 0.2 and punishment rate is 0.07. Fig. 4 shows algorithm completely.

When packs arrive to sink from source, sink also confirms message of receiving the packs. Therefore, if there are some separate paths from source toward sinks, there will receive confirmation message of receiving from some different directions. By receiving separate messages, source produces updating pack on the available Paths. This updating pack informs intermediate nodes that the same data is being sent from several paths and it asks nodes to amend their sending path.

Each Node has its Learning Automata. After receiving this message, each intermediate node executes its learning automata. The number of action of each automata is based on its surrounding region. For example, as you see in Fig. 5 node Y has six actions because there are 6 regions around it. In continuation, each supervisor node (yellow node) sends a pack to its neighbors (surrounding regions). This region is broadcasted. Supervisor states its HOP in this pack. Each neighbor also sends this pack to all till the time that this pack reaches to the node which has the same ID with this pack.

When this packets reached to sink, sink send a packet as an answer to source of packets. Pack sending is broadcast. Pack sending is broadcast. Therefore, sender node receives response message pack from some different paths. If numbers of steps of sender to original source is less than steps of respondent a positive reward is considered for the zone from which pack is received and sender requests to have relation with message respondent.

The paths which have more steps than sender are punished. See the below example for clarification. Source Node which is named S is sending pack to its sinks. Source finds out that the same data is being transferred in several paths. Therefore, it sends updating place to intermediate nodes. For example node Y executes its automata. This node sends its step distance to source and also its ID to its neighbors. After several broadcasting pack Reach to node X. Node X also sends number of its steps to Node Y. since number of both nodes from source is 2, no path is formed (Fig. 5). In Another Example in Fig. 6 distance of Node Y to source is more than distance of Node X to source. Therefore, Node Y sends a pack to Node X in order to make a Path. Node X received the pack from several paths. Node X rewards the Path through which it has received faster the pack from Node Y. As a result, after some repetitions of This Actions a Path from Node X to Node Y is made whose steps to source are much less. Consider this for all supervisors of Path (Fig. 7). After sometimes sending and receiving the message and learning its automata, multiple paths are removed and eventually a region named intermediate region is made whose steps are much less in comparison with the time before execution of algorithm. At the end, when an optimal path is made, previous path will be removed in such a way that current node for example Node Y sends a message to source Node. Source Node doesn't send a message through the previous path. Fig. 7-b shows that after complete execution of algorithm, long multiple paths are removed and almost one path with one intermediate has been made.

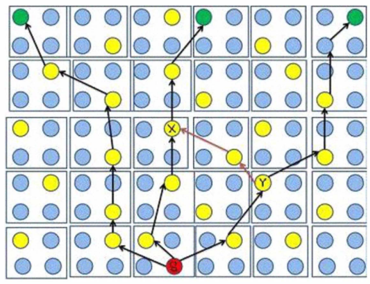
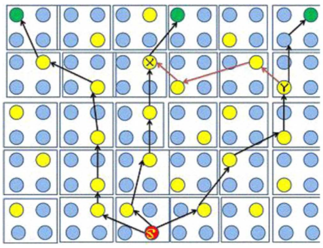
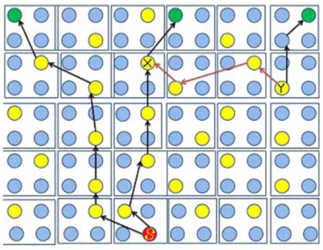


Figure 5. Distance(s, y) = distance(y, x) = 2, so x would not selected as Interface node

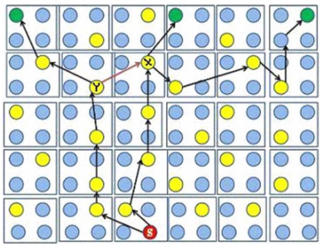


a

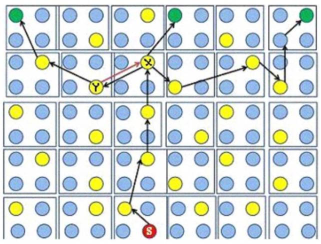


b

Figure 6. a: Distance(s, y)=4 > distance(y, x) = 3, so x would selected as Interface node b: eliminate all region from s to y



a



b

Figure 7. a: Distance(s, y)=4 > distance(y, x) = 1, so x would selected as Interface node b: eliminate all region from s to

As a result, as it has been showed in Fig. 7, number of engaged nodes for sending message has become much less in comparison with the past and also message arrives to sinks by less steps.

In next Session We Explain Simulation and Comparison Conclusion.

