

Research Article

A Fuzzy Data Fusion Solution to Enhance the QoS and the Energy Consumption in Wireless Sensor Networks

Mario Collotta,¹ Giovanni Pau,¹ and Alexander V. Bobovich²

¹*Faculty of Engineering and Architecture, Kore University of Enna, 94100 Enna, Italy*

²*SUAI, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia*

Correspondence should be addressed to Mario Collotta; mario.collotta@unikore.it

Received 1 April 2017; Revised 8 June 2017; Accepted 21 June 2017; Published 20 July 2017

Academic Editor: Paolo Barsocchi

Copyright © 2017 Mario Collotta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wireless Sensor Networks (WSNs) are formed of various nodes that gather parameters in a monitored environment. These nodes interact with each other or can be arranged into clusters controlled by a cluster head that has the task of rerouting the acquired data to a base station. Both the Quality of Service (QoS) and low data quality are common issues in WSNs, mainly prompted by the data fusion mechanism, where a certain amount of low-quality data may affect the overall fusion result negatively. In this paper, a fuzzy-based solution for data fusion in WSNs is presented to provide a better QoS and to reduce the energy consumption. The suggested approach can aggregate only true value rather than process the full data. This purpose is accomplished thanks to a Fuzzy Logic Controller (FLC) implemented within nodes. Besides, the data, which have been separated, are aggregated by a cluster head which also has the responsibility of determining the probability that an event has happened in the monitored environment. Finally, the base station estimates whether an event has occurred and, eventually, raises an appropriate alarm. The results of a real testbed scenario reveal that the proposed method achieves encouraging performance.

1. Introduction

Wireless Sensor Networks (WSNs) are growing increasingly in everyday life. Home automation, health, industry, and road monitoring are some examples of application fields [1]. In recent studies, conducted in various scenarios, in which WSNs are applied, researchers accomplished that WSNs design has to meet specific features, such as low implementation cost [2], scalability [3], low power consumption [4], Quality of Service [5], and real-time constraints satisfaction [6].

In WSNs, sensor measures refer to the values that are gathered. These values are associated strictly with an event that is occurring in the monitored area. It is valuable to highlight that an error can happen if the data sensed by the device diverges from the real value. The value collected by the sensor may be wrong, and, therefore, this error occurs at sensor node level. In some cases, a sensor node itself might gather incorrect data. Moreover, the duplication of data might also lower the QoS. Obviously, both these

issues have a different (more or less) impact on the QoS. So, it is evident that specific solutions can be applied to resolve one or both these problems. Regarding errors caused by the sensing device, it is clear that a possible solution is to detect more data (or duplicate them), to understand the correlation between them. However, this mechanism, as mentioned above, may have a relevant impact on the QoS. As a consequence, it is appropriate to apply specific data fusion solutions, supported by different techniques, such as fuzzy logic, with the aim to solve the joint problem (but also individually). For instance, the authors of [7] implement a framework for advanced driver assistance systems, based on fuzzy logic and multisensor data fusion techniques, to reduce the driver's workload and the risk of road incidence. Their approach aims to decrease the load introduced by multiple sensors on board the vehicle through a combined approach between fuzzy logic and data fusion. So, it is a possible solution to the problem of detecting an incorrect value by the sensor and the duplication of the data. It is clear that the goal of data fusion solutions is to decrease the network load while

maintaining a certain reliability threshold. However, it is necessary to note that these values (network load, reliability) cannot be determined a priori applying them to all data fusion applications (wireless or wired), but they are closely related to the specific application, which has its own QoS requirements. Regarding the application fields of WSNs, the main aim of the data fusion mechanism is to achieve a greater QoS and to make proper decisions about the events of interest [8, 9]. In data fusion mechanisms, the data collected from various sensors is processed and, subsequently, new information that a single sensor cannot get is created. In fact, data fusion assures that the data quality of the WSN is improved, but also the energy consumption can be decreased since unnecessary information is discarded [10, 11].

In the literature, several data fusion mechanisms have been introduced with the goal of decreasing the energy consumption in the WSNs [12, 13]. These mechanisms are based on different methods such as probability theory [14], possibility theory [15], rough set theory [16], and Dempster-Shafer evidence theory [17]. Most of these approaches can eliminate the duplicated data in the fusion process. Nevertheless, these methods do not distinguish specific limitations of the sensor devices. In fact, most of them assume that the entire data, acquired by sensors, are the true value, which is an unreasonable hypothesis. Furthermore, all existing transmission schemes involve both necessary and unnecessary data to the base station. If, on the one hand, it is true that authors' approaches are interesting and promising, in terms of data fusion performance and energy efficiency, on the other hand, it is evident that additional intelligent and innovative solutions can be developed to further improve both the whole data fusion process and, primarily, the energy efficiency.

In recent years, many research works implement soft computing techniques also in data fusion applications [18]. Considering that input and output values of Fuzzy Logic Controllers (FLCs) [19] are real variables mapped with a nonlinear function, they are suitable for numerous engineering problems, mainly for combined problems where traditional control methods do not reach comparatively favorable outcomes. For this reason, in this paper, a fuzzy-based solution for data fusion in WSNs is introduced to provide a better QoS and to decrease the energy consumption. In the proposed data fusion method, a Fuzzy Logic Controller (FLC) is implemented in each sensor of the WSNs to differentiate the real data values from measured data values. Afterward, cluster heads combine data [20]. They are also able to estimate the probability that an event has happened in the monitored environment. During this process, additional data is erased and, consequently, the energy consumption is minimized. The main contribution of the proposed solution is to improve the system performance concerning the QoS and to minimize the power consumption by transferring only the determined probability of the events instead of the entire fused data. However, it is necessary to note that, in some WSN applications, such as, safety systems and industrial automation systems, all the available data must be transmitted. In these application scenarios and real-time WSNs, the solution introduced in this paper could not be suitable. In fact, in real-time or industrial WSNs [21], the main aim is

to meet requirements regarding QoS parameters, bounded latency, high reliability, and dependability. The employment of a data fusion mechanism can undoubtedly raise problems regarding the constraints of the real-time application. For this reason, the fuzzy-based solution introduced in this paper is highly appropriate in applications that do not require strict real-time constraints. Besides, the proposed approach is certainly usable even in soft real-time contexts, where the deadline overruns do not imply irreparable damage to the entire system but only a slight deterioration in performance. A possible application scenario of the solution proposed in this paper can be a generic environmental monitoring, indoor and outdoor, such as those considered in [22–25]. In fact, this considered reference scenario is not related to real-time constraints since the possible parameters to monitor are the temperature, the humidity, and the carbon monoxide. When their values exceed a certain threshold, then it is evident that an event has occurred in the monitored environment. The purpose of the solution introduced in this paper is to provide a better QoS but also, fundamentally, to reduce the energy consumption of battery-powered wireless devices.

