

Smart-City-based Data Fusion Algorithm for Internet of Things

Jawad Khan¹, Muhammad Amir Khan², N. Z. Jhanjhi^{3,*}, Mamoona Humayun⁴ and Abdullah Alourani⁵

¹Department of Computer Science, IIC University of Technology, Phnom Penh, Cambodia

²Department of Computer Science, COMSATS University Islamabad, Abbottabad Campus, 22060, Pakistan

³School of Computer Science and Engineering, SCE, Taylor's University, Subang Jaya, Selangor, Malaysia

⁴Department of Information Systems, College of Computer and Information Sciences, Jouf University, al-Jouf, KSA

⁵Department of Computer Science and Information, College of Science in Zulfi, Majmaah University, Al-Majmaah, 11952, Saudi Arabia

*Corresponding Author: N. Z. Jhanjhi. Email: noorzaman.jhanjhi@taylors.edu.my

Received: 01 January 2022; Accepted: 23 March 2022

Abstract: Increasingly, Wireless Sensor Networks (WSNs) are contributing enormous amounts of data. Since the recent deployments of wireless sensor networks in Smart City infrastructures, significant volumes of data have been produced every day in several domains ranging from the environment to the healthcare system to transportation. Using wireless sensor nodes, a Smart City environment may now be shown for the benefit of residents. The Smart City delivers intelligent infrastructure and a stimulating environment to citizens of the Smart Society, including the elderly and others. Weak, Quality of Service (QoS) and poor data performance are common problems in WSNs, caused by the data fusion method, where a small amount of bad data can significantly impact the total fusion outcome. In our proposed research, a WSN multi-sensor data fusion technique employing fuzzy logic for event detection. Using the new proposed Algorithm, sensor nodes will collect less repeated data, and redundant data will be used to increase the data's overall reliability. The network's fusion delay problem is investigated, and a minimum fusion delay approach is provided based on the nodes' fusion waiting time. The proposed algorithm performs well in fusion, according to the results of the experiment. As a result of these discoveries, It is concluded that the algorithm describe here is effective and dependable instrument with a wide range of applications.

Keywords: Internet of things; data fusion; wireless sensor network; fuzzy inference system

1 Introduction

The Internet of Things (IoT) intends to create a future in which things in the physical world and cyberspace can be connected and integrated. There are high expectations for the future of this developing technology, which is predicted to connect everything (including physical objects and as well as automobiles, vehicles, appliances, etc.) and allow data to flow between them. The Internet of Things (IoT) is critical in various real-world control systems and remote perception. It has penetrated



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To make our daily lives easier, the Internet of Things (IoT) connects billions of sensors and gadgets in real-time. Globally, 43 billion gadgets are expected to be connected by 2022. Data fusion and analytics are used to increase a well understanding of the surroundings by using a variety of sensors. IoT data differs from other types of data in that it contains unique properties (e.g., trust recommendation in pervasive social networks). Multi-phase developments of IoT can be shown as in Fig. 2.