V. SIMULATION AND COMPARISON CONCLUSION

In this part, we see result of Simulation for each scenario which has been discussed in previous part and relevant figure is shown. For Simulation software of NS2[13,14] has been used whose model of energy is considered 0.660W for sending and 0.395W for receiving data which is according to energy in card of PCM-CIA-WLAN which is introduced in NS2. Then results of each Simulation have been assessed.

The important point is that we have implemented the algorithm by standard method of automata and then we have compared all of them to each other. In all methods, we turn on a chronometer and make decision based on it. So, in suggested algorithm, if the time of delivered pack is less than previous mode desirable situation has been occurred but if the time becomes more, undesirable situation has occurred. In these models we have considered 0-2 as reward and 0.07 as punishment. In this experiment, number of nodes is 100 and they are distributed in the area of 160*160 sqm. Sinks and one source are placed in this region. Here we have compared one-stage attraction algorithms with suggested algorithm. Compared parameters are network overhead rate, calculation of average delay of packs, losing and failure rate of packs, calculation of remained energy.

A. Network Overhead Rate

Fig. 8 shows network overhead rate per unit of time. Two-Phase Pull method has 3 phases. So, it has the most overhead in comparison with others. One- Phase Pull method in discovery phase is much faster than Two- Phase Pull method. Therefore, it has less overhead. But, in suggested method, at First, there is more overhead due to construction of region and execution of Learning Automata machine, but, after sometime and finding optimal path, overhead rate decreases quickly.

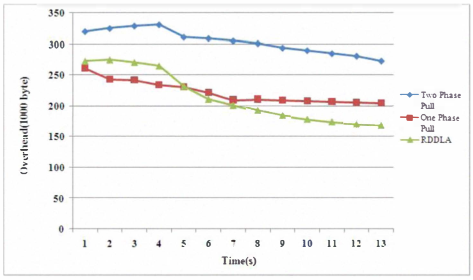


Figure 8. Overhead Rate in Time

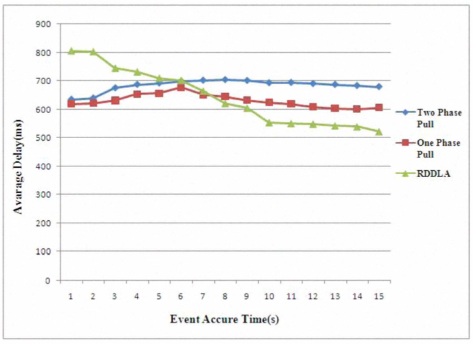


Figure 9. Packet Average delay in time

B. Packs Average Delay

First, average delay of packs in our method is more than two other methods and the reason is that calculation rate in RDDLA is more than others. After passing the time and increasing the packs in the network, primary effect of delay decreases and eventually it reaches its Minimum possible amount. By event accurate time we mean the time that packs are sent toward Sinks and average delay means the time that it takes the packs to arrive to Sinks. Fig. 9 has shown data delivery delay per unit of time.

C. Losing and Failure Rate of Packs

There are several paths for data in One-Phase Pull and Two-Phase Pull methods, so, more nodes are engaged in transferring in comparison with suggested method. Therefore, probability of failure of packs and losing them in suggested method is much less in comparison with two previous methods because we deal with less nodes and also less regions. This matter is very obvious in wireless networks because error rate in wireless networks is very high. In Fig. 10, a few packs in RDDLA has been failed, this amount has had 60% decreases in comparison with two previous methods. So, more than 40% of packs have arrived to Sinks completely and without any deficiency over the time.

D. Network Consumed Energy

Fig. 11 shows consumed energy per until of time and One-Phase and Two-Phase Pull methods have various paths for delivery of packs, so because of engagement of more nodes consumed energy in the network is high. But in the suggested method we have reduced energy consumption for 30% by deletion of extra regions in sending the packs.