This paper is organized as follows. Section 2 presents an analysis of existing approaches in the literature. Section 3 describes the system requirements and introduces the proposed approach. Section 4 shows the performance obtained with proposed approach and, finally, in Section 5 the paper is summarized, reporting the conclusions.

2. Related Works

Data fusion approaches can be categorized based on several features. In fact, the similarities among the input data can be utilized to classify data fusion into valuable data, redundant data, and corresponding data. All aggregating protocols are classified into two main categories, which are structure-based and structure-free. In the following subsections, both fuzzy-based and structure-based/free methods for data fusion presented in the literature are briefly reviewed.

2.1. Fuzzy-Based Data Fusion Approaches. A method based on fuzzy-evaluation for data fusion, implemented to an environmental monitoring system with multisensors, is introduced in [26]. At first, the amount of the humidity, temperature, ventilation, and other parameters, collected by multiple sensors, is fuzzed. The scheme proposed by authors not only guarantees the accuracy of the monitoring system but also enhances its intelligence. In fact, simulation results reveal that the performance of this technique overcomes that of conventional diagnosis methods, for instance, with single sensors. Nevertheless, the authors do not address the power consumption issue of sensor nodes and do not establish how to implement their approach in real devices.

The authors of [27] suggest a new fuzzy data fusion algorithm, based on variable weight, for WSNs to enhance the reliability and efficiency of the global data fusion method. In the algorithm proposed by the authors, the weight of all cluster head nodes is not fixed in the entire fusion mechanism. In fact, considering that the amount of data,

the time delay, and the trustworthiness of each cluster head affect the weight of the final fusion, these weights are defined by a fuzzy inference system. Simulation results show that this algorithm can produce more reliable and accurate fusion results compared to other literature approaches, particularly when there are undetected data or compromised nodes. Nevertheless, this approach has been only simulated by the authors. As a result, it is not possible to consider what would be its impact on an implementation in real devices. In fact, the performance may differ significantly from those obtained through simulations.

A method for fire monitoring and forecasting, based on the fuzzy theory for multisensors data fusion, is presented in [28]. Even in this case, the results confirm the effectiveness of a fuzzy-based monitoring method. On the contrary, a fuzzy logic approach for data fusion to detect faults within a WSN, adopting a statistical process control and a clustered covariance method, is introduced in [29]. In this case, the proposed method aims to overcome uncertainty and false-positives within the fault discovery process in a WSN. The obtained results are promising. However, in the solutions suggested in [28, 29], the power consumption of sensor nodes is not addressed, and, moreover, the authors do not consider the QoS as the main requirement for their application.

The authors of [30] present a new hierarchical system for fault determination in a smart power grid, considering the fusion of information obtained by multidata resources. The system is based on fuzzy Petri net. Moreover, it uses dual-data resources to investigate fault diagnosis based on the information of the switch and the electricity. The solution proposed by the authors can ascertain the network topology and the directional weighted fuzzy Petri nets. The fuzzy *c*-means clustering method is employed to outline the final diagnosis decision. Simulation results show that the algorithm not only can effectively solve the drawbacks by using the single data resource for diagnosis but also can achieve the exact result in the case of incomplete information. Nevertheless, the fuzzy approach, proposed in [30], requires more computation compared to a type-1 fuzzy controller, such the one suggested in this work. Therefore, its implementation may require powerful and expensive hardware, which is not feasible in WSNs.

A new method to obtain a decision through data fusion even when the sources are dissimilar is presented in [31]. In the proposed method, instead of indicating one single weight to each classifier, the distribution of reliability of each source, over the input space, is determined using a fuzzy approximation function. The latter is an assuring framework for dealing with uncertain and incomplete information. The results gained through simulations demonstrate the advantages of the proposed method in reaching decisions with disparate information sources. However, even in this case, the energy consumption is not taken into account.

2.2. Structured-Based and Free-Based Fusion Methods. The structure-based techniques can be grouped into three categories, which are clustering, tree, and grid. It is useful to note that, usually, in cluster-based data fusion approaches, the sensor nodes are clustered into distinct groups each of

which has its cluster head. These nodes deal with the merging of collected data to decrease information transmission in wireless networks [32]. In [33], the goal of the authors is to simplify the data fusion mechanism in a WSN. To this end, they propose a protocol for sensor information systems in which all nodes must be placed in a stationary linear chain. Moreover, these nodes need to be conscious of each other positions to determine the closest neighbors. It is useful to note that in the protocol proposed in [33] all the sensors need to combine the obtained data, to provide the same size of collected data, and transmit it to the nearest neighbor. This method constantly happens to reach the closest node to the sink, named leader node (i.e., the cluster head), that fuse the whole acquired data packets. The results presented in [33] prove that the proposed protocol can be employed in a large-size WSN, but it requires significant energy consumption.

The main aim of tree-based methods is to arrange all sensor nodes into a tree and make them capable of aggregating the data [32]. Subsequently, a reduced amount of data is forwarded to the root node of the tree. The central feature of tree-based approaches is labeling the data through the information of the whole network. There are some benefits in employing tree-based techniques, such as the reduction of the communication among neighboring nodes. Nevertheless, there are also lacks. In fact, not only the power consumption is considerably high but also the delay introduced by the protocol is not negligible, notably in a large-size network. A suboptimal tree-based fusion protocol is presented in [34] trying to overcome these difficulties. The main purpose of the authors is to produce a plausible minimum tree towards the base station. The results obtained through simulations are promising. However, it is important to highlight that the proposed solution does achieve reliable performance when network conditions are fickle.