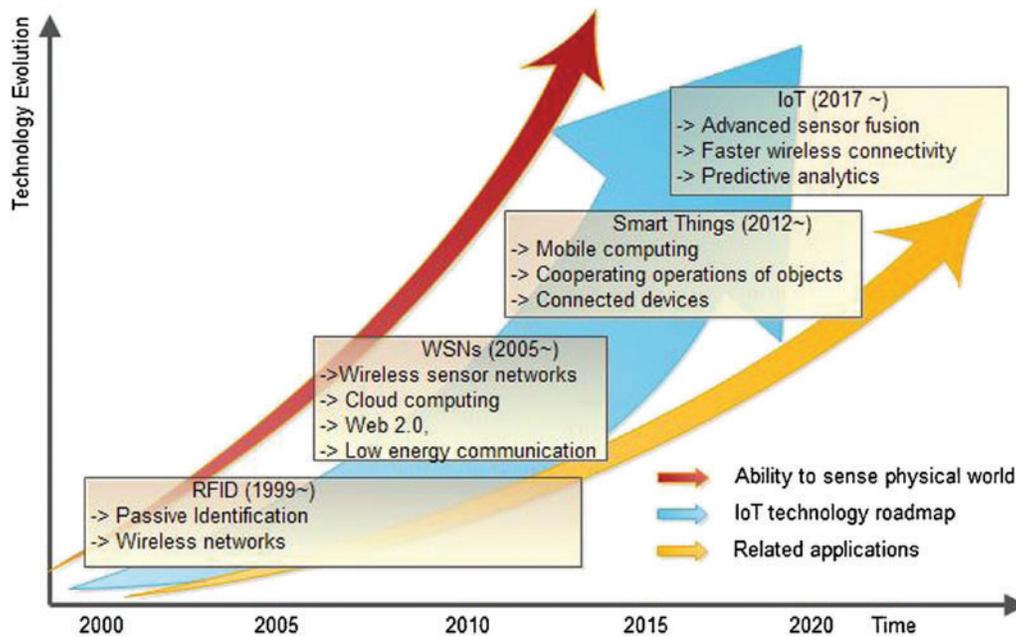


Figure 2: Multiphases of IoT [6]

Nowadays, communication technology is the backbone of each IoT application because it offers an easily transferable data channel for applications. We know in IoT, there exists a lot number of sensor nodes that are used for communication with each other. IoT is deployed in an unattended environment, and also we know that when the sensor node is applied, battery replacement is nearly impossible. Therefore, the main problem in IOTs is how to perform complicated data monitoring, data accuracy, data aggregation delays, network lifetime maximization, and reporting in energy-limited sensor nodes. The existing arrangements demonstrate that the energy consumption transmission between sensor nodes is far higher than the computer energy consumption.

Raw sensor data from IoT sensors can be seen to contain enormous amounts of messy and meaningless data. As a result, raw sensor data must be cleaned before data investigation can be achieved to extract important facts and figure from the prepared IoT sensor. Furthermore, in a regional IoT sensor network, a huge amount of unwanted and worthless data might result in high computing costs and resource overutilization. Data outlier detection, data denoising, aggregation, and imputation are the most frequent data processing techniques. It is therefore very important to develop a structure to condense the transmission of data between sensor nodes. Due to the potential that sensors collect data with a better cover and power efficiency of communication platforms. With large quantities of data in an IoT, data recovery techniques are readily available to translate and therefore degrade the network's energy consumption.

There is only one way to raise the quality of the data fusion result, and that is by using multiple sensors to expand the width of the information collected. Because of this, a wide variety of sensors are typically employed to detect and perceive various types of information (such as infrared data, environmental data, medical data, etc.). The problem with disseminating all data is that it wastes network bandwidth and electricity due to multi-source heterogeneity and a large amount of sensory data [7]. As a result, data fusion becomes a key technique for extracting vital information from a wide range of sensory and gathered data to increase data quality and make decision-making easier.

Data fusion has several definitions. It is defined as “the study of successful ways for translating data from various sources and points in time into a depiction that enables human or automated decision making” [8]. Data fusion can minimize data dimension and size, optimize data volume, and make sense of it. It helps remove data imperfections and collaboration of data from many sensors. IoT data fusion has been extensively studied recently. But it still has drawbacks like privacy leaking and power usage.

Data Fusion creates a single data from multiple sources. The merged data contains more event information than any single event source. Sensors’ features are merged to provide more precise data. Adding analogous and complementary information improves the information’s reliability and overall detail. The primary objective of WSN’s data fusion mechanisms is to increase QoS to ensure reliable and accurate decision-making about interest events. The QoS here means that precise, complete, and reliable information is delivered reliably. Fusing data ensures that it increases the data quality of the IoT sensor network and reduces the consumption of energy by removing redundant information.

To reduce energy use in IoT sensor networks, there are many data fusion mechanisms. These mechanisms are centered on Dempster-Shafer’s evidence theory, rough set theory, probability theory, and possibility theory. In the fusion process, the majority of these approaches can remove duplicate information. These methods, on the other hand, do not account for the limits of certain sensor devices. They believe, for example, that sensor nodes always function well and produce accurate data. These are implausible assumptions that have the potential to alter the environment. The precision of the sensor node’s functions, for example, can be affected by temperature. The existing approaches also transfer sensitive data to a processing hub, both necessary and unnecessary, resulting in excessive energy consumption. While the authors’ methods to data fusion performance and energy economy are intriguing and promising, it is evident that further clever and novel solutions can be created to improve the entire data fusion process, and particularly energy efficiency.

In the modeling of inaccurate data, fuzzy set theory provides an appropriate tool. Many research activities have also been using soft computing techniques in data fusion applications in recent years. Considering that the Fuzzy Inference Engine or System (FIE) input and output values are true non-linear variables, they can be used in many technical problems, especially combined problems, in which traditional control methods do not achieve relatively good results Input and output values.

