

# ENERGY OPTIMIZATION USING ENERGY-EFFICIENT SELF-ORGANIZED CLUSTERING MODEL WITH SPLITTING AND MERGING FOR WSN

S.Kirubakaran , M.Karthik

Department of IT, Sri Venkateshwara College of Engineering, Pennalur, India.

## ABSTRACT

*This paper deals with a novel forwarding scheme for wireless sensor networks aimed at combining low computational complexity and high performance in terms of energy efficiency and reliability. In this paper, we propose an energy-efficient self-organized clustering model with splitting and merging (EECSM), which performs clustering and then splits and merges clusters for energy-efficient cluster-based routing. EECSM uses information of the energy state of sensor nodes, in order to reduce energy consumption and maintain load balance. An analytical model for estimating the energy efficiency of the scheme is presented, and several practical issues such as the effect of unreliable channels, topology changes, and MAC overhead are discussed. The results obtained show that the proposed algorithm outperforms traditional approaches in terms of power saving, simplicity, and fair distribution of energy consumption among all nodes in the network.*

*Key Terms: - EECSM, energy efficiency, packet splitting, reliability, wireless sensor networks.*

## I. INTRODUCTION

A wireless sensor network (WSN) is composed of a large number of low-cost devices distributed over a geographic area. Sensor nodes have limited processing capabilities, therefore a simplified protocol architecture should be designed so as to make communications simple and efficient. Moreover, usually the power supply unit is based on an energy-limited battery, therefore solutions elaborated for these networks should be aimed at minimizing the energy consumption. the battery of each wireless sensor node cannot be recharged or replaced, and thus all batteries have to be well managed, in order to provide long network lifetime and reduce energy consumption for WSNs [1, 2]. In general ,data packets from the source node have to traverse multiple hops before they reach the destination. In recent years, in order to guarantee the WSNs survivability and increase the network lifetime in such special-purpose environments, various energy-efficient schemes have been proposed in the literature [9-15]. To this purpose, several works have shown that energy consumption is mainly due to data transmission, and accordingly energy conservation schemes have been proposed aimed at minimizing the energy consumption of the radio interface.

For the energy-efficient transmission of information obtained by sensor nodes, many routing protocols have been developed. The routing protocol decides the transmission route

for a data packet, from the sensor node to the external final destination BS. In general, routing protocols in WSNs can be divided into *flat routing*, *cluster-based routing*, and *location-based routing*, depending on the network structure. Sensor nodes of flat routing typically have identical roles, whereas sensor nodes of cluster-based routing have different roles, in order to create hierarchy. Meanwhile, location-based routing uses the location of sensor nodes for routing. Among these routing protocols, cluster-based routing protocol is the most energy efficient [3]; therefore, we study cluster-based routing in this paper.

In cluster-based routing, a network consists of several clusters, and each cluster is comprised of a *cluster head* (CH) and many *cluster members* (CMs). Due to the limitation of a wireless sensor node, a centralized clustering and routing method is almost impossible in WSNs. Considering these limitations, a distributed clustering and routing method is normally used for WSNs. In distributed clustering and routing, each sensor node can reduce energy consumption, by clustering when transmitting the data packets. In this paper, we propose an *energy-efficient self-organized clustering with splitting and merging* (EECSM) model for WSN. In order to increase energy efficiency, EECSM performs clustering, by considering the energy state information and also by deciding on the split or merge clusters. This action is performed by monitoring the size of the clusters in self organized ways. In addition, several protocols and approaches, for minimizing energy consumption, have been proposed by authors such as [1-8].

## II. METHODOLOGY

### A. Minimizing the energy consumption

The aim of reducing energy consumption while taking the algorithmic complexity into account, proposed a novel approach that burst the original messages into several packets, such that each node in the network will forward only small sub packets. The splitting procedure is achieved applying the EECSM algorithm. The sink node, once all sub packets (called splitting components) are received correctly, will recombine them, thus reconstructing the original message. The splitting procedure is especially helpful for those forwarding nodes that are more solicited than others due to their position inside the network.

