AN EFFICIENT APPROACH TO DETECT FAULTY READINGS USING FUZZY LOGIC AND READ VECTOR ALONG THE SUBSTRATE ACCESS WIRELESS LONG-THIN SENSOR NETWORKS

S. Jalaleddin Dastgheib

Young Researchers Club, Shiraz Branch,
Islamic Azad University, Shiraz, Iran

Email: dastqeib@gmail.com

Abstract

Wireless sensor networks (WSN)s consist of hundreds or thousands tiny nodes that work together and connected to each other to do some special tasks. Detecting nodes with faulty readings is one the important issues in WSNs. The existing algorithms to detect faulty readings using weighted voting and are divided in tow category; existing algorithms that using correlation of two nodes read vectors as weight and algorithms that using inverse of distance as weight. The first category algorisms are costly and second category algorithms have weaknesses in accuracy of calculations. This paper proposes a new fuzzy-based algorithm to detecting faulty readings in WSNs. We propose the new method based on LTN however, it is applied in the most of WSN structure. Using an effective fuzzy inference system can improve the decision-making algorithm, which using for detecting faulty readings in WSNs. We use of entire read vector without any additional cost to the network. The experimental results show that the proposed algorithm imposes very low cost to the network; in addition, the accuracy of the results is improved when compared to the other algorithms.

Keywords: *Wireless sensor networks, Long-Thin, Faulty readings, Fuzzy logic, read vector.*

I. INTRODUCTION

Wireless sensor network (WSN) consists of hundreds of small nodes (sensors) that work together to accomplish a network task. One of the important issues that need to be considered in WSNs is detecting nodes with faulty readings. In WSN, the faulty nodes give wrong reports on environment status, which resulted in incorrect interpreting on the environment data. Furthermore, the environmental noise can affect the accuracy of the reports, too [1]. In this paper, the arbitrary and noisy readings are considered as faulty readings. This is very important to detect and filter the faulty readings in the network, as these failures can be the source of attacks and may prevent the benefit of WSN.

In this paper, a method to achieve fault tolerance in the structure of LT network will be proposed. Long-Thin Network (LTN) is a specific type of network topology that widely used in wireless sensor applications. In the LTN, nodes may form several long backbones, which extend the network to intended coverage areas. A backbone is a linear path, which may contain tens or hundreds of routers. We propose the new method based on LTN however, it is applied in the most of WSN structure.

In this paper, the focus is on detecting faulty readings in WSNs. Clearly, collecting all readings and sending them to the sink is not a complex method and by performing statistical analysis, the sink can determine which readings are away. Furthermore, due to limitations of energy sources in nodes this centralized method will be impractical because packet forwarding directly to the sink will increase node energy consumption.

The fact that readings data from neighboring nodes are similar can be expressed by the spatial correlation. The set of neighboring nodes is called set of witnesses. Therefore, an idea to detect faulty readings is using this correlation space. In other words, if a node (s_j) receives an unusual reading, this node asks neighbors if the reading is faulty or not by sending suspicious reading to them (referring to the set of witnesses). Based on the classic voting of majority:

1. Each node within the set of witnesses (e.g. node s_i) makes a judgment by comparing its reading with an abnormal reading submitted by the suspect node (s_j).
2. If the difference between these two readings exceeds from a predefined threshold, s_i considers the reading posted by s_j as faulty and gives a negative vote to s_j. Otherwise, s_i claims that s_j is normal and a returns positive vote to s_j.
3. After collecting votes from the neighbors, s_j decides if the reading is faulty or not. If the number of negative votes is more than positive votes, reading reports of unusual s_j is known as a faulty reading. Otherwise, this reading is assumed as observed happening [2].

Nevertheless, this simple voting method does not work well when the number of faulty nodes is large. To solve this problem in this research a weighted voting algorithm is proposed [3, 4].

Assuming that closer nodes have more readings similarity, weighted voting algorithms allocate more weight to closer neighbors. (E.g. weights of reverse of distance of a node with its neighbors.)

There are two category of weighted voting method:

- The fist category is base on inverse of distance as weight.
- The second category is base on correlation between sensors readings as weight.

