Cluster Head Enhanced Election Type-2 Fuzzy Algorithm for Wireless Sensor Networks

J. C. Cuevas-Martinez, A. J. Yuste-Delgado, and A. Triviño-Cabrera

***Abstract*— This approach presents a fully distributed clustering solution for wireless sensor networks. It relies on the results of an interval type-2 fuzzy logic system that gives each node the chance to be a cluster head. Taking into account the limited computational resources of the sensors, this inference system has been carefully adapted to be run in each node through a sampling process of the entire solution space of the fuzzy system. The input variables of the system are obtained from the information that each node derives from its performance metrics and those related to its neighbors. The acquisition of these last data does not incur in any additional control packets. The results obtained show a significant improvement in the network lifetime when compared with other recent approaches. This improvement takes place even when contrasting with centralized methods.**

***Index Terms*— Wireless sensor network, fuzzy systems, distributed systems, fuzzy control.**

I. Introduction

WIRELESS sensor networks (WSNs) constitute an active research topic due to the still open techniques to be used in these networks and their wide variety of applications [1]: military, surveillance, disaster control, health care, home automation, etc. All those applications of WSNs result from their versatility, which is closely related to the main characteristics of the nodes: small isolated processing units with wireless communication capabilities and a limited power supply. Therefore, nodes usually work under strict energy constraints. Moreover, most applications rely on random node deployments because those sensors are deployed in the wilderness or environments where infrastructures for power supply are not feasible [2]. Despite these limitations, the main purpose of most nodes in the WSN is to monitor the surrounding environment and to send the measurements to a sink or base station to which is not usually connected directly [3]. As a consequence, techniques like sleep scheduling (i.e. setting idle periods) and power control (i.e. tuning the transmission power) are common remedies in WSNs to prolong the network lifetime [4]. Network lifetime refers to the period of time during which the nodes can accomplish their duties, mainly, sensing some parameters, transferring the information and retransmitting others' data. This period is usually limited by the batteries of the nodes. Thus, the techniques to be applied in WSN have to find a balance between extending the network lifetime and ensuring precision, that is, the number of measurements done and transmitted over time and across space. In those terms, and in order to reduce the number of transmissions, routing techniques, like clustering, have been employed over the last years with a varying level of success. Clustering is based on setting some nodes as cluster heads (CHs) and making them aggregate the data collected from their neighbors. However, the nodes acting as CHs may suffer from early battery depletion as they have to send a lot more messages than any other non-CH nodes. Dynamic clustering techniques are used to solve this problem by changing the nodes that act as CHs in every round. According to this approach, Low-Energy Adaptive Clustering Hierarchy (LEACH) [5] is one of the most relevant protocols. It sets the election of the CH on a simple random process in every round. This method has been proven to work fairly well for randomly deployed networks based on homogeneous nodes. Consequently, and also due to its simplicity, nowadays it is kept as a reference for most approaches to dynamic cluster formation in WSN [6]. However, one of the problems arising from that kind of pure stochastic method is that there is a chance of not selecting any node in one round which eventually could lead to data losses. Moreover, in LEACH several important parameters are not taken into account in the CH election process. For instance, this protocol does not consider the residual energy in the CH or how many surrounding nodes the candidate CH should potentially serve. The relevance of these issues has only been investigated recently [7]. However, most of the proposed algorithms are either centralized [8], [9] or usually rely on additional precise location information, which needs to be provided by a built-in GPS. Thus, those approaches call for extra energy to deal with these required additional control data and the location hardware, which also raises the sensor price. In order to avoid the additional control data and the hardware required by those solutions, we present a new fully distributed, interval type-2 fuzzy logic controller (IT2 FLC) based clustering algorithm that relies on the information that each sensor is able to gather by itself, avoiding any extra control packets, which helps to save energy for the power-constrained nodes. Fuzzy logic controllers are widely employed in engineering to characterize a system whose inputs and outputs have extremely complex relationships, making the definition of any mathematical function between them impossible. Furthermore, due to the great uncertainty involving the use of only local information, a type-2 fuzzy logic approach is preferred. In particular, each node infers the probability of being a CH based on the output interval of a carefully designed IT2 FLC and on the past behavior of the node.