In grid-based data fusion approaches, sensor nodes cannot communicate with each other. In this case, they compose a particular virtual grid used to forward data packets only towards a predefined data aggregator. The authors of [35] propose a hybrid system consisting of in-network and grid-based fusion techniques. The goal is to increase the performance of the data aggregation. In fact, an aggregator node is determined when an event happens, whereas the other nodes still preserve the log of previous occurrences in a specific table. On the contrary, when a sensor recognizes an event, it examines its table (past situations) and distinguishes the nature of the current event. Although this proposed hybrid system is attractive and has sufficient performance, all the data from each node is required to be elaborated and forwarded. As a consequence, the energy is not utilized efficiently. Besides, an additional mechanism to limit data repetition is not taken into account.

Structure-free data fusion approaches are the appealing solution that can be employed in WSNs with periodic variations [32]. In fact, this method is likewise suitable for monitoring and control systems that do not expect any planned arrangement. It is useful to highlight that various studies for data aggregation have been introduced in this domain. For instance, the algorithm presented in [36] applies both fuzzy inference and neural networks. The approach,

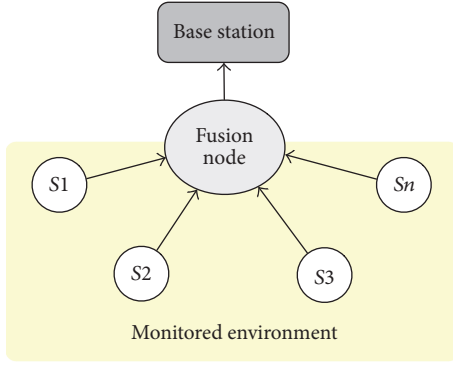


FIGURE 1: Fusion mechanism diagram.

after collecting and analyzing some environmental information, such as smoke density and temperature, outlines whether a fire has happened. More in detail, the solution proposed by the authors improves the correction precision, but both the power consumption and data compression are not taken into account. Still in the research field of structure-free data fusion approaches, in [27] a data fusion approach for WSNs, based on variable weight fuzzy inference system, is suggested by the authors. The goal is to enhance the reliability and the efficiency of the data fusion mechanism in a Wireless Sensor Network. In this case, unfixed and distinctive fusion weights are designated to each cluster head. These weights are calibrated using a fuzzy logic inference system, considering several features, such as the integrity of the cluster head, the quantity of data, and the delay in network communication. As a result of the mechanism proposed by the authors, the acquired and fused data packets are more accurate and reliable. However, the energy consumption issue is not addressed even in this case.

Section 2 has revealed explicitly that in the literature various techniques aim to improve the QoS of the network through the data fusion. Nevertheless, most of them did not take into account the message overhead, the energy consumption, and, mainly, the implementation on real devices.

3. The Proposed Approach

3.1. Requirements. A data fusion process can be applied in a WSN to bypass heavy traffic and to save energy. These problems are caused by the transmission of redundant data packets to the base station from each sensor node. For this reason, a data fusion scheme can be adopted to aggregate the collected data from the other sensors in every node belonging to a generic WSN. On the contrary, regarding data fusion mechanism introduced in this work, a generalized block diagram is shown in Figure 1. The sensor nodes $S1, S2, \dots, Sn$ gather data from the monitored environment. These nodes, then, forward the obtained data to a fusion node through a direct point-to-point wireless link. Afterward, this node will produce one single internal representation of the monitored environment from received inputs. The base station then utilizes this unique representation. As a consequence, the base station, in general, does not have access to the individual

sensor measurements. The conditions of both the monitored environment and the sensor nodes can also change swiftly, and, for this reason, the data acquired by the base station might not be the true values. Various useless data packets are generated and forwarded from each sensor node in every round of data gathering, but, in many circumstances, a sensor node is not able to acknowledge them, while it creates data packets for further process, whereby these data are transmitted to the fusion node and, consequently, aggregated with the other true value data. As a result, they can have negative effect on making proper decisions. Moreover, this can lead to quite high energy consumption in the WSN. So, it is necessary to better manage the unnecessary data, in addition to redundant data transferred from sensor nodes.

Another requirement to consider is the data quality. In fact, to enhance the QoS in the WSN, it is essential to make sure that the readings and the measurements of the sensor are true values to be used. Several environmental characteristics can affect a sensor node behavior. These factors can decrease or increase the output signal of the sensors. Furthermore, nonoperating environmental limits, such as a high or low temperature, usually influence the sensors' performance.

Data redundancy is another major requirement. Specifically, the redundancy in WSNs is defined as the use of redundant data, for instance, extra bits to describe the events of interest. In most cases, the quality of the obtained data can decrease if information, for a specific event in a location, is received from different sources. Typically, sensor nodes in a WSN are deployed densely, and thus they provide a significant amount of redundant and duplicate data. This feature may cause packet collisions, bandwidth waste, and energy consumption. On the contrary, the removal of additional data improves the overall quality of the acquired data and decreases the number of transmissions from source nodes to the base station. However, the fuzzy-based solution introduced in this paper is highly suited for applications that do not require strict real-time constraints. In addition, the proposed approach is certainly useful even in soft real-time contexts, where the deadline overruns do not imply irreversible damage to the entire system but only a slight deterioration in performance. It is necessary to use proper techniques and approaches to overcome the issue of redundant and duplicate data. The key idea is that sensor nodes combine and transfer their data generating a lower traffic, instead of sending messages for each event individually. Besides, it could be useful to determine only the probability that an event has happened at the base station, whereby only the necessary information related to estimated probabilities can be sent out to the base station. If the probability of existing events is more than a predetermined threshold, the system transmits the information to the base station.

3.2. Fuzzy Logic Controller. It is necessary to specify that the application context of the solution presented in this paper refers to a general environmental monitoring. In this environment, different physical quantities are gathered, such as, temperature, humidity, or carbon monoxide. The scheme of the proposed data fusion solution is depicted in Figure 2. At the first step, the sensor nodes gather data ($D1, D2, \dots, Dn$)

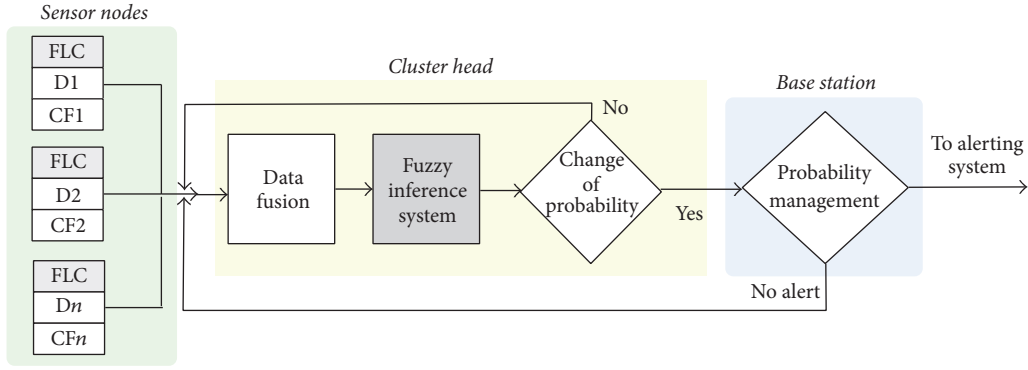


FIGURE 2: The proposed data fusion solution.