The proposed approach, almost all nodes operate as in a classical forwarding algorithm and, with the exception of the sink, a few low-complex arithmetic operations are needed. The sink node is computationally and energetically more equipped than the other sensor nodes, the overall complexity remains low and suitable for a WSN. Moreover, the proposed technique does not require the use of disjoint paths. Some preliminary results of this approach have been presented but they were empirical only and obtained through simulations on a sensor network where it is assumed that an ideal communication among neighbour sensors occurs, and all the burst components can be received correctly between a pair of nodes.

Through analytical model allows us to derive accurate results regarding energy consumption and complexity. The effect of important parameters such as nodes density and transmission range through both extensive simulations and an analytical study of the trade off between energy saving, complexity, and reliability of the proposed technique. In WSN, data

generated by each node, are to travel from multiple sources to a data recipient or sink rather than the communication between nodes.

### B. Packet Forwarding

A sensor network where sensor nodes periodically send messages to a sink node through a multi hop transmission. The basic idea is to burst the messages sent by the source nodes so that a reduced number of bits are transmitted by each forwarder node. In order to better understand the main idea, the example in Figure 1. Nodes A and B have to forward a packet to the sink S and can do it through nodes X, Y, and Z, which are all in the coverage range of A and B. If a normal forwarding scheme is adopted, two cases, Case a) A and B select different next-hop nodes. This happens with probability  $2/3$ . Case b) A and B select the same next-hop node. This happens with probability  $1/3$ .

If there are  $\omega$  bits for each packet, the maximum numbers of bits transmitted by a node belonging to the set  $\{X, Y, Z\}$  is  $w$  bits in the case a) and  $2\omega$  bits in the case b). Let now assume that each node in the set  $\{X, Y, Z\}$  knows that A and B have three possible next-hops and that a different forwarding scheme is adopted, as shown in Figure 1. In particular, when X, Y, and Z receive a packet, they split it and send to the sink only a part. In this case, X, Y, and Z have to transmit at most bits each. While comparing the two forwarding methods, it is concluded that the last one reduces the maximum number of bits transmitted by a node belonging to the set  $\{X, Y, Z\}$ . More precisely, the reduction factor is  $1 - 2/3 = 1/3$ . When comparing the bursting procedure with the procedure shown in case a), and  $(2 - 2/3) \cdot 1/2 = 2/3$  when the bursting procedure is compared to the procedure shown in case b). Summarizing, an average reduction factor of  $4/9$  is obtained.

Accordingly, the lifetime of a sensor network increases as the energy consumption is more distributed among the nodes. It is worth remarking that the bursting procedure has to be performed in a simple manner, and consequently with low energy consumption.

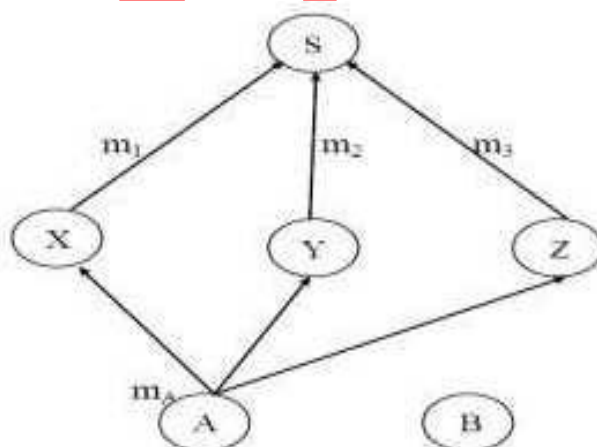


Figure.1. Packet forwarding using different next-hops

The total amount of transmitted bits does not change (bits are transmitted anyway, either with or without splitting); by bursting of packets it is possible to reduce the maximum number of transmitted bits per node and therefore the mean energy that each node consumes for the transmission..

Finally, it can be observed that if a perfect balancing is possible, which occurs when the number of next-hop nodes is a factor of the number of transmitted messages (i.e., the number of messages is exactly divisible by the number of next-hops), the energy consumed by nodes will be the same either with or without bursting. However, if this is not the case, using a bursting technique makes the number of forwarded bits significantly reduced. Moreover, the reduction increases if the ratio “message length over number of components” decreases.