The second batch methods that explain later are very complex and expensive as energy. The first category has some weaknesses such as faulty node in near the voting node can devastating impact on voting. Debraj de proposes a weighed voting algorithm [5] and tries to cover these weaknesses but has not been fully resolved. In this algorithm, weight is composed of inverse distance of two nodes and the confidence number between them. Debraj de uses of confidence number, which is obtained from the localization error detection algorithm, to increase accuracy of voting.

In the second category, weight is considered as correlation of two nodes read vectors. These algorithms have high accuracy but are very costly. We in paper [6] try to reduce the energy consumption of this method however, this problem still exist.

Based on the above observations we propose a novel approach to detect faulty reading using fuzzy logic. Fuzzy logic, which will be explained briefly in section three, is used in the proposed method due to can reduce computational complexity, delay, and energy consumption, improve accuracy and performance. Some of the areas fuzzy logic has been applied to are cluster-head election [7,8], security [9,10], data aggregation [11], routing [12,13], MAC protocols [14], and QoS [15,16]. However, not much work has been done on using fuzzy logic for detecting faulty nodes.

Since the proposed method considers suitable sequence of data to detect faulty readings like the second category has high accuracy, in addition by using fuzzy logic just a few bit transfer in the network, so this method is energy efficient. We refer to the proposed method feature in next sections.

The rest of the paper organized as: LTN is introduced in section 2; A briefly overview is done on fuzzy logic in section 3; Section 4 proposes the new method; At the end of this paper, in parts 5 and 6 comparing and conclusion is done.

II. LONG-THIN NETWORK

The Long-Thin Network (LTN) is a type of network topologies, which widely used in wireless sensor network applications [17-19].

The form of nodes distribution in the Long-Thin Network (LTN) causes each node to have fewer neighbors. Few neighbors will causes having fault in network. Number of neighbors should not be so little that compromise the health of network. Long-Thin structures are usually used in environments that are included in the restrictions. These restrictions limit the number of neighbors. In this structure, failure of some close together nodes may pull some parts of network into isolation, or in a worse case the entire network may stop working. A structure for LTNs is shown below (Figure 1). This structure is an optimal deployment for the sensor nodes within the LTN, and is useful in most practical applications [5].

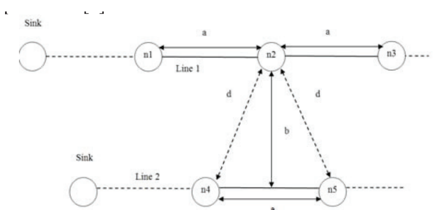

Figure 1. Fault tolerant sub-structure for Long-Thin network

The nodes distribution has been explained in figure 1; this infrastructure is repeated throughout the network. In this infrastructure, the increase in number of lines depends on space limitations and other issues. Number of neighboring nodes is four based on this infrastructure.

III. FUZZY LOGIC

Before discuss about the proposed method to detect faulty readings it is necessary to do an overview on fuzzy logic. Fuzzy Logic (FL) is defined as the logic of human thought, which is much less rigid than the calculations computers generally perform. Fuzzy Logic offers several unique features that make it a particularly good alternative for many control problems. It is inherently robust since it does not require precise, noise-free inputs and can be programmed to fail safely [20, 21].

The output control is a smooth control function despite a wide range of input variations. Since, the FL controller processes user defined rules governing the target control system; it can be modified and tweaked easily to improve system performance. Fuzzy Logic deals with the analysis of information by using fuzzy sets, each of which may represent a linguistic term like “Warm”, “High”, etc. The range of real values over which the set is mapped, called domain and the membership function describes fuzzy sets. A membership function assigns a truth (crisp) value between 0 and 1 to each point in the fuzzy set’s domain. Depending upon the shape of the membership function, various types of fuzzy sets can be used such as triangular, trapezoidal, beta, PI, Gaussian, sigmoid, etc. We use triangular and trapezoidal membership functions due to they are suitable for real-time operation because they do not complexity computations and are having enough accuracy [22,23].