An algorithm using the fuzzy rule-based system is therefore proposed in this approach. Uncertain quality information is handled by including fuzzy techniques, fuzzy logic, fuzzy set theory, arithmetic and mathematics, fuzzy control, and fuzzy decision-making. Rule-based fuzzy operators are a new type of operator designed specifically to apply the approximate reasoning principles.

In the proposed data fusion process, the actual data values are distinguished from measuring data values by the Fuzzy Inference Engine FIE in each IoT sensor. Cluster heads subsequently combine information [9]. You can also estimate the probability of an outcome in the supervised environment. Further data will be deleted during this process, which will help mitigate energy consumption. The main concept of the proposed solution is to improve the QoS system performance and reduce

electricity consumption by simply conveying the determined chance of procedures rather than all fused data.

In the short following contribution are made:

- a) Suggest a data fusion technique to expand the performance of IoT sensors.
- b) Further data are deleted, and the energy consumption is consequently minimized.
- c) The proposed approach is powerful regarding events of interest concerning the node failures of the sensor
- d) We evaluated the proposed facts fusion technique's effectiveness to a baseline approach.

The following paper is arranged. The analysis of existing literature approaches is presented in Section 2. The system requirement and data fusion frame work is introduced in Section 3. Proposed data fusion methods is described in Section 4, and finally, Section 5 describe fuzzy control model. Results and analysis have been describe in Section 6.

2 Related Works

For too many applications, the large burden of transmission has arisen because of too much data generation, and data fusion is increasingly necessary to reduce redundant data in the routing design of WSNs. Multiple features can classify the approaches to data fusion. Indeed, the fusion of data can be divided into valuable data, redundant data, and relevant data based on similarities between input data. All procedures are classified into two categories: structure-free and structure-based. The fuzzy-based and structure-based/free data fusion strategies proposed in the literature are briefly reviewed in the following paragraphs.

Numerous literary works show that the use of the Internet of things network is exceptionally interconnections between energy utilization and assets. A huge amount of redundant data are spawned by the Internet of Things applications such as surveillance and recognition. Many methods have been employed with data regrouping to filter redundant sensing data. All these methods can be categorized as tree-based, grid-based, and structure-free approaches based on clusters. The hybrid clustering of data discusses a methodology that blends static and dynamic clustering methods [10].

In cluster Wireless Sensor Networks (WSN), an event detection technique using the fuzzy Artificial Intelligence (AI) logical branch was suggested in [11]. Environmental monitoring is one of the main applications of WSNs in the broader context of IoT-based smart companies. In this method, the author proposed multiple sensors for sensing different parameters. Further show that the fusion of signals is executed by a system centered on fuzzy rules. In the cluster heads, signals are received, and the fusion process is then performed. Multi-sensor fusion has the advantage of enhancing accuracy and reliability. Results also indicate that the suggested protocol can be used in a huge WSN format but that energy utilization is significant.

For data fusion, a hybrid class of Kalman and fuzzy filter systems are used efficiently. One of these studies [12] proposes fuzzy logic-based adaptive Kalman filter (FL-AKF) used Kalman filter. The noise covariance matrices measurement is changed to appropriate noise profile numbers for incoming data for all local FL-AKF. A fuzzy adjustment inference system is exercised. Using fuzzy logic can help manage inaccurate data. The Kalman filter improves the accuracy of estimations by filtering the covariance matrix. The technique obtained is, therefore, more precise. Additionally, for fusion in data for autonomous vehicle guidance, the integration of fuzzy logic and Kalman filter is being developed [13]. In this technique, there is a lack of adaptability that limits this technique.

Data fusion technique in [14] provide a multi-sensor environment monitoring system based on a fuzzy data fusion evaluation method. In his study, he used different variations like humidity, ventilation, temperature, and other parameter composed by multisensory and later fused. The system suggested by the authors also increases its intelligence. The results of simulation show that the performance of this technique, e.g., by single sensors, overcomes that of conventional diagnostic methods. The authors nevertheless do not address the problem of sensor nodes' consumption of energy or how their approach is implemented in real devices.

In [15], a hybrid clustering-based data aggregation methodology is discussed that simultaneously combines dynamic and static clustering algorithms. In his proposed method, the algorithm selects an appropriate clustering way based upon network status and also noted the whole data collected from each node is the real worth. Obviously, this supposition reduces the WSN packet transmission ratio.

To increase the lifetime of the IoT network, the node placement strategy and routing mechanism have been contested and efficiently merged into one network architecture. Environmental data is collected by sensor knots, and data aggregation and path calculation are done for the energy-efficient and reliable path computation to limit the number of retransmissions in this suggested network architecture [16].