### III. ENERGY-EFFICIENT SELF-ORGANIZED CLUSTERING WITH SPLITTING AND MERGING

We propose EECSM for a WSN. To achieve energy efficiency, EECSM has the following five considerations, using the local information. First, sensor nodes that have the most remaining energy from neighbours become the candidates for the CHs, since the CH consumes a lot of energy. Second, each sensor node, except CHs, selects the nearest CH, in order to minimize energy consumption during the data transmission phase. Third, a cluster that has a number of CMs that is less than the merging threshold merges into other clusters, in order to minimize energy consumption of transmitting the data packets from the CHs to the BS. Fourth, a cluster that has a number of CMs that is larger than the splitting threshold splits into two clusters, in order to reduce the overhead of the CH. Fifth, to prevent any breakdown of CHs, a CH-backup mechanism selects a new CH that has maximum energy in the cluster.

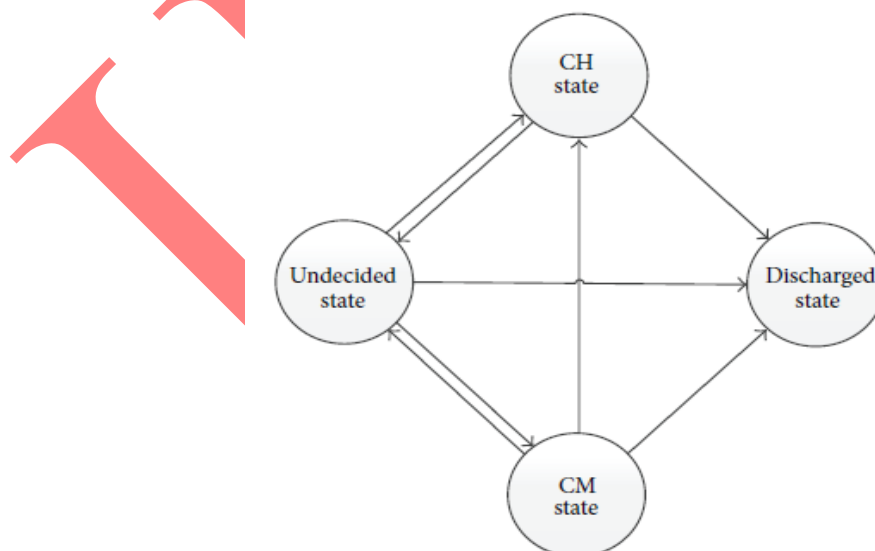
EECSM is comprised of three phases and a CH-backup mechanism; the *clustering phase* forms the clusters for the WSN; the *merging cluster phase* decides whether to merge clusters or not; the *data transmission phase* sends/receives the data packets; and the *CH backup mechanism* elects a new CH when the CH is a breakdown.

**1. States of Sensor Nodes.** Each sensor node of EECSM performs its roles while being transformed into the following four states (*undecided state*, *CH state*, *CM state*, and *discharged state*); if a sensor node becomes a node with the discharged state, it does not change its state.

**2. Clustering Phase.** The clustering phase commences when the sensor nodes are first scattered in the sensor field or after the completion of the “data transmission phase.” This phase decides new CHs to form new clusters for the WSN. To achieve energy efficiency, the criterion in the selection of CH is the remaining energy of CMs. To reduce the load (or overhead) of CHs, the EECSM regulates the size of the clusters. The clustering phase is comprised of four steps: *broadcasting step*, *splitting step*, *CH selection step*, and *clustering step*, as follows.

**A. Broadcasting Step.** When the sensor nodes are scattered in the sensor field at period zero, there is no CH. Thus, all sensor nodes must decide on the CHs only on their own. The EECSM always uses the remaining energy, in order to decide the next CHs, except in period zero. Because the initial energies of all sensor nodes are identical in our assumption, all sensor nodes become CHs in period zero. To avoid this wasteful situation, EECSM uses another CH selection method just once at period zero. We define the *broadcasting range*, before the explanation of the other CH selection method. The broadcasting range is the reachable range of a data packet transmitted from a sensor node. The CH selection method at period zero uses the number of neighbourhood sensor nodes within the broadcasting range, in order to decide who will be the CHs for the first round. In selecting a sensor node having the most neighbours within the broadcasting range as a CH, distances between the CH and CMs are reduced. This reduces the transmission energy between the CH and its CMs. Therefore, in order to decide CHs, we use the number of neighbourhood sensor nodes within the broadcasting range just once at period zero.