Manuscript received March 29, 2017; revised April 29, 2017; accepted May 4, 2017. Date of publication May 12, 2017; date of current version September 8, 2017. The associate editor coordinating the review of this paper and approving it for publication was M. A. Imran. *(Corresponding author: A. J. Yuste-Delgado.)*

J. C. Cuevas-Martinez and A. J. Yuste-Delgado are with the Department of Telecommunication Engineering, Universidad de Jaén, 23700 Linares, Spain.

A. Triviño-Cabrera is with the Department of Electrical Engineering, Universidad de Málaga, 29071 Málaga, Spain (e-mail: ajyuste@ujaen.es).

Digital Object Identifier 10.1109/LCOMM.2017.2703905

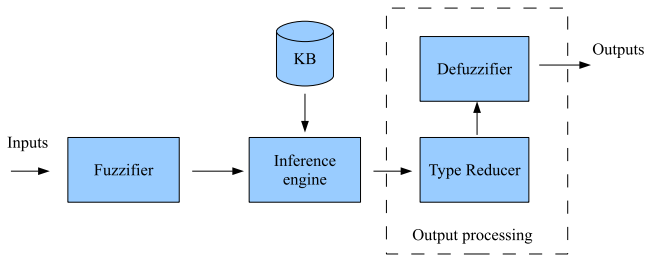

Fig. 1. Interval type-2 fuzzy logic controller.

II. Methodology

The presented cluster head enhanced election type-2 fuzzy algorithm (CHEETAH) dynamically elects the CH in a WSN following a distributed approach that is based on an interval type-2 fuzzy logic controller. IT2 FLCs link inputs and output through several stages as can be seen in Fig. 1. The main difference in the architecture shown in Fig. 1 between type-1 and type-2 fuzzy controllers lies in the "output processing" block. The other blocks are (i) a fuzzifier, which normalizes input values from a crisp value, which in type-2 fuzzy logic implies a more complex processing; (ii) the knowledge base (KB), which defines the relationships between inputs and output with IF THEN clauses; (iii) the inference engine which binds the inputs and the knowledge base to infer a solution; (iv) the output processing block which produces a single value for type-1 fuzzy logic controllers and an interval for type-2 fuzzy logic controllers (in this last case, a new block, "type reducer", is used to obtain the output) and (v) the defuzzifier, which gives a crisp value through denormalization.

The reasons for using an interval type-2 fuzzy logic [10] can be summarized as follows [11]: (i) it is easier to design fuzzy rules from expert knowledge and from natural language because the membership grade of an IT2 is an interval instead of a crisp number, which increases the robustness of the system, and (ii) IT2 FLCs are more adaptive than type-1 fuzzy systems due to the complex relationships between their inputs and outputs, which is of paramount importance in systems with high levels of uncertainty like WSNs.

A. Input Variables

The input variables of the interval-based type-2 fuzzy system are the following:

- residual energy (R) of the node, measured as a percentage of the maximum energy of the battery.
- relative distance (D) to the base station (BS) estimated from the received signal strength indication (RSSI) of the messages that come from the BS;
- historical contribution as a CH (H), i.e., the mean number of nodes attended by that node when acting as CH in the past per round; and
- efficiency (E), which stands for the estimated power saved by a group of nodes when they relay their data to the CH instead of sending them directly to the BS.

The output interval of the IT2 FLC (o_l, o_h) represents the minimum and the maximum chance for a node of being elected as a CH, respectively. As can be noted, variables D and E involve a high degree of uncertainty because they cannot be obtained simply by measuring a parameter within the sensor. Alternative methods are proposed in this letter for their estimation. Variable D is obtained from the division of the distance to the BS (estimated from the messages RSSIs) into the distance to the BS from the furthest area of the deployment field of the WSN. The maximum value considered for D is 1 so that normalized inputs can be used in the fuzzy system. Historical Contribution (H) gives information about the goodness of a node acting as a CH. Therefore, if a node serves a number of nodes close to or even greater than the optimum number of CH for each round k obtained in LEACH [5], the node is considered as a suitable candidate to be a CH. The exact computation is given in Eq. 1.

$$H = \frac{\sum_{i=0}^{r} h_i}{r * k * n} \tag{1}$$

where:

- r is the total number of times that the node was a CH.
- h_i is the number of nodes that chose that node as their CH in round i.
- $k = 0.05$ which is the optimum factor established by LEACH.
- n is the total number of nodes in the WSN.