A new approach has been used in [17], who develop an energy dispersion system at the user level in smart homes using stacking denoise encoders, have considered the problem of intelligent energy consumption. They use power data for each device and the overall power consumed at home to disperse it to a local cloud. Data is then broken down to provide the user with feedback for various devices.

3 Data Fusion Framework for Event Detection

One practical scheme for WSNs in which sensors form clusters based on certain predefined criteria is clustering. Each cluster's cluster head selects a sensor node and is responsible for transmitting data to a base station (sink node). The clustering divides a large coverage range of a WSN between several clusters and is therefore much smaller than before the cluster in each of the sensor nodes. If the same sensor sends data straight to a base station, the amount of power used to transmit data from the sensor node to the cluster head is much less than that required. To improve network performance, clustering also allows for load balancing in each cluster. In the IoT, data fusion can be used to bypass and save heavy traffic. The transmission to the base station of redundant data packets from each sensor node causes these difficulties. Therefore, a data fusion scheme can be adopted in every node of a generic WSN to aggregate the data collected from the other sensors.

3.1 Energy Model

We assume in this proposed paper that the sink nodes and base station are powerful enough. Compound operations are performed from sink nodes, such as data fusion. These nodes are powered by three states: idle, transmitting, and receiving. Their energy consumption is derived. For sensor nodes, energy consumption is not considered because data deletion is simply redundant. The consumption of energy in idle is insignificant; spending energy (E_{TX}) and receiving energy (E_{RX}) from the transmission of an n-bit message to a d-distance are deliberate as follow:

$$E_{TX}(n, d) = n.E_{elec} + n.d^k.\epsilon_{amp} \quad (1)$$

$$E_{RX}(n) = n.E_{elec} \quad (2)$$

where E_{elec} refers to each transmitter or recipient's energy consumption, where ε_{amp} denotes the amplifier consumes one square meter per bit of energy, and k represents the circulation attenuation index [18]. The formula is as follows, taking into consideration the energy usage of data fusion:

$$E_R = k(E_{elec} + E_{DF}) \quad (3)$$

E_{DF} is the energy usage calculated after each data bit has been fused? In the ideal fusing state, data fusion nodes can fuse n input data packets of similar duration into an output packet while saving $(n - 1)/n$ energy.

In WSN, cluster heads are used as sink nodes, and the other cluster nodes consist of children's communications nodes. The clustering ends when the network topology ends. The cluster-head nodes have enough ability to be entertained as sink nodes to facilitate the studies on data fusion, so we need no recurring cluster-head node selections. The data acquired by the sensor nodes for the cluster's offspring is routinely integrated and sent to their parental sink nodes. All the sensor nodes in every cluster should be contained so that the whole cluster is covered with potential happening attributes and the abstract of communications of like sensor nodes in every cluster. Humidity, Temperature, smoke density, light intensity, Carbon Monoxide, and oxygen are used to collect data and disperse it to Cluster Head (CH). The CHs shall merge and transmit the data to the Base Station. Every node then routines Fuzzy Logic Controller (FLC) to assign weight to its data based on its current situation. The data is then forwarded to the wirelessly connected CH of a BS.

To its huge surface area, the Base Station (BS) is typically moveable, necessitating the use of a distinct protocol to control its movement. However, because this study aims to address data fusion, the base station must be assumed to be stationary to enable further research and to stress the research focus. All discharge nodes and sensor nodes on this network system can connect directly with the base station. Sink nodes can connect with the base station either directly or via many hops. Sink and sensor nodes in the same cluster can communicate directly and across multiple hops, and sensor nodes in the same cluster can communicate as well. We know that communication is the main energy usage of sensor nodes. When sensor nodes submit data to the sink node as early as it receives it, sensor nodes in regions with frequent event changes have a shorter life cycle since monitoring values are supplied more frequently. In this paper, therefore, proposes an algorithm based on fuzzy logic to resolve the above-mentioned issues, and the data fusion algorithm principle with the IoT network is illustrated.

A comprehensive block diagram is shown in Fig. 3 concerning the data fusion mechanism introduced in this work. The sensor nodes S_1, S_2, \dots, S_n collects data from the environment examined. The data received is then transmitted via a point-to-point wireless connection to a fusion node. This node then produces from the received inputs one individual internal representation of the controlled environment. This unique representation is then used by the base station. The base station has not generally had access to the measurement of each sensor.