**B. Splitting Cluster Step.** After each cluster configuration, EECSM considers the suitable number of clusters. First, EECSM tries to split a large cluster into two small clusters. The disadvantage of a large cluster is the high energy consumption of its CH, when the data packets are transmitted from its CM to its CH via a relatively long transmission path. When receiving the data packets, the more CMs the cluster has, the more its CH consumes energy. A processing bottleneck can also occur at the CH in a large cluster. After the data transmission phase of the previous clustering round, in order to decide the number of clusters in the next clustering round, EECSM selects the next step, according to the value of the splitting threshold.



Transition diagram of sensor node states.



If the number of CMs of a cluster is more than the splitting threshold value, the EECSM executes the splitting cluster step. Since a cluster is split into two clusters at the splitting cluster step, in the next round, two CHs for the split cluster are identified as the First CH and the Second CH, respectively.

**C. CH Selection Step.** If the number of CMs of a cluster is less than the splitting threshold value, the EECSM executes not the splitting cluster step, but the CH selection step. In this step, the CH decides only one CH for the next clustering round.

**D. Clustering Step.** Each CH broadcasts a CH-signal packet to the entire sensor field for clustering. The sensor nodes, except for CHs, or nodes with discharged state receive, store, and list the CH-signal packets. Due to the fact that information can be used at the re-clustering step of the merging cluster phase, the EECSM stores and lists the information temporally. This can reduce overhead of re-clustering. To minimize energy consumption of CMs during the data transmission phase, the sensor nodes select the nearest CH as its CH.

**3. Merging Cluster Phase.** After the clustering phase of the current clustering round, the EECSM checks the size of the clusters. Since the EECSM checks and splits the large clusters at the previous phase, it needs to check only the small clusters at this phase. The disadvantage of a small cluster is inefficient energy consumption, particularly when the data packets are transmitted from the CH to the BS, since energy consumption from a CH to the BS is much more than the energy consumption from a CM to the BS for longer transmission. For this reason, energy consumption of a small cluster is not energy efficient, and therefore the small clusters need to be merged into large clusters. The merging cluster phase on EECSM is comprised of two steps: *merging cluster step* and *re-clustering step*.

**4. Data Transmission Phase.** If the merging cluster phase is finished, clustering is complete. The EECSM enters into the data transmission phase, as EECSM starts to inform the situation of the sensor field to the external BS, by sending gathered data. The data transmission phase is divided into 2 steps: CM transmission step and CH transmission step.

**5. CH Backup Mechanism.** WSNs are useful in inaccessible and dangerous areas, such as military targets, disaster regions, and hazardous environments: they protect humans from danger. Using sensor nodes in those areas may cause a breakdown of individual sensor nodes. This can reduce the network lifetime. Furthermore, the breakdown of CHs can affect the network lifetime, as well as entail a loss of information. If the CHs break or fail, information received from its CMs to the BS may be lost. In order to reduce the degradation of the network lifetime, EECSM has a "CH backup mechanism." If the CH breaks or fails, the CMs that are close to their CH can recognize the breakdown of their CH. The reason why the CMs can realize the breakdown of their CH is that the CMs can recognize whether the data packet is transmitted from their CH to the BS. The advantage is that there is no additional overhead for recognizing the state of the CH. The CH-backup mechanism of EECSM decides a new

CH, after the breakdown of the CH. The CH backup mechanism is comprised of the following 2 steps: *CH reelection step* and *cluster recovery step*.