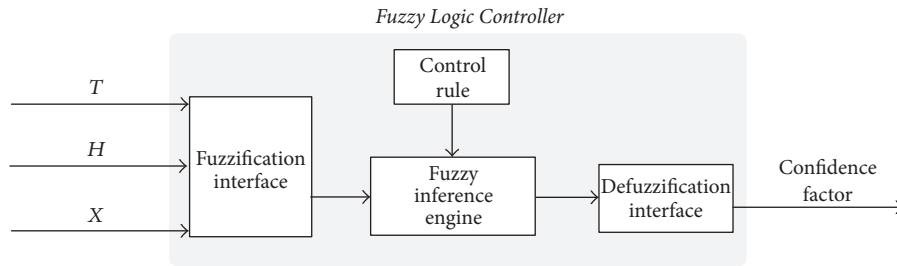


FIGURE 3: Proposed Fuzzy Logic Controller.

and determine a confidence factor (CF) for each collected data packet in each round. A Fuzzy Logic Controller (FLC) is implemented on each sensor node to calculate the CF individually. The purpose of this FLC is to obtain a confidence level of the obtained data considering the current condition of the sensor nodes.

The FLC, implemented in the sensor node, is shown in Figure 3. The controller receives crisp input values of temperature (T), humidity (H), and carbon monoxide (X) gathered by the sensors, which are converted into linguistic values by using a chosen set of membership functions. The purpose of a FLC is to refine these linguistic values through an inference system, by using a set of if-then rules. In detail, these rules are coupled in the FLC, which returns a membership function, represented, in this paper, by Gaussian functional shapes [37]. A Gaussian membership function is defined as follows:

$$f(x, \sigma, c) = e^{-(x-c)^2/2\sigma^2}, \quad (1)$$

where c is the membership function center and σ defines the membership function width. The linguistic output can be outlined through this function. The Gaussian model takes crisp inputs and provides crisp outputs. This mechanism is carried out by user-defined fuzzy rules and taking into account also user-defined fuzzy variables. The procedures of the Gaussian rules can be defined in four steps:

- (i) Mapping each of the crisp inputs into a fuzzy variable (fuzzification)

- (ii) Determination of the output of each rule given its fuzzy antecedents
- (iii) Determination of the aggregate output(s) of all of the fuzzy rules
- (iv) Mapping of the fuzzy output(s) to crisp output(s) (defuzzification)

Gaussian functional shapes have been adopted because the precision increases considerably, without degrading the computational performance, as confirmed in [38]. The FLC produces a confidence factor (CF_1, CF_2, \dots, CF_n) for each sensor data (D_1, D_2, \dots, D_n) that is obtained in real-time. The FLC determines if the temperature, the humidity, and the carbon monoxide values are in the satisfactory range. To achieve that, FLC examines its input measures with the desired range for each sensor. The latter, for each sensor, can be found on its specified datasheet. Anyway, the occurrence of an event must meet some conditions (thresholds are exceeded), such as temperature ($T > 100^\circ\text{C}$), humidity ($H < 50$ (RH%)), and carbon monoxide ($X > 100$ (ppm)).

It is essential to note that there is no issue where the sensor nodes behave abnormally, for instance, by interacting with an event out of context, since they are planned only to gather environmental parameters and to treat them through the fuzzy controller. The output of the FLC, for each sensor, can be 100% only if the environmental factors are in the acceptable range. If the obtained value is outward the range, the FLC returns a confidence factor ($0\% \leq CF_n < 100\%$) for the gathered data. Also, in this case, each node compares the

estimated CF with a decided threshold. If the generated CF is less than the threshold value, then the data is disregarded. Contrarily, it is forwarded to the cluster head in a data message. This action limits the occupation of data that does not present the true values. The final step of the FLC consists in the conversion of the received CF value into crisp logic decisions (defuzzification). The defuzzification process of the proposed FLC is based on Mizumoto's functions centroid and maximum [39].

As depicted in Figure 2, at the second step, the cluster heads are responsible for aggregating the acquired data from the cluster members (sensor nodes). The cluster heads start the data fusion process at the end of each round of data collection. To aggregate the full obtained data from cluster members, the following equation is adopted:

$$AD = \frac{(CF1 * D1) + (CF2 * D2) + \dots + (CFn * Dn)}{CF1 + CF2 + \dots + CFn}, \quad (2)$$

where Dn is the acquired data from the same kind of cluster members, CFn is the confidence factor of the obtained data, and AD (Aggregated Data) is a mixture of the data from the same kind of sensor nodes. In fact, AD is a combination of the data with a higher confidence as it combines the entire data received correctly based on their confidence factors. Considering various sources of one kind of sensor nodes in distinct locations, AD provides an accurate view of the monitored environment. It is necessary to note that AD is estimated for the three classes of deployed sensor nodes separately (T, H, X) and, for this reason, there is a set of AD that is saved in a matrix with one row and n columns. The following equation represents this matrix:

$$M_{AD} = \{AD1, AD2, \dots, ADn\}. \quad (3)$$

For instance, let us consider various sensors (T, H, X) deployed on a cluster-based approach. Over a period, the temperature sensors can gather the values of 25°C, 20°C, and 15°C with the confidence factors of 0.82, 0.77, and 0.58, respectively. As a result, considering (2), the aggregated temperature data (AD_T) is 20.55. In the same way, the aggregated humidity data (AD_H) and the aggregated carbon monoxide data (AD_X) are 37.35 and 28.18, respectively. Considering (3), in this case, the matrix would be

$$M_{AD} = \{20.55, 37.35, 28.18\}. \quad (4)$$

As depicted in Figure 2, the obtained M_{AD} is processed by another fuzzy inference system in the cluster head. In this case, the output is the probability of occurring events in the monitored environment. This Fuzzy Interface System analyzes M_{AD} based on the provided fuzzy rules. The FLC merges these rules and returns a membership function, represented, in this paper, by Mamdani triangular functional shapes [37]. As a result, it is feasible to set, through the inference mechanism, the correct output according to the fuzzy inference rules presented in Table 1.