The key reason why a data aggregation mechanism harms its performance is the misbehavior of sensor nodes. The behavior of a sensor node may be influenced by different environmental factors, such as sudden humidity or temperature shifts. These environmental aspects lower or grow the sensor output, which creates ultra-low frequency noise within the transmitted signal. Furthermore, non-operating environmental confines such as low or high air temperatures neighboring the sensor nodes generally affect the performance of the sensor.

The quality of the data is the main requirement in the whole scenario. It is important to ensure that readings and sensor measurements are true values to improve QoS in IoT networks. Data redundancy is also an extra key condition. Explicitly, redundancy is defined in WSNs, for example, by using

redundant data to describe the events. The removal of duplicate data, by contrast, increases the total quality of attained data and reduces the number of spreads data to the base station from source nodes. The fuzzy logic solution presented in this paper is, however, very suitable for applications without rigorous real-time restrictions.

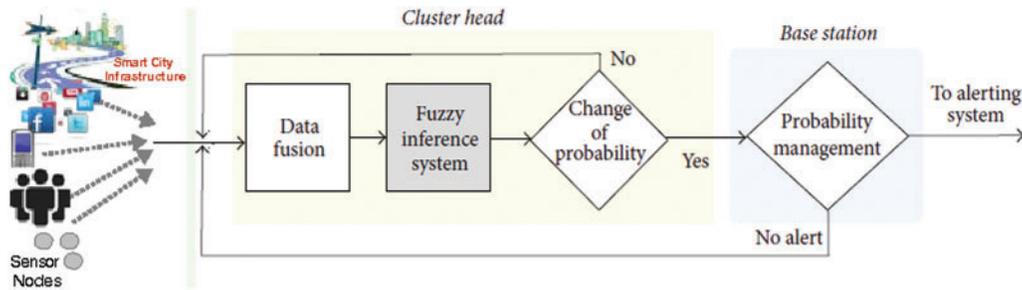


Figure 3: Proposed scheme data fusion mechanism

Let R be the percentage that BS has received of the correct data. The proportion of data packets received by the BS is shown under Eq. (3) as:

$$R = \frac{V_p}{L} \times 100 \quad (4)$$

Here in the above equation L means the total quantity of app-defined events and V_p is the total quantity of messages that contain solely the accurate value of the data that was received. It is necessary to state that general environmental monitoring is the application framework of the solution presented in this paper. In the proposed smart city environment, different physical measures are collected, such as humidity, temperature, light intensity, smoke density, carbon monoxide, and oxygen. Fig. 3 depicts the scheme for the proposed solution for data fusion.

3.2 Basic Data Fusion Model

In 1965, Lotfi Zadeh stated the first Fuzzy Logic System (FLS), which in contrast to logic variables, introduces the concept of linguistic variables. FLS can have a range of approximate values between 0 and 1 in a continuous value. Fuzzification, interference, and defuzzification are the basic structure of FLS. As in fuzzification, crisp inputs are changed into fuzzy inputs. A Member Function (MF) shows the truth of every output and input value. MF has numbers 0 to 1 over a crisp variable interval. Triangular, trapezoidal, and bell curves are the most commonly used for member functions. Each MF consists of different variables depending on the number of variables predefined for each input/output. The five inputs that we have used in the proposed smart city IoT application are temperature (t), humidity (h), light intensity (li), smoke density (sd), carbon monoxide (cm), and oxygen (o). MF has five variables for the output, which is the probability of a smart city environment: very low (VL), low (L), medium (M), high (H), and very high (VH).

4 Proposed Data Fusion Method

The recommended data fusion approach is explained in Algorithm 1. This method collects data from the five sensors till the events are detected. The method takes nodes A, B, C, D, \E as inputs. The sensory data that has been fused is the output of the required input.

Algorithm 1: IOT based Data fusion Algorithm

INPUT: (A: temperature, B: humidity, C: light intensity, C: smoke density, D: carbon monoxide, and E: oxygen)
OUTPUT: Fused Data
START
WHILE (Event NOT Sensed)
 FOR all the cluster members in one kind
 $CI_n \leftarrow FLC(A, B, C, D, E)$
 IF ($CI \geq T \ \&\& \geq 0$)
 The integrated Data will be referred to Cluster Head
 ELSE
 Data will be not consider
 END IF
 END FOR
 A minimal-delay fusion path approach is proposed.
 $Tw(i)$: Set waiting time for each fused node
 Resulting $\leftarrow FIS(Tw_1, Tw_2, Tw_3 \dots Tw_n)$
 IF resulting not changed
 Disregards the data fusion waiting time
 ELSE
 Send the result to BS
 END IF
END WHILE
STOP