### **ENERGY REDUCTION FACTOR**

It is important to observe that the set of prime numbers with can be arbitrarily chosen provided that. Therefore, the number of bits needed to represent can be reduced by choosing the prime numbers as small as possible. As a consequence of this choice, the MERF is maximized. Throughout, it indicates with Minimum Primes Set (MPS) the set of the smallest consecutive primes that satisfy the condition. For instance, if and is a 40-bit word, the MPS will be (this is the set of smallest four consecutive primes that satisfies the condition). The MERF in this case is 0.725. However, when the primes set are chosen as above, the message can be reconstructed if and only if all the burst components are correctly received by the sink. Let's consider another primes set.

The choice it is possible to reconstruct the original message even if component is lost (i.e., if we have one failure). In fact, whatever the lost component is, the product of the primes associated with the received components satisfies the condition, and therefore it respects the hypothesis of the burst theorem. For instance, if the last component is not received, it is again possible to obtain as, where is the product of the first three primes, and the first three burst coefficients computed for the MPS-1.

1. The number of components is not changed (i.e., the same number of forwarders is needed).
2. The MERF obtained with the new set is 0.65, i.e., MERF is reduced by about 11%.

Because usually sensor nodes have simple processing units, it is mandatory to have a low complex procedure for obtaining values. This goal can be easily reached if is fixed or takes only a few values. This is a very fast procedure.

## **IV. SYSTEM IMPLEMENTATION**

### **A. FORWARDING ALGORITHM**

The forwarding algorithm is based on two temporal phases, the Initialization phase and the Forwarding phase. The initialization phase is to split data and the forwarding phase performs forwarding.

### **B. INITIALIZATION PHASE**

This phase organizes the network in clusters and also has the advantage of minimizing the number of hops needed to reach the sink. The Initialization phase has been described in detail and it is realized through an exchange of initialization messages (IMs) starting from the sink that is supposed to belong to the cluster. Retransmit the IM. On the basis of the received IMs, at the end of the procedure each node in the network will know its own next-hops, which other nodes will use it as a next-hop, and into how many parts the received packets.

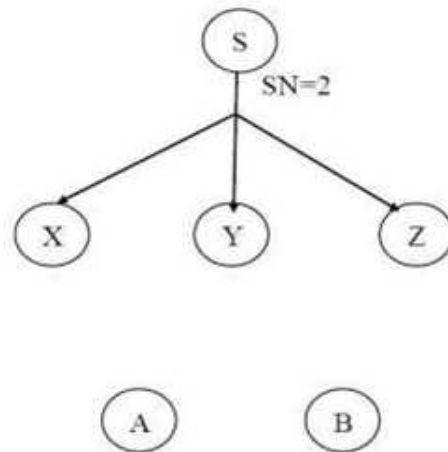
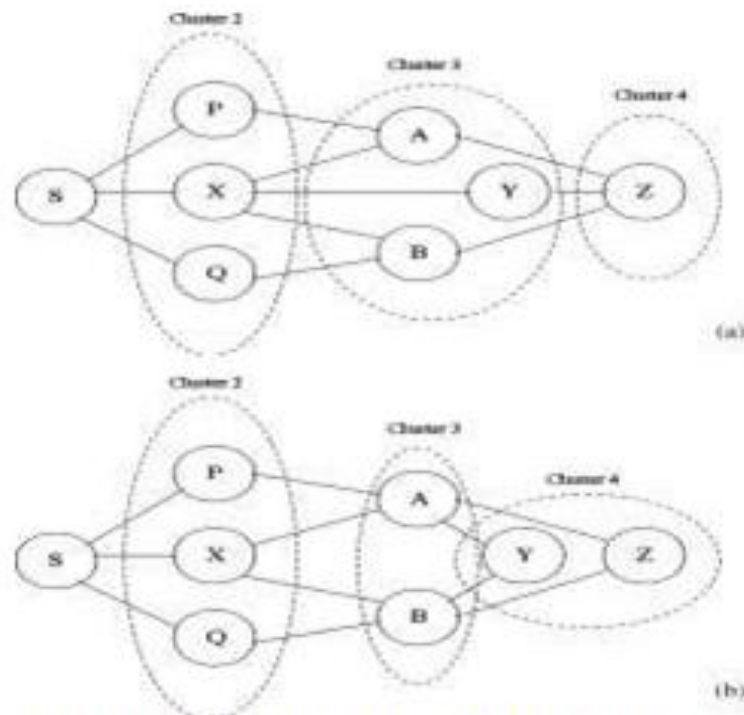


Figure.2. Sink sends the first IM

The sink sends the first IM for the initialization phase each node can obtain the MPS and select a different prime number of the OMPS-f by considering the order of the addresses specified in the IM. Recall that in order to obtain the MPS- f, it is sufficient to know to be either fixed or specified within the IMs, while the number of primes corresponds to the number of possible next-hops (that each forwarding node knows on the basis of the received IMs) Because could be different for each source node , it use the notation .