The fuzzified values are processed by the inference engine, which consists of a rule base and various methods for inferring the rules. One of the fuzzy systems that used in the inference engine of the expert system is the Mamdani fuzzy system. The Mamdani fuzzy system is a simple rule-base method that does not require complicated calculations which can employ the IF…THEN… rules to control systems [24]. All the rules in the rule-base are processed in a parallel manner by the fuzzy inference engine. The defuzzifier performs defuzzification on the fuzzy solution space. That is, it finds a single crisp output value from the solution fuzzy space. Some techniques are introduced for deffuzification like Center Of Area (COA), mean of maximum and etc. COA is most suitable technique for WSN so we use this technique for defuzzification [25]. In this paper, the crisp value adopting the COA defuzzification method was obtained by the following formula.

$$CrispOutput(\alpha) = \frac{\int_z \mu_A(x) z dz}{\int_z \mu_A(z) d} \qquad (1)$$

Where α is the crisp value for the "z" output and $\mu A(z)$ is the aggregated output membership function.

IV. PROPOSED METHOD

As mention above, the voting methods to detect faulty readings have some weaknesses. To solve these problems we introduce a novel approach using fuzzy logic. Consider node 2 in development of node in figure 1 as suspected node that wants to know the status of its readings thereupon requests from its neighboring nodes to run their fuzzy systems. Meanwhile node 2 also runs its fuzzy system. The component of fuzzy system, which each node has it, is shown in figure 2.

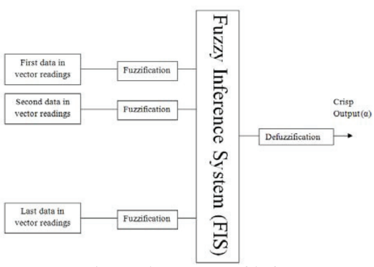

Figure 2. The components of the fuzzy system

Suppose that all readings of a node (e.g. node i) include a sequence of readings inside the sliding window Δt, this sequence readings is called Read Vector. The node i readings can be expressed as follows:

$$b_i(t) = \{x_i(t - \Delta t + 1), x_i(t - \Delta t + 2), ..., x_i(t)\} \qquad (2)$$

That x_i (t) is sensed value by node i in time **t** [2]. We suppose that the length of read vector is eight because it can provide the necessary accuracy and increase of this number is easily possible. The fuzzy membership functions for inputs and output is shown in figure 3.

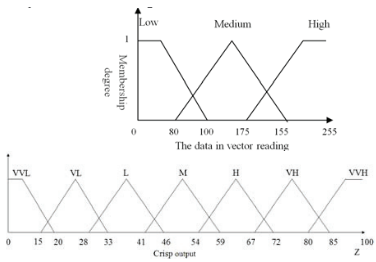

Figure 3. Fuzzy membership functions

As was mentioned is section two, we use of mamdani method to do fuzzy inference. Some of the existing rules, which are used in mamdani inference system, are listed in table I.

Table I. Some of the existing rules

1	2	3	4	5	6	7	8	Output
L	L	L	L	L	L	L	L	VVL
L	L	H	H	L	L	H	H	M
L	H	H	H	H	H	L	M	VH
M	H	L	M	H	L	M	H	M
H	H	H	L	H	H	H	H	VVH
H	M	L	L	L	L	L	L	VVL
H	H	M	M	H	H	L	L	H
L	L	M	M	L	L	H	H	L

Legend: L =Low;M=Medium;H=High;V=Very

When the suspected node (node 2) wants to know its readings is faulty or not, runs its fuzzy system and sends the crisp output to its neighboring nodes. Each of these nodes runs the same fuzzy system then compares the output of this fuzzy system with the suspected node crisp output (α_2) and gives a vote to the suspected node. Each of neighboring nodes (e.g. node j) has a weight to participate in voting which is obtained from following formula. The weight is similarity between the suspected node fuzzy system output and the neighboring node (e.g. node j). We use of Jaccard similarity function [26] to obtain the similarity between α_2 and α_j.

$$Weight_j = \frac{\alpha_j * \alpha_2}{\alpha_j^2 + \alpha_2^2 - \alpha_j * \alpha_2} \quad (3)$$

All neighboring nodes send their vote and weight to node 2 then node 2 does the voting by the following formula:

$$Voting_2 = \sum_j vote_j * weight_j \quad (4)$$

The suspected node (node 2) can determine the status of its readings.

$$Node_2\ Readings = \begin{cases} if\ voting_2 \le 0 & faulty \\ Otherwise & correct \end{cases} \quad (5)$$

As can seen the proposed method using fuzzy logic and read vector athwart the other method imposes very low cost to the network and has very high accuracy.