Let us consider an example to determine an event occurrence. Considering rule 3 in Table 1, if the measured valued of AD_T is "low" and the others related to AD_H and AD_X

TABLE 1: FLC inference rules.

Rule	AD_T	AD_H	AD_X	Probability
1	Low	Low	Low	Low
2	High	Low	Low	Low
3	Low	High	Low	Low
4	Low	Low	High	Low
5	High	High	Low	High
6	High	Low	High	High
7	Low	High	High	High
8	High	High	High	High

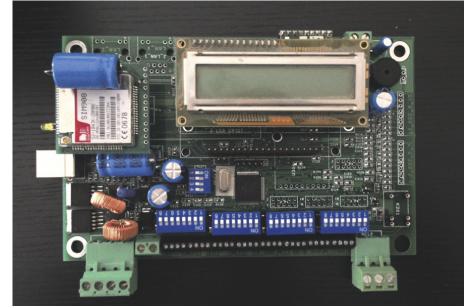


FIGURE 4: Used prototyping board.

are "high" and "low," respectively, the probability that an event has occurred is "low." In other words, if the calculated probability in the clustered nodes has not changed, the cluster head does not forward the data packet to the base station. On the contrary, the probability of the events in their controlled areas is forwarded by the cluster heads. However, the change in the probability does not guarantee an accurate detection, and it is only acknowledged as a potential event in the monitored environment. For this reason, to make sure that the detection is correct, the base station requires controlling and processing regularly the collected probabilities that are produced over the time. In fact, all the obtained probabilities are constantly treated by the base station. If a continuous variation is noticed, the event is notified to the alerting subsystem by the base station. On the contrary, the algorithm proceeds in collecting data. It is necessary to note that the mechanism in cluster heads to communicate their data to the base station can operate even when the channel condition is poor because only a few data that do not require the optimal conditions in the channel are forwarded. In this way, the base station can collect updates from cluster heads and then assess whether there have been changes in the probabilities.

4. Experimental Results

A real test-bed scenario has been deployed to evaluate the soundness of the proposed solution. The processing unit of the used prototyping boards (Figure 4) is the Microchip PIC24FJ256GB108 microcontroller [40], which combines the control characteristics of a microcontroller unit with features

TABLE 2: Operating range of the sensors used in the test-bed.

Sensor	Operating range
Temperature	-40 to 125 [°C]
Humidity	0% to 100% [RH]
Carbon monoxide	0 to 10000 [ppm]

typical of a digital signal processor (DSP). This microcontroller is perfect for low power applications. Moreover, another important feature is the availability of many serial ports and autonomous timers. Furthermore, it is suitable for embedded control and monitoring applications, due to the large amounts of memory for buffering and flashing. The prototyping boards deployed in the test-bed scenario have been provided with the IEEE 802.15.4 MRF24J40MB transceiver from Microchip Technology [41], temperature, and humidity sensors (SHT21 Humidity/Temperature Sensor) from SENSIRION [42] and an electrochemical carbon monoxide sensor (TGS 5042) from FIGARO. [43]. Table 2 shows the operating range of the used sensors. An amount of 14 nodes has been used: 12 sensors, 1 fusion node, and 1 base station. The same prototyping board, equipped with the transceiver [41], has been used for all nodes.

The main aim of the experimental scenario has been to analyze the expected amount of truly collected data also evaluating the QoS in the WSN. In each round, several data packets have been produced and forwarded. The percentage of correctly received data, the average value of transmitted packets, and the ratio of Throughput to Workload (TW) have been measured. TW value has been obtained according to the following equation:

$$TW = \frac{Th}{Wl}, \quad (5)$$

where Th (throughput) is the number of packets transmitted by the device and Wl (workload) is the total number of packets that the device has to send. Moreover, the energy consumption, in each round of data collection, has been estimated according to the approach proposed in [44]. The obtained results have been displayed on an LCD screen connected to the prototyping board. Besides, the microcontroller continuously sent the output data to a computer by a serial cable to measure and plot the performance. The results have been compared with a variable weight based fuzzy data fusion algorithm for WSN proposed by Wang et al. in [27] because it is a distributed fuzzy-based data fusing algorithm that enhances the QoS in WSNs, similar to the one introduced in this paper. Moreover, the proposed solution has also been compared with the fuzzy-based data fusion approaches presented in [28] by Wang et al and in [29] by Shell et al. However, it is necessary to note that in the latter two methods the power consumption of sensor nodes is not addressed, and the authors do not consider the QoS as the main requirement for their application. At last, it is necessary to note that the duration of the test campaigns has been 90 minutes with a sampling time set to 1 second.

The first obtained results regard the percentages of correct received data by the base station, using both the proposed

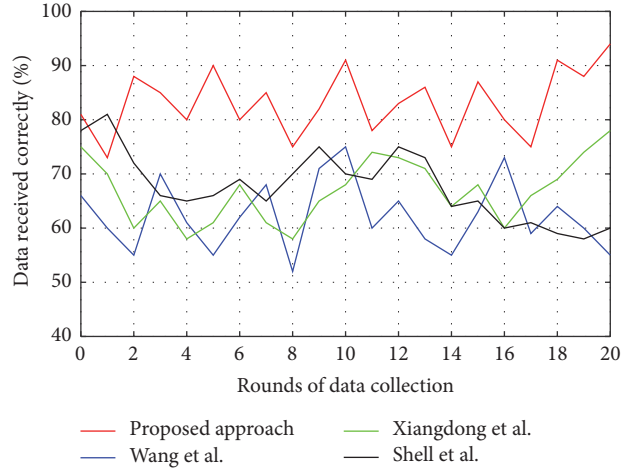


FIGURE 5: Data successfully received by the base station.

solution and Wang et al.'s. one, in each round of data gathering. For simplicity, in Figure 5, only the first 20 rounds of data collection are depicted. As it is possible to note, the proposed method provides a greater percentage of valid data compared with the other solution. These results have been attained due to the removal of incorrect data to prevent the corruption of valid data throughout the fusion process. Besides, the data of fusion process have been forwarded into the cluster head to evaluate whether there was a difference associated with the collected data. At the end, the base station, processing the received data, can discover if an event is recognized precisely. It is important to highlight that Wang et al. method achieved lower performance since the quality of gathered data is estimated based on their designated weights. Nevertheless, these weights are counted on all data, covering both true and incorrect values of data. Consequently, in Wang et al., since the cluster head aggregates the entire collected data, the true values might be affected by the incorrect data in the fusion process. The solutions of Xiangdong et al. and Shell et al. achieve a lower performance since they do not consider the QoS as the main requirement for their application.