However, the sink, in order to reconstruct the messages, also needs to know the index of the received component (i.e., for each). For this purpose, it will assume that in the header of each packet there is a field called mask. The mask could be the binary representation of the index followed by the number of components [i.e., pair] or a “one-hot” coding bit sequence followed by a tail bit .Its assumed that the overhead introduced by the mask is negligible. According to the previous initialization procedure shown in Figure .It will receive the IM with from the node X, and it decides to belong to, node Y will have only one next-hop (i.e., X) because Y is at the end of the coverage range of nodes belonging to. Now, we consider what happens with the modified procedure.





**Figure.3. Clustering of burst packets**

The proposed initialization procedure can be further refined in order to increase in some cases the number of possible next hops that a node can use as forwarders. In particular, when a node receives very few IMs with, it does not choose immediately to belong to the cluster, but it waits for the IMs with the next sequence number in order to belong to the last cluster if it is more convenient.

When node Y receives the IM with, it postpones its decision to belong to other node. After some time, it will receive two new IMs with from nodes A and B, and therefore it decides to belong to in order to have two possible next-hops instead of one as shown in Figure 3. Basically, that anode can postpone its decision to belong to a particular cluster if the number of IMs received is less than a chosen threshold, but this can be done just one time in order to avoid an increase of the number of hops needed to reach the sink. The threshold can be a constant value (either specified in the IM or pre stored in the node memory, and therefore already known by the node). Moreover, if the number of received IM is less than the threshold, it is possible to use the conventional shortest path approach (SP) that keeps working also with our technique.

The initialization procedure is performed only when the network is activated for the first time, and it is not necessary to run it when either a new node joins the network or a node runs out of energy. In both cases, it is sufficient that few IMs are exchanged between the node and its neighbours belonging to the near clusters. More details about the operations needed for the above cases have been described. Moreover, in order to consider the unreliability of the channel, which causes loss of IM packets and consequently nodes with an insufficient number of neighbours, each node can start periodically a new joining procedure.

### C. FORWARDING PHASE

Once the network has been organized, the Forwarding phase is applied. Basically; all nodes follow the same forwarding rule: If there is a number of neighbours at least equal to, and the packet has not previously split, then split the packet; else use conventional shortest path approach. Let's consider the network shown, where clusters are obtained according to the initialization procedure already described in the previous section. The messages sent by each node when the source node H sends a message to the sink S. According to the initialization procedure, node G knows that it is the only next-hop of node H, and therefore it must forward the packet without performing a splitting procedure. It is worth highlighting that it is not necessary for G to specify the list of the destination addresses in the packet. In fact, in the initialization phase, nodes have already received the IM message, and therefore they know that node G has four next-hops and that all of them have to split the messages received from G into  $N_G=4$  parts.

Therefore, when they receive the packet, according to both the packet size and they independently select the prime numbers 3 and send the components, together with a proper mask, to one of the possible next-hops. When the sink receives a component, it identifies the number of expected components on the basis of the mask, and therefore it calculates the MPS-f and the coefficients needed to reconstruct the original message.

Concerning the complexity of the algorithm, it is worth mentioning that the message splitting is performed only one time by the nodes that are the closest to the source and have the opportunity to do it (e.g., if they are in proximity of a number of neighbours higher than the threshold specified for the initialization phase), whereas the other sensor nodes in the network will just forward the sub packets. Moreover, only the sink node will reconstruct the original message through more complex operations as described, but this can be consider that usually the sink node is computationally and energetically more equipped than the other sensor nodes. Obviously, in the case of very large packets, it is possible to split the packets recursively, but in order to keep the complexity of the proposed algorithm very low ,its consider that a packet can be split only one time.