V. ANALYZING AND COMPARING

This section compares existing methods to detect faulty readings with the proposed method. The existing method generally use of weighted voting to detect faulty readings. The existing weighted voting algorithms to detect faulty reading can be classified into the two categories. The first category is established based on inverse of distance as weight and the second category is established based on the correlation between read vectors.

The algorithms based on distance inverse are very vulnerable to faulty nodes. Debraj De faulty reading detection algorithm belongs to the first category. He tries to overcome this problem but we will show that the problem still exist in this algorithm.

The algorithms based on correlation as weight are very complex and costly. The complexity rate of these algorithms is high and it is overall about $O(n^3)$ [2]. We proposed a method in [6] that reduced the complexity rate of correlation algorithms to an acceptable level. The complexity rate of this algorithm is about $O(n^2)$. Therefore, this algorithm is still costly.

A) Weakness of the first category algorithm (Debraj de)

It is easily proved that our approach is much more accurate than the first category algorithms. Note the figure 3.

It is easily proved that our approach is much more accurate than the first category algorithms. Note the figure 3.

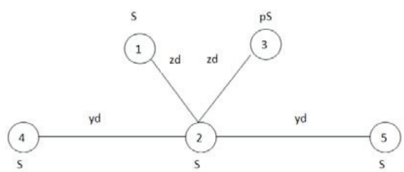

Figure 3. LTN Network

In this figure, S is sensed value by the nodes and d indicated the distance. Z, y, p are fix numbers. According to Debraj de voting method that is in the first category, the following relations are obtained. These relations are used to detect faulty readings of node 2. Node with Number 3 is a faulty node that is going to destroy the voting result.

$W_{12}=3/zd$ $\quad W_{23}=1/zd$ $\quad W_{24}=W_{25}=3/yd$

$Vote_i = (3/zd)*S - (1/zd)*pS + (6/yd)*S \rightarrow$

$$vote_i = \frac{S}{d} * (\frac{3}{z} - \frac{p}{z} + \frac{6}{y})$$

To node 3 can destroy voting must: $\frac{3}{z} - \frac{p}{z} + \frac{6}{y} < 0$

$\xrightarrow{\text{Multiply xy in both sides}} 3y+py-6z<0$

We can write wz instead of y, and then we have:

$$3wz - wpz + 6z<0 \xrightarrow{\text{Discard z}} 3w + 6<pw$$

As was seen, in this case if unequal is established, voting is wrong! For example, if w=2 then:

$$3*2z - 2pz + 6z < 0 \rightarrow 6 < p$$

According to this example, if p is greater than six voting failed. Consider that the accuracy is very low and very high risk of inaccuracy. With decreasing the distance between faulty node and voting node (increasing w) effect of the faulty node on voting brutally increases. The following Figure 4 shows this.

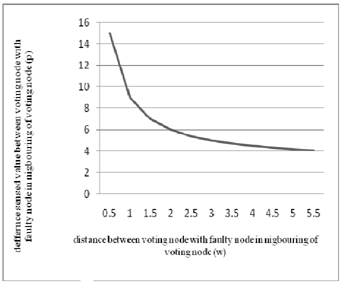

Figure 4. Effect of faulty nodes on Debraj de voting

To resolve this problem we can use algorithms that use the amount of their read vectors similarity or correlation as weight instead of the distance between two nodes. These algorithms are the second category of voting algorithms. They are often time consuming and expensive.

If the proposed fuzzy-based fault detection algorithm applied on the same data, the results are:

$Weight_1 = Weight_4 = Weight_5 = 1$
Absolutely $Weight_3 < 1$
$\rightarrow Voting_2 = +1+1+1- Weight_3 > +2$

The result is correct, which shows the high accuracy in fault detection. In addition, using fuzzy logic can individually improve the accuracy.