The second test campaign has been focused on determining the average of the data packets that have been forwarded to the base station. Figure 6 presents the results obtained with the proposed solution compared to those of Wang et al., Xiangdong et al., and Shell et al. In some cases, during the measurements, the values obtained with the proposed method are lower because the entire data do not need to be transmitted. In fact, they are not all true values. In the proposed approach, the incorrect values are excluded from transfer and modeling processes. Furthermore, the cluster head does not expect to transport the fused data as it can discover the probability that an event has happened. The cluster head sends the assessed probability only if irregularity is identified, after a comparison with historical data. In this way, the data that is addressed to the base station represents only the chance that an event has occurred. Even in this case, the approaches of Wang et al., Xiangdong et al., and Shell et al. have obtained inferior performance because their

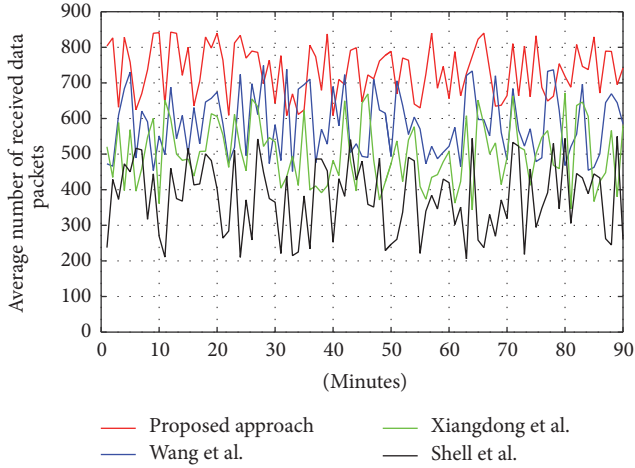


FIGURE 6: Average of the data packets transmitted to base station.

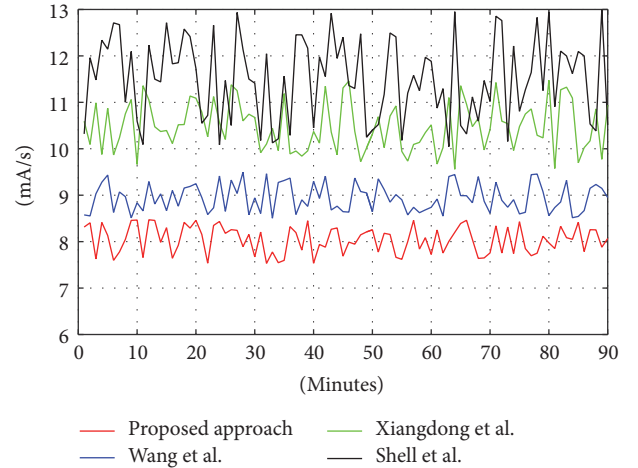


FIGURE 8: Energy consumption comparison.

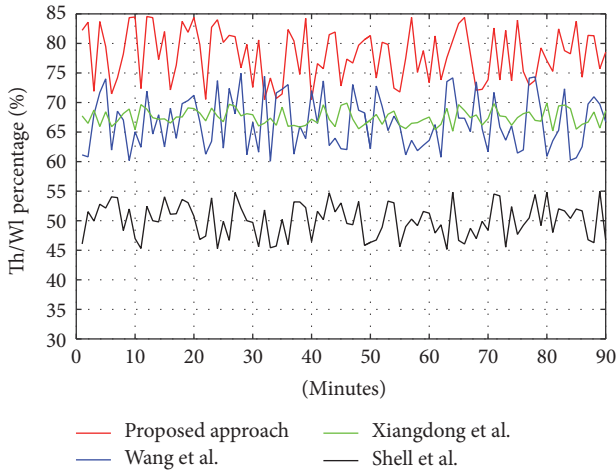


FIGURE 7: Th/Wl behavior.

suggested network topology is cluster-based and the obtained data are mixed based on the prescribed weights on clusters. As a consequence, these designated weights strengthen the impact of the incorrect data on the true values during the fusion process.

In Figure 7, the Th/Wl behavior obtained with the proposed solution compared to those of Wang et al., Xiangdong et al., and Shell et al. is depicted. As it is possible to note using the suggested method, Th/Wl fluctuates between 85% and 70%. These values are satisfactory mainly in those application scenarios with a moderate or high variation of data, for example, temperature, humidity, or carbon oxide detection, where the essential thing is to assure a high Th/Wl performance because alarms must be raised in real-time. The Th/Wl behavior using Wang et al. proposal shifts within 75% and 60%, while those of Xiangdong et al. shift within 70% and 65%. Although these values may be admissible in a monitoring context, even in this case, they achieve lower performance than the approach introduced in this paper. The worst performance has been obtained with the approach of

Shell et al. since Th/Wl fluctuates between 55% and 45%. In this case, these values could not be suitable for a monitoring context.

Finally, the energy consumption has been measured, and the obtained results are shown in Figure 8. These results are very encouraging and prove how the proposed fuzzy-based solution allows a significant decrease in the sensor's energy consumption through a careful management of their activities. In the proposed approach the amount of data packets within the network is decreased, and, for this reason, the power consumption is reduced significantly. As presented in Figure 8, the proposed solution minimizes the energy consumption in the WSN, gaining better performance than the other methods. The reason is that a lower number of packets are transferred in the network. The approaches of Xiangdong et al. and Shell et al. achieve the worst performance, regarding the energy consumption, because, as mentioned above, they do not address the power consumption.