The maximum value of H, in order to be normalized for the fuzzifier, is $k * n * 1.5$.

The value of $efficiency$ is updated each time a node acts as a CH in the following way:

$$E = 1 - \min\left(\frac{\sum_{i=1}^{h} e_i}{\sum_{i=1}^{h} e_i^{'}}, 1\right) \tag{2}$$

where:

- h is the number of nodes that chose that node as their CH in one round.
- e_i = energy spent by a node to send data to the CH.
- $e_i^{'}$ = estimated energy spent if the same data sent to the CH were sent to the base station directly.

Therefore, the lower the efficiency, the worse a node acts as a CH in terms of power saving. Variable E can only be estimated on the basis of the RSSI received from each node. Each node transmits at maximum power the first time it sends data to a new CH to allow the computation of efficiency. All the variables are set to 0 in the first round. Variable R is updated in every round, whereas the other three are computed by a node once it has been selected as a CH. The fuzzy set layout of all the input variables is similar for all of them. Each input variable has three fuzzy sets —low, medium and high— with a trapezoidal shape as depicted in Figure 2. Thus, for a variable, the width of the membership function reflects the degree of uncertainty of that variable. Therefore, the wider the area, the higher the uncertainty of the variable, so that its contribution could avoid the effects of noisy inputs (i.e. the width for variable R is narrower than that of D because R is directly measured by the node whereas D is estimated from the RSSI measurements). For the output interval, two more fuzzy sets have been employed, very low and very high.

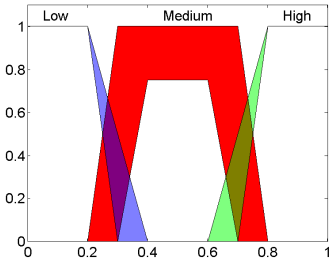

Fig. 2. Fuzzy set layout for variable R.

The knowledge base has 81 different rules. An extract of them is presented as follows: **if** R is low and D is high and H is low and E is low **then** $output$ is low. Therefore, the IT2 FLC of each node will infer a low probability of being selected as a CH if the remaining energy on the node is low (R low), it is far from the BS (D high), its historical behavior as a CH is not significant because it was selected by few nodes (H low) and the energy saved by nodes that elect it as a CH is not significant compared with the energy used by sending the data directly to the BS (E low).

However, if a node has an almost-full battery (R high), it is close to the BS (D low), in previous rounds it was selected as a CH by an important number of nodes (H high), and a significant amount of energy was saved by those nodes (E high), the probability of being elected as a CH will be inferred as high as it is stated in the following rule: **if** R is high and D is low and H is high and E is high **then** $output$ is high.

B. Cluster Head Election Algorithm

CHEETAH is ruled by three phases (similar to LEACH): CH selection, CH advertisement and data transmission. These phases are explained below:

- CH selection: each node individually computes its own probability P based on the four different variables that are the inputs of the type-2 fuzzy system.
- CH advertisement: Then each node is randomly assigned a value ranging from 0 to 1. If this value is lower than or equal to P, then the node sends a CH advertisement message (ADV) to the network at the maximum transmission power. Thus, each non-CH node selects as its CH the node whose ADV is received with the highest RSSI.
- Data transmission: each node adjusts the transmission power to reach the selected CH and sends the information to it. The CH then aggregates the received data from all its contributing nodes and forwards it to the sink.

The calculus of P for each node in a round i, is based on the result of the interval-based type-2 fuzzy system as shown in equation (3).

$$P(i) = \begin{cases} o_h & \text{if ch = 0} \\ o_l & \text{if ch = 1} \end{cases} \tag{3}$$

where:

- o_h is upper limit of the interval of the type-2 fuzzy system
- o_l is lower limit of the interval of the type-2 fuzzy system.
- ch is set to one when a node has already been CH once for that epoch, 0 otherwise.