B) Weakness of the second category algorithm (The previous proposed method [6])

We for simplification of calculations consider the following assumptions:

- K: The number of sensed value, vote along weight and crisp output bits
- L: The length of read vector
- K=8;L=8

We in paper [6] tried to improve the energy consumption of the second category of voting by reducing the length of read vector but this method is still very costly because many bits is transferred in network to detect faulty readings. The following formula and Figure 5 shows this matter. Since the length of read vector in the previous proposed method is limited, we consider four as maximum length of read vector.

The number of transfer bits in the previous proposed method = L*K*4 + 4*K

The number of transfer bits in the proposed method = 8*K

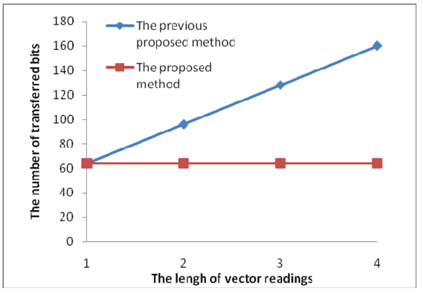

Figure 5. Transferred bits

It is concluded from what was said that the proposed method is low complex, energy efficient, low-delay (fast), accurate, distributed and scalable. See the table 2.

Table II. Comparing between faulty readings detection algorithms

Computational complexity	Energy consumption	Accuracy	Features / Method
High	High	High	**The previous method**
Low	Low	Low	**Debraj de**
Low	Low	Very High	**The proposed method**

VI. CONCLUSION

The main goal of this study was to propose an effective fuzzy-based faulty reading detection approach in wireless sensor networks. The entire read vector is used in the proposed method to improve the results accuracy. For improve the energy consumption for sending this vector, a fuzzy system is run on it and the crisp output is send instead of read vector. Therefore, the proposed method has very high accuracy than the first category voting algorithm and very low energy consumption than the second category algorithm.

We showed that the proposed method is energy efficient, accurate, distributed, low-delay (fast), scalable and robust approach to detect faulty readings.

REFERENCES

[1] E. Elnahrawy and B. Nath, "Poster Abstract: Online Data Cleaning in Wireless Sensor Networks," Proceedings of the 1st International conference on Embedded networked sensor systems, Los Angeles, pp. 294-295, 2003.

[2] X. Xiao, W. Peng, C. Hung, and W. Lee, "Using SensorRanks for In-Network Detection of Faulty Readings in Wireless Sensor Networks," In Proc. of MobiDE, pp. 714-721, 2007.

[3] T. Sun, L.J. Chen, C.C. Han, and M. Gerla, "Reliable Sensor Networks for Planet Exploration," In: The IEEE International Conference On Networking, Sensing and Control (ICNSC), pp. 816–821, 2005.

[4] B. Krishnamachari, and S. Iyengar, "Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks," IEEE Transactions on Computers, vol. 53, no. 3, pp. 241–250, 2004.

[5] D. De, "A distributed algorithm for localzation error detection correction, use in in-network faulty reading detection: applicability in long-thin wireless sensor networks," In Proc. of the IEEE Wireless Communications and Networking Conference, pp. 1-6, April 2009.

[6] A. Barati, S.J. Dastgheib, A. Movaghar, I. Attarzadeh, "An Optimised Algorithm to Detect Faulty Readings along the Substrate Access Wireless Long-Thin Sensor Networks", UKSim 5th European Symposium on Computer Modeling and Simulation, IEEE computer society, pp. 372-377, 2011.

[7] I. Gupta, D. Riordan and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," In: Proceedings of the 3rd Annual Communication Networks and Services Research Conference, IEEE Computer Society, Washington, DC, USA, pp. 255—260, 2005.

[8] J. Kim, S. Park, Y. Han and T. Chung, "CHEF: cluster head election mechanism using fuzzy logic in wireless sensor networks," In: 10th International Conference on Advanced Communication Technology (ICACT), Institute of Electrical and Electronics Engineers (IEEE), USA, pp. 654-659, 2008.