5. Conclusions

A fuzzy-based solution for data fusion in WSNs to provide a better QoS has been proposed in this paper. The proposed method applied a cluster-based network, where cluster members gather data from the monitored environment constantly until a target event is identified. The certainty that the data to be forwarded within the network matches that truly gathered ones is a key feature in a WSN. For this reason, in the proposed approach, a Fuzzy Logic Controller has been realized in each sensor node. In fact, this FLC taking as input the collected values, such as temperature, humidity, and carbon monoxide, defines a confidence factor for each obtained data. Each sensor can group its data into true and untrue data utilizing the calculated confidence factor. Afterward, for further processing, these data are forwarded to the cluster head whose main task is to aggregate the collected data. The cluster head, instead of transmitting the summed data to the base station, estimates and forwards the probability that a specific event has occurred in the monitored

environment. This probability is determined through a fuzzy inference system realized within the cluster head. If the estimated probability is greater than a set threshold, then it is delivered to the base station that verifies whether the obtained data is accurate enough before alerting the alarm system.

Several test campaigns using real prototyping boards have been carried out to assess the validity of the suggested solution. The percentage of correctly received data, the average value of transferred packets, the ratio of Throughput to Workload, and the energy consumption have been determined. The results achieved with the proposed method are very encouraging and, besides, have been compared to another approach in the literature, always getting better performance.

The usefulness of the proposed system on real components has been proven through an implementation on the Microchip PIC24FJ256GB108 microcontroller [40], a COTS device available at affordable price. The possible impact of the proposed solution is broad as it can be widely and efficiently applied in practice, being a nonexpensive system to realize.

In the future, the solution introduced in this paper will be examined on networks with a higher density. Its complexity and the possibility of introducing further input variables will be assessed to improve the inference mechanism of the approach and to measure how it enhances both the whole network life-cycle and the QoS.

Another objective for future research on the system here addressed is to augment the solution suggested in this paper with a neural network capable of forecasting the environmental conditions, that is, to predict the requirements at different times of the day or on different days of the week. This combination would enable the fuzzy controller to perform its choice taking into account not only the current status as detected by the WSN but also the possible short-term evolution of the environmental conditions.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] C. Zhu, L. Shu, T. Hara, L. Wang, S. Nishio, and L. T. Yang, "A survey on communication and data management issues in mobile sensor networks," *Wireless Communications and Mobile Computing*, vol. 14, no. 1, pp. 19–36, 2014.
- [2] T. Matsui and H. Nishi, "Analysis and implementation of WSN with route selection considering energy consumption," in *Proceedings of the 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm '16)*, pp. 1–7, Sydney, Australia, November 2016.
- [3] Y. Liu and B. Seet, "Optimal network structuring for large-scale WSN with virtual broker based publish/subscribe," in *Proceedings of the 2017 2nd Workshop on Recent Trends in Telecommunications Research (RTTR '17)*, pp. 1–5, Palmerston North, New Zealand, February 2017.
- [4] H. Khanmirza and N. Yazdani, "Game of energy consumption balancing in heterogeneous sensor networks," *Wireless Communications and Mobile Computing*, vol. 16, no. 12, pp. 1457–1477, 2016.
- [5] T. Semprebom, C. Montez, G. M. De Araujo, and P. Portugal, "Skip game: an autonomic approach for QoS and energy management in IEEE 802.15.4 WSN," in *Proceedings of the 20th IEEE Symposium on Computers and Communication (ISCC '15)*, pp. 1–6, July 2015.
- [6] R. A. Lara-Cueva, R. Gordillo, L. Valencia, and D. S. Benitez, "Determining the main CSMA parameters for adequate performance of wsn for real-time volcano monitoring system applications," *IEEE Sensors Journal*, vol. 17, no. 5, pp. 1493–1502, 2017.
- [7] M. R. Ghahroudi, M. Sarshar, and R. Sabzevari, "Introducing a sensor network for advanced driver assistance systems using fuzzy logic and sensor data fusion techniques," *Ad-Hoc and Sensor Wireless Networks*, vol. 8, no. 1-2, pp. 35–55, 2009.
- [8] A. Zaidi, B. Ould Bouamama, and M. Tagina, "Bayesian reliability models of Weibull systems: state of the art," *International Journal of Applied Mathematics and Computer Science*, vol. 22, no. 3, pp. 585–600, 2012.
- [9] J. Liang, Z. Wang, and Q. Liang, "Adaptive sensor selection for multitarget detection in heterogeneous sensor networks," *Ad-Hoc and Sensor Wireless Networks*, vol. 12, no. 3-4, pp. 325–342, 2011.
- [10] L. Li and W.-J. Li, "The analysis of data fusion energy consumption in WSN," in *Proceedings of the International Conference on System Science, Engineering Design and Manufacturing Informationization (ICSEM '11)*, vol. 1, pp. 310–313, IEEE, October 2011.
- [11] X. Wang, H. Qi, and S. S. Iyengar, "Collaborative multimodality target classification in distributed sensor networks," in *Proceedings of the 5th International Conference on Information Fusion (FUSION '02)*, vol. 1, pp. 285–290, July 2002.
- [12] A. Abdelgawad and M. Bayoumi, "Data fusion in WSN," *Lecture Notes in Electrical Engineering*, vol. 118, pp. 17–35, 2012.
- [13] J. M. Bahi, A. Makhoul, and M. Medlej, "A two tiers data aggregation scheme for periodic sensor networks," *Ad-Hoc and Sensor Wireless Networks*, vol. 21, no. 1-2, pp. 77–100, 2014.
- [14] D. Gavalas, A. Mpitziopoulos, G. Pantziou, and C. Konstantopoulos, "An approach for near-optimal distributed data fusion in wireless sensor networks," *Wireless Networks*, vol. 16, no. 5, pp. 1407–1425, 2010.
- [15] E. J. Wright and K. B. Laskey, "Credibility models for multi-source fusion," in *Proceedings of the 2006 9th International Conference on Information Fusion (FUSION '06)*, pp. 1–7, July 2006.
- [16] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Springer, Dordrecht, Netherlands, 1991.
- [17] R. Haenni and S. Hartmann, "Modeling partially reliable information sources: a general approach based on Dempster-Shafer theory," *Information Fusion*, vol. 7, no. 4, pp. 361–379, 2006.
- [18] B. Ma, N. Li, C. Wang, Z. Han, and T. Guo, "Based on fuzzy neural network of multi-agent data fusion," in *Proceedings of the 2012 International Conference on Modelling, Identification and Control (ICMIC '12)*, pp. 975–980, June 2012.
- [19] J.-Q. Gao, L.-Y. Fan, L. Li, and L.-Z. Xu, "A practical application of kernel-based fuzzy discriminant analysis," *International Journal of Applied Mathematics and Computer Science*, vol. 23, no. 4, pp. 887–903, 2013.
- [20] Y. Sun, H. Wang, K. Zhang, and X. Yang, "Associated clustering strategy for wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 2014, Article ID 817234, 2014.