To allow an effective and real-time computation of the IT2 FLC to infer the probability of being a CH, without the need of high-performance microprocessors, the input variables and the solution space have been sampled so that any complex process in the sensor node is avoided. Each input is represented by 10 values, which gives a four dimension matrix of 10^4 possible solutions for the IT2 FLC that need around 480 Kbytes for double precision float values, which can be easily stored in a node. Although sampling reduces precision, it can be assumed due to the uncertainty of the input variables, and the intrinsic characteristics of the IT2 FLC, like the interval definition for fuzzy sets. Hence, when a node has to obtain its probability (P) of being a CH, it has to check the right entrance of a matrix stored in its memory with the quantified values of R, D, H and E. Consequently, it gets the interval output and decides which (lower or upper) limit is set to P according to its past behavior. The node then decides whether it has to act as a CH when the result of a random function is lower than or equal to P. In addition, if a CH does not receive any message from other nodes to join its cluster, it converts itself into a contributing node of the closest CH. The proposed method has been tested under several network configurations composed of a different number of nodes. The results, which are promising, are explained in further detail in the next section.

III. Results

CHEETAH has been implemented in Matlab in order to compare it with two recent approaches that also employ computational intelligence techniques and have been demonstrated to outperform LEACH among other solutions. For both test scenarios, we have used symmetric transmission channels to allow an effective RSSI estimation and the first-order radio model [5]. Therefore, the energy expended to transmit an n-bit message a distance d appears in (4) whereas (5) shows the energy expended to receive a message.

$$\begin{aligned} E_{Tx}(n,d) &= E_{Tx-elec}(n) + E_{Tx-amp}(n,d) \\ &= \begin{cases} nE_{elec} + l\epsilon_{fs}d^2 & d < d_0 \\ nE_{elec} + l\epsilon_{mp}d^4 & d \geq d_0 \end{cases} \end{aligned} \tag{4}$$

$$E_{Rx}(n,d) = E_{Rx-elec}(n) = nE_{elec} \tag{5}$$

where:

- $E_{elec} = 50\ nJ/bit$, is the energy dissipated by the electronics in the transmitter.
- $\epsilon_{fs} = 10\ pJ/bit/m^2$, is the energy amplifier per bit for free space propagation model.
- $\epsilon_{mp} = 0.0013\ pJ/bit/m^4$ is the energy amplifier per bit for multi-path propagation model per bit.
- $d_0 = \epsilon_{fs} / \epsilon_{mp}$ meters.

Advertisement messages have a length of 200 bits whereas data packet length is set to 4000 bits and the energy expended in data aggregation is set as $E_{DA} = 5\ nJ/bit/signal$. It is supposed that every node can send data to any other node

TABLE I
SIMULATION RESULTS FOR THE EXPERIMENTS IN TS-1

	LEACH	T2FLCA	CHEETAH
FND	689	1012	1099
HNA	813	1074	1502
LND	881	1133	2055

TABLE II
SIMULATION RESULTS FOR THE EXPERIMENTS IN TS-2

	LEACH	DUCF	CHEETAH
FND	416	715	1564
HNA	671	767	2943
LND	791	988	3758

within the network without needing routing. For each test scenario, there are three performance metrics:

- the round when the first node depletes its battery, known as first node dies (FND),
- the number of rounds when 50% of the nodes are dead, called half of the nodes alive (HNA), and
- the round when the last node dies (LND).

The results of our proposal for both the experiments are the mean of 30 different simulations.

A. First Scenario

In the first test scenario (TS-1), we compare our proposal with T2FLCA [12]. This approach is a centralized type-2 fuzzy clustering system based on ant colony optimization (ACO). The ACO subsystem relies on some control packets of 200 bits and 100 bits in length. The experiments have been developed over a fixed network of 200 m $\times$ 200 m where 200 nodes were randomly deployed. The BS is located at the center of the sensor field (100 m, 100 m), which corresponds to the configuration where T2FLCA has its best performance. The initial energy of the nodes is set to 0.5 J. The mean results for FND, HNA, and LND obtained in TS-1 are detailed in Table I. As can be noted, T2FLCA gets a better performance than LEACH. Nevertheless, our proposal clearly improves the results obtained using these two solutions. This improvement is achieved by the convenient selection of the CH and the reduced amount of control packets required by our proposal.