[9] H.Y. Lee and T.H. Cho, "Fuzzy logic based key disseminating in ubiquitous sensor networks," In: 10th International Conference on Advanced Communication Technology (ICACT), Institute of Electrical and Electronics Engineers (IEEE), USA, pp. 958-962, 2008.

[10] B. Kim, H. Lee and T. Cho, "Fuzzy key dissemination limiting method for the dynamic filtering-based sensor networks," ICIC, pp. 263-272, 2007.

[11] B. Lazzerini, F. Marcelloni, M. Vecchio, S. Croce and E. Monaldi, "A fuzzy approach to data aggregation to reduce power consumption in wireless sensor networks," In: Fuzzy Information Processing Society, NAFIPS, Annual Meeting of the North American, Institute of Electrical and Electronics Engineers (IEEE), USA, pp. 436-441, 2006.

[12] J.M. Kim and T.H. Cho, "Routing path generation for reliable transmission in sensor networks using ga with fuzzy logic based fitness function," In: Proceedings of the 2007 International Conference on Computational Science and Its Applications – Volume Part III, Springer-Verlag, Berlin, Heidelberg, pp. 637-648, 2007.

[13] S.-Y. Chiang and J.-L. Wang, "Routing analysis using fuzzy logic systems in wireless sensor networks," In KES, pp. 966-973, 2008.

[14] Q. Ren, and Q. Liang, "Fuzzy logic-optimized secure media access control (FSMAC) protocol wireless sensor networks," In: Proceedings of the 2005 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety (CIHSPS), Institute of Electrical and Electronics Engineers (IEEE), USA, pp. 37-43, 2005.

[15] S.A. Munir, Y.W. Bin, R. Biao and M. Jian, "Fuzzy logic based congestion estimation for QoS in wireless sensor network," In: Wireless Communications and Networking Conference (WCNC), Institute of Electrical and Electronics Engineers (IEEE), USA, pp. 4336-4341, 2007.

[16] F. Xia, W. Zhao, Y. Sun and Y.-C. Tian, "Fuzzy logic control based QoS management in wireless sensor/ actuator networks," Sensors 7 (12), pp. 3179-3191, 2007.

[17] M.S. Pan, H.W. Fang, Y.C. Liu and Y.C. Tseng, "Address assignment and routing schemes for ZigBee-based long-thin wireless sensor networks," in IEEE Vehicular Technology Conference, pp. 173–177, 2008.

[18] F.J. Wu, Y.F. Kao and Y.Ch. Tseng, "From wireless sensor networks towards cyber physical systems," Pervasive and Mobile Computing, Volume 7, Issue 4, Pages 397-413, August 2011.

[19] G. Montenegro, S. Dawkins, M. Kojo, V. Magret, and N. Vaidya, "Long Thin Networks", RFC 2757, January 2000.

[20] T.J. Ross, "Fuzzy Logic With Engineering Applications," Third Edition, A John Wiley and Sons, Ltd., Publication, ISBN: 978-0-470-74376-8, 2010.

[21] L.A. Zadeh, "Fuzzy logic," Computing & Processing (Hardware/Software), IEEE, Volume: 21 Issue: 4, pp. 83 – 93, 1988.

[22] Zhao, J., Bose, "BK: Evaluation of membership functions for fuzzy logic controlled induction motor drive," In: IECON Proceedings (Industrial Electronics Conference), vol. 1, pp. 229–234, 2002.

[23] W. Pedrycz, "Why triangular membership functions?," Fuzzy sets and Systems, Elsevier, Volume 64, Issue 1, pp. 21-30, 1994.

[24] Sh. Shamshirband, S. Kalantari, Z. S. Daliri and L. Sh. Ng, "Expert security system in wireless sensor networks based on fuzzy discussion multi-agent systems," Scientific Research and Essays Vol. 5(24), pp. 3840-3849, 18 December, 2010.

[25] A. Patel and B. Mohan, "Some numerical aspects of center of area defuzzification method," Fuzzy Sets and Systems, vol. 132, pp. 401-409, 2002.

[26] J.V. Sickle, "Using mean similarity dendrograms to evaluate classifications," Journal of Agricultural, Biological, and Environmental Statistics, Vol. 2, No. 4, pp. 370-378, Dec. 1997.