- [21] M. Collotta, G. Pau, and V. Maniscalco, "A fuzzy logic approach by using particle swarm optimization for effective energy management in IWSNs," *IEEE Transactions on Industrial Electronics*, no. 99, pp. 1-1, 2017.
- [22] P. Barsocchi, E. Ferro, L. Fortunati, F. Mavilia, and F. Palumbo, "EMS@CNR: an energy monitoring sensor network infrastructure for in-building location-based services," in *Proceedings of the 2014 International Conference on High Performance Computing and Simulation (HPCS '14)*, pp. 857-862, July 2014.
- [23] R. Bolla, R. Rapuzzi, M. Repetto, P. Barsocchi, S. Chessa, and S. Lenzi, "Automatic multimedia session migration by means of a context-aware mobility framework," in *Proceedings of the 6th International Conference on Mobile Technology, Application Systems (Mobility '09)*, pp. 1-36, ACM, New York, NY, USA, 2009.
- [24] P. Barsocchi, M. G. C. A. Cimino, E. Ferro, A. Lazzeri, F. Palumbo, and G. Vaglini, "Monitoring elderly behavior via indoor position-based stigmergy," *Pervasive and Mobile Computing*, vol. 23, pp. 26-42, 2015.
- [25] D. Bacciu, S. Chessa, C. Gallicchio, A. Micheli, and P. Barsocchi, "An experimental evaluation of reservoir computation for ambient assisted living," *Smart Innovation, Systems and Technologies*, vol. 19, pp. 41-50, 2013.
- [26] H. Sun, W. Wang, Y. Cao, S. He, and X. Yan, "Application of fuzzy data fusion in multi-sensor environment monitor," in *Proceedings of the 2009 International Conference on Computational Intelligence and Security (CIS '09)*, pp. 550-553, December 2009.
- [27] Q. Wang, H. Liao, K. Wang, and Y. Sang, "A Variable Weight Based Fuzzy Data Fusion Algorithm for WSN," in *Ubiquitous Intelligence and Computing*, vol. 6905 of *Lecture Notes in Computer Science*, pp. 490-502, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [28] X. Hu and X. Wang, "Application of fuzzy data fusion in multi-sensor fire monitoring," in *Proceedings of the 2012 International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA '12)*, vol. 1, pp. 157-159, August 2012.
- [29] J. Shell, S. Coupland, and E. Goodyer, "Fuzzy data fusion for fault detection in wireless sensor networks," in *Proceedings of the 2010 UK Workshop on Computational Intelligence (UKCI '10)*, pp. 1-6, September 2010.
- [30] Y. N. Wang, J. F. Ye, G. J. Xu, Q. M. Chen, H. Y. Li, and X. R. Liu, "Novel hierarchical fault diagnosis approach for smart power grid with information fusion of multi-data resources based on fuzzy petri net," in *Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '14)*, pp. 1183-1189, July 2014.
- [31] F. Fatemipour, M.-R. Akbarzadeh-T, and R. Ghasempour, "A new fuzzy approach for multi-source decision fusion," in *Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '14)*, pp. 2238-2243, July 2014.
- [32] V. Pandey, A. Kaur, and N. Chand, "A review on data aggregation techniques in wireless sensor network," *Journal of Electronic and Electrical Engineering*, vol. 1, no. 2, pp. 1-8, 2010.
- [33] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924-935, 2002.
- [34] Z. Zhou, M. Xiao, L. Liu, Y. Chen, and J. Lv, "An improved DV-HOP localization algorithm," in *Proceedings of the 2009 2nd International Symposium on Information Science and Engineering (ISISE '09)*, pp. 598-602, December 2009.
- [35] K. Vaidyanathan, S. Sur, S. Narravula, and P. Sinha, "Data aggregation techniques in sensor networks," Tech. Rep. OSU-CISRC-11/04-TR60, The Ohio State University, 2004.
- [36] S. Chen, H. Bao, X. Zeng, and Y. Yang, "A fire detecting method based on multi-sensor data fusion," in *Proceedings of the System Security and Assurance*, pp. 3775-3780, IEEE International Conference, October 2003.
- [37] H. Preuss and V. Tresp, "Neuro-fuzzy," *Automatisierungstechnische Praxis*, vol. 36, pp. 10-24, 1994.
- [38] V. O. Olunloyo, A. M. Ajofoyinbo, and O. Ibidapo-Obe, "On development of fuzzy controller: the case of gaussian and triangular membership functions," *Journal of Signal and Information Processing*, vol. 2, no. 4, pp. 257-265, 2011.
- [39] M. Mizumoto and K. Tanaka, "Some properties of fuzzy sets of type 2," *Information and Control*, vol. 31, no. 4, pp. 312-340, 1976.
- [40] M. Technology, "Pic24fj256gb110 family data sheet," <http://www.microchip.com>.
- [41] "Mrf24j40mb data sheet," <http://ww1.microchip.com/downloads/en/DeviceDoc/70599B.pdf>.
- [42] SENSIRION, "Sht21 humidity/temperature sensor," <http://www.farnell.com/datasheets/1780639.pdf>.
- [43] FIGARO, "Electrochemical carbon monoxide (co) sensor tgs 5042," <http://www.figaro.co.jp/en/product/entry/tgs5042-b00.html>.
- [44] M. Collotta, G. Scatà, S. Tirrito, R. Ferrero, and M. Rebaudengo, "A parallel fuzzy scheme to improve power consumption management in Wireless Sensor Networks," in *Proceedings of the 19th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA '14)*, pp. 1-4, September 2014.