B. Second Scenario

The second test scenario (TS-2) includes the comparison with a distributed load balancing unequal clustering in wireless sensor networks using the fuzzy approach (DUCF) described in [13]. This method employs three input variables for CH election: (i) residual energy of a node, (ii) node degree: number of nodes within the transmission radius, and (iii) distance to BS. DUCF also needs short control packets to work (25 bytes). The experiments have been developed over a fixed network of 200 m $\times$ 200 m where 100 nodes were randomly deployed. The BS is located at the center of the sensor field (100 m, 100 m). The initial energy of the nodes is set to 1 J. Table II shows the results obtained for TS-2. As in the previous scenario, the values obtained for our proposal show that it can significantly extend the network lifetime by eluding the use of control packets to select the CH.

IV. CONCLUSION

This letter has presented an interval type-2 fuzzy logic control to perform a distributed clustering in WSN so that the network can operate for a longer time on limited batteries. The algorithm counts on a sampling process to adequate its function to computational constrained devices. It also eludes the transmission of any additional control packets required to define the control inputs so that the network lifetime is maximized. As evident from Table I and Table II, the results obtained by the presented approach significantly outperform the solutions used for the comparison in two scenarios. This is achieved without transmitting any control information to the BS or any other node in the network. Therefore, our approach can significantly improve the lifetime of the whole network without the need for any complex process in the nodes or any central computation. In future work, we will address multi-hop routing and heterogeneity of the nodes using this approach.

REFERENCES

[1] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*. New York, NY, USA: Wiley, 2010.
[2] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.
[3] A.-F. Liu, P.-H. Zhang, and Z.-G. Chen, "Theoretical analysis of the lifetime and energy hole in cluster based wireless sensor networks," *J. Parallel Distrib. Comput.*, vol. 71, no. 10, pp. 1327–1355, Oct. 2011.
[4] X. Zhang, X. Ding, S. Lu, and G. Chen, "Principles for energy-efficient topology control in wireless sensor networks," in *Proc. 5th Int. Conf. Wireless Commun., Netw. Mobile Comput. (WiCom)*, Sep. 2009, pp. 1–3.
[5] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.
[6] X. Liu, "A survey on clustering routing protocols in wireless sensor networks," *Sensors*, vol. 12, no. 8, pp. 11113–11153, 2012.
[7] S. Tyagi and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 36, no. 2, pp. 623–645, Mar. 2013.
[8] G. Hacioglu, V. F. A. Kand, and E. Sesli, "Multi objective clustering for wireless sensor networks," *Expert Syst. Appl.*, vol. 59, pp. 86–100, Oct. 2016.
[9] F. Zhang, Q.-Y. Zhang, and Z.-M. Sun, "ICT2TSK: An improved clustering algorithm for WSN using a type-2 Takagi–Sugeno–Kang fuzzy logic system," in *Proc. IEEE Symp. Wireless Technol. Appl. (ISWTA)*, Sep. 2013, pp. 153–158.
[10] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Inf. Sci.*, vol. 8, no. 3, pp. 199–249, 1975.
[11] D. Wu, "Two differences between interval type-2 and type-1 fuzzy logic controllers: Adaptiveness and novelty," in *Advances in Type-2 Fuzzy Sets and Systems* (Theory and Applications). New York, NY, USA: Springer, 2013, pp. 33–48.
[12] Q.-Y. Zhang, Z.-M. Sun, and F. Zhang, "A clustering routing protocol for wireless sensor networks based on type-2 fuzzy logic and ACO," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ)*, Jul. 2014, pp. 1060–1067.
[13] B. Baranidharan and B. Santhi, "DUCF: Distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach," *Appl. Soft Comput.*, vol. 40, pp. 495–506, Mar. 2016.