In Algorithm 1, $T = nt$ represents the time it takes for the sink node to send messages to the BS and the sensor node to send messages back to the sink node. In addition, two scenarios can be employed instead of delivering signals to the sink node every t seconds from each sensor node in the cluster: First, The sensor node is woken up by an event in the monitored region, and it delivers the data it has collected to the sink node.; second, data collected by the sensor node changes during the event occurrence process, differing from the data shifted previously period t , so the data must be relocated again; otherwise, it will not be shifted again. Longer sensor node life and shorter network longevity are both possible with this technique. Because each WSN node has FLC implemented, it is possible to tell the difference between measured and true data. The data is then merged by cluster heads. FLC is used to calculate the range of a node's confidence interval.

Each attribute value's interval values will be calculated by the sink node, who will then divide them into ten equal parts and assign the ten intervals the values 0–0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, 0.4–0.5, 0.5–0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, 0.9–1.0 correspondingly. To finish the standardization and integration process, obtained data will be assigned an interval value based on its size. To minimize data delay in the network, a fusion path approach with the least delay is proposed. In this study, other cluster members, in addition to the cluster heads, can become fusion nodes. The time between the commencement of data fusion and the end of fusion node i in the event area is known as the data fusion waiting time.

$$T_w(i) = \partial \frac{di}{ki} \in N \quad (5)$$

where $T_w(i)$ = waiting time fusion, ∂ waiting time coefficient of data fusion, di denotes the degree of relation nodes and ki denotes the hop between node i and cluster heads, and N represents the number of members in clusters. In our proposed algorithm, we have calculated the waiting time of each fused node as used in the DFCTA algorithm [19].

Algorithm 1 improves the overall stability of the network while also lengthening the network's lifespan. This is due to some factors, the most significant of which are: During the waiting time between each fused node, the BS distributes a huge number of data fusion jobs among the sink nodes. This allows the sink nodes to work on challenging data fusion tasks simultaneously. In this approach, time is saved, and the overall real-time performance of the system is improved, as well as the prediction results are made available to the terminal staff in a timely manner. A wide monitoring region is common for IoT network systems, with expected events most likely to occur in a certain location. According to this study, regional monitoring improves the prediction system's accuracy and allows people to take more targeted defense measures by making it easier to determine which area has the greater probability that an event will occur at a specific time (such as when a section of the building is most expected to catch fire). Finally, the communication durations of the entire network are reduced, extending the network's life cycle because the sensor and sink node times are evenly distributed in this article.

5 The Fuzzy Control Method

Earlier, we established that fuzzy control models have the principles of crisp input, fuzzy model rules, model fuzzification, a model inference engine, model defuzzification, and finally, assessment. In our proposed scheme, as illustrated in Fig. 4, the fuzzy control model is intended to simulate fire intensity in our smart city IoT scheme and then decide on the right fire probability based on a set of specified parameters: temperature, humidity, light intensity, smoke density, Carbon Monoxide, and oxygen. Using a fuzzy control model for initial fire recognition, we distinct six multisensory input parameters and their parallel fuzzy membership values, such as

- I: temperature variation (DT) = very low;
- II: humidity variation (DH) = dry;
- III: smoke intensity variation or carbon monoxide (DCO);
- IV: smoke intensity variation or carbon monoxide concentration (SI) = low; and
- V: smoke intensity variation or carbon monoxide flame existence is defined as "true" in this study.

Member function for each input crisp is defined. The proposed fuzzy-based approximation model uses three types of Mamdani membership functions. Among them are the Gaussian (gauss), triangular (trimf), and trapezoidal (trapmf) MFs. In [20] explains how to use MF in model design. After following the fuzzification process, a microcontroller unit is used to process the fuzzier values (MCU). A rules inference engine is included in this component. All inference estimation rules that will be used for the suggested fuzzy-based model are included here. The database engine serves as a fuzzy associative matrix (FAM) storage for all inference rules that will be used by the recommended model, resulting in the creation of a knowledge base. We then establish the minimum chance of a fire outbreak using the stated fuzzy domain and the "AND" operator, and the IF::: THEN structure. The weighted function ($W = 1$), which will be determined by FIS, will denote that all evaluation inference rules are equivalent and, as a result, all evaluation inference rules receive equal precedence in the estimation of the fuzzy inference system.