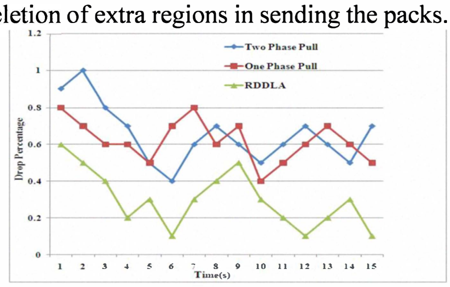


Figure 10. Packet loss in network

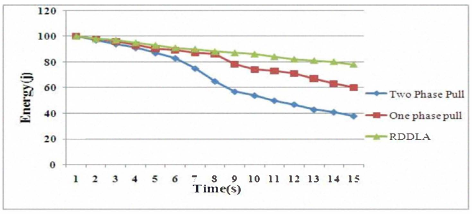


Figure 11. Energy remaining in time

VI. CONCLUSION AND FUTURE WORK

Our assessment shows that basic directed diffusion has the discussed problems and limitations. After checking basic diffusion it has been revealed that this protocol doesn't do any attempts for preventing multiple paths. Therefore, when using it energy of all network nodes are consumed simultaneously. But, in case of existence a supervisor which is chosen by learning automata, we can reduce energy consumption in the whole network and this is an issue which has been proved by Simulation. In fact about 30% energy in the network has been saved which can make the network more stable. We could even choose the supervisor such that is caused the nodes to change to quite status for a while when sending data and this also resulted in energy saving and prevented nodes evacuation.

For Future work that we work on it, we can raise this idea that supervisor can change place from one region to another in order to cause new supervisor to attend the region with more energy and more energy balance is made in the network and supervisor nodes endure more.

REFERENCES

[1] Ch. Intanagonwiwat, R. Govindan, D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks", IEEE/ACM Transactions on Networking (TON), Vol. 11, Issue 1, ISSN: 1063-6692, pp. 2-16, 2003.

[2] I.F. Akyildiz, W. Su, E. Cayirici, *"Wireless sensor networks: a survey"*, Vol. 38, Issue 4, pp. 393-422, 2002.

[3] J. Heidemann, F. Silva, D. Estrin, *"Matching Data Dissemination Algorithms to Application Requirements"*, Proceedings of the First ACM International Conference on Embedded Networked Sensor Systems (SenSys'03), ISBN: 1-58113-707-9, LA, USA, pp. 218-229, 2003. DOI: http://doi.acm.org/10.1145/958491.958517

[4] Y. Sh. Chen, Yau-Wen Nian, J.P. Sheu, " An Energy-Efficient Diagonal-Based Directed Diffusion for Wireless Sensor Networks", Proceedings of the 9th International Conference on Parallel and Distributed Systems(ICPADS'02),IEEE Computer Society, ISBN:0-7695-1760-9, USA, pp.445-460, 2002.

[5] A. Köpke, Ch. Frank, H. Karl, V. Handziski, W. Drytkiewicz, "Improving the Energy Efficiency of Directed Diffusion Using Passive Clustering", Proceedings of First European Workshop in Wireless Sensor Networks (EWSN), ISSN:0302-9743, ISBN: 978-3-540-20825-9, Springer Berlin, pp.172-187, 2004.

[6] M. Lee, V.W.S. Wong, "LPT for Data Aggregation in Wireless Sensor Networks", Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), BC, USA, Vol. 5, pp.29-74, 2005. DOI: http://dx.doi.org/10.1109/GLOCOM.2005.1578302

[7] A. Marcucci, M. Nati, C. Petrioli, A. Vitaletti, "Directed diffusion light: low overhead data dissemination in wireless sensor networks", 61st IEEE Conference in Vehicular Technology (VTC 2005-Spring), ISSN: 1550-2252, Vol. 4, pp. 2538-2545, 2005. DOI: http://dx.doi.org/10.1109/VETECS.2005.1543793

[8] A. Marcucci, M. Nati, C. Petrioli, A. Vitaletti, "Directed diffusion light: low overhead data dissemination in wireless Sensor Networks", 61st IEEE Vehicular Technology Conference, VTC 2005-Spring, ISSN: 1550-2252, Vol. 4, pp.2538-2545, 2005.

[9] K. S. Narendra, M. A. L. Thathachar, Learning automata: An introduction, Prentice Hall, ISBN: 0-13-485558-2, Pages: 476, 1989.

[10] H. Beigy, Intelligent Channel Assignment in Cellular Networks: A learning Automata Approach, A Dissertation for Doctor of Philosophy in CS Amirkabir University, Tehran, Iran, 2004.

[11] A.S. Poznyak, K. Najim, Learning Automata and Stochastic Optimization, ISBN 3-540-76154-3 Springer-Verlag Berlin Heidelberg New York, 1997.

[12] H. Garcia-Molina," Election in a Distributed Computing system", IEEE Transaction on Computers, C-31(1),pp.48-59, 1982.

[13] D. Ganesan, R. Govindan, S. Shenker, D. Estrin, "Highly Resilient Energy-efficient Multi path Routing in Wireless Sensor Networks", Proceedings of ACM MOBIHOC, ISSN:1559-1662, Vol.5, Issue 4, pp. 11-24, 2001. DOI: http://doi.acm.org/10.1145/509506.509514

[14] Ns – network simulator version 2, http://www.isi.edu/nsnam