## V. PERFORMANCE EVALUATION

In this section, the comparison of the performance of EECSM in terms of energy consumption to those obtained by SP. Moreover, provided some results obtained comparing the EECSM to the most naive splitting scheme, a simple packet division into chunks.

The results have been obtained through a custom MATLAB simulator. A comparison between the results obtained through the analysis and those obtained through the simulator. Then, we analyze some other parameters in order to show the advantages of the proposed technique.

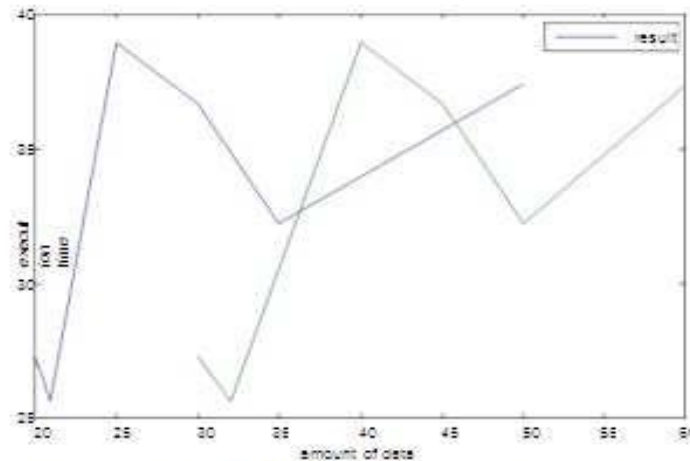


Figure.4. Amount of Data vs Execution Time

Figure 5 shows the values of transmission 1 and transmission 2 are taken by simulating the packet forwarding. Let consider a sensor network where nodes are randomly distributed in a square area of size  $m$ , with density nodes/ $m$ . Sensor nodes are assumed to be static as usual in most application.

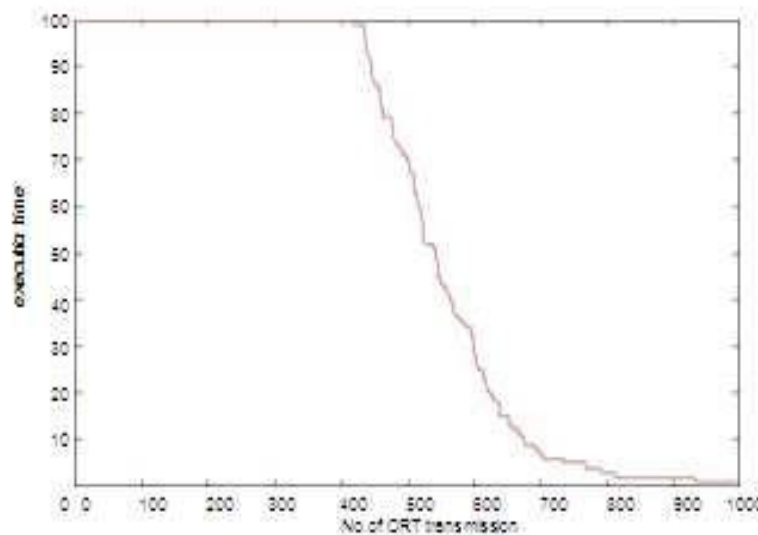


Figure.5. No of Transmission vs Execution Time

Figure.5, shows the values of transmission 1 and transmission 2 are taken by simulating the packet forwarding. In each simulation, the sink node is located in the center of the square grid, and each sensor node has a transmission range equal to  $m$ .

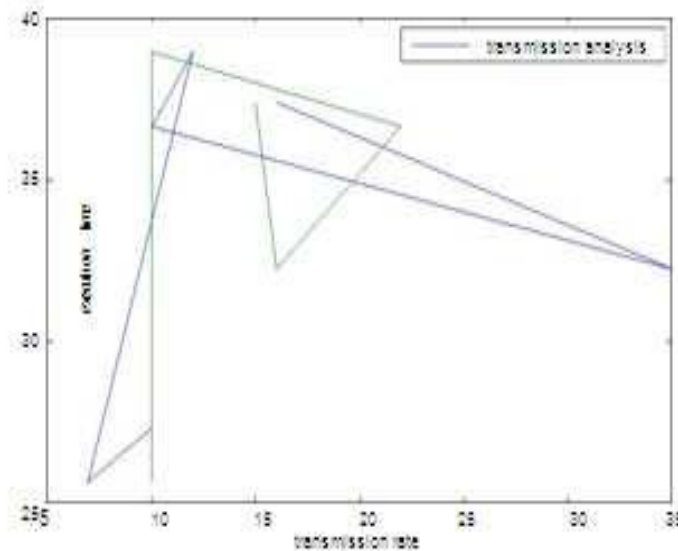


Figure.6 Transmission Rate vs Execution Time

Figure.6 shows the evaluation between transmission rate and execution time. The network is organized in clusters numbered in ascending order starting from the cluster where is located the sink node, which is identified with the events randomly occur in faraway clusters. Simulations neglect the effect of collisions and retransmissions at the MAC layer.

## VI.CONCLUSION

A novel forwarding technique for wireless sensor networks based on the EECSM algorithm is able to predict the energy efficiency of the process. The choice of the clustering algorithm parameters in order to keep the processing complexity low, and then the trade off between energy consumption and reliability are measurable. Finally, the overhead introduced in terms of packet header size in MAC layer will be reduced. Simulation results have confirmed the results obtained and have shown that applying the EECSM-based technique significantly reduces the energy consumed for each burst node, reduces computational complexity and consequently increases the network lifetime.

## REFERENCES

- [1] J. Yick, B. Mukherjee, and D.Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52,no. 12,pp. 2292–2330, 2008.
- [2] S.Mahfoudh and P.Minet, "Survey of energy efficient strategies in wireless ad hoc and sensor networks," in *Proceedings of the 7th International Conference on Networking (ICN '08)*, pp. 1–7, April 2008.
- [3] J.N.Al-Karaki andA. E.Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, 2004.

- [4] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005.
- [5] N. Israr and I. U. Awan, "Multilayer cluster based energy efficient routing protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 4, no. 2, pp. 176–193, 2008.
- [6] M. Younis, M. Youssef, and K. Arisha, "Energy-aware management for cluster-based sensor networks," *Computer Networks*, vol. 43, no. 5, pp. 649–668, 2003.
- [7] F. Dressler, *Self-Organization in Sensor and Actor Networks*, JohnWiley & Sons, New York, NY, USA, 2007.
- [8] C. Prehofer and C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 78–85, 2005.
- [9] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [10] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "How to prolong the lifetime of wireless sensor network," in *Handbook of Mobile Ad Hoc and Pervasive Communications*. Valencia, CA: American Scientific Publishers, 2007, ch. 6.
- [11] T. R. Burchfield, S. Venkatesan, and D. Weiner, "Maximizing throughput in ZigBee wireless networks through analysis, simulations and implementations," in *Proc. Int. Workshop Localized Algor. Protocols WSNs*, Santa Fe, NM, Jun. 2007, pp. 15–29.
- [12] G. Campobello, A. Leonardi, and S. Palazzo, "On the use of Chinese Remainder Theorem for energy saving in wireless sensor networks," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 2723–2727.
- [13] G. Campobello, A. Leonardi, and S. Palazzo, "A novel reliable and energy-saving forwarding technique for wireless sensor networks," in *Proc. ACM MobiHoc*, New Orleans, LA, May 18–21, 2009, pp. 269–278.
- [14] R. Crepaldi, A. F. Harris, III, M. Rossi, G. Zanca, and M. Zorzi, "Fountain reprogramming protocol: A reliable data dissemination scheme for wireless sensor networks using fountain codes, demo abstract," in *Proc. ACM SenSys*, Sydney, Australia, Nov. 2007, pp. 389–390.
- [15] B. Deb, S. Bhatnagar, and B. Nath, "ReInForM: Reliable information forwarding using multiple paths in sensor networks," in *Proc. 28<sup>th</sup> Annu. IEEE LCN*, Bonn, Germany, Oct. 2003, pp. 406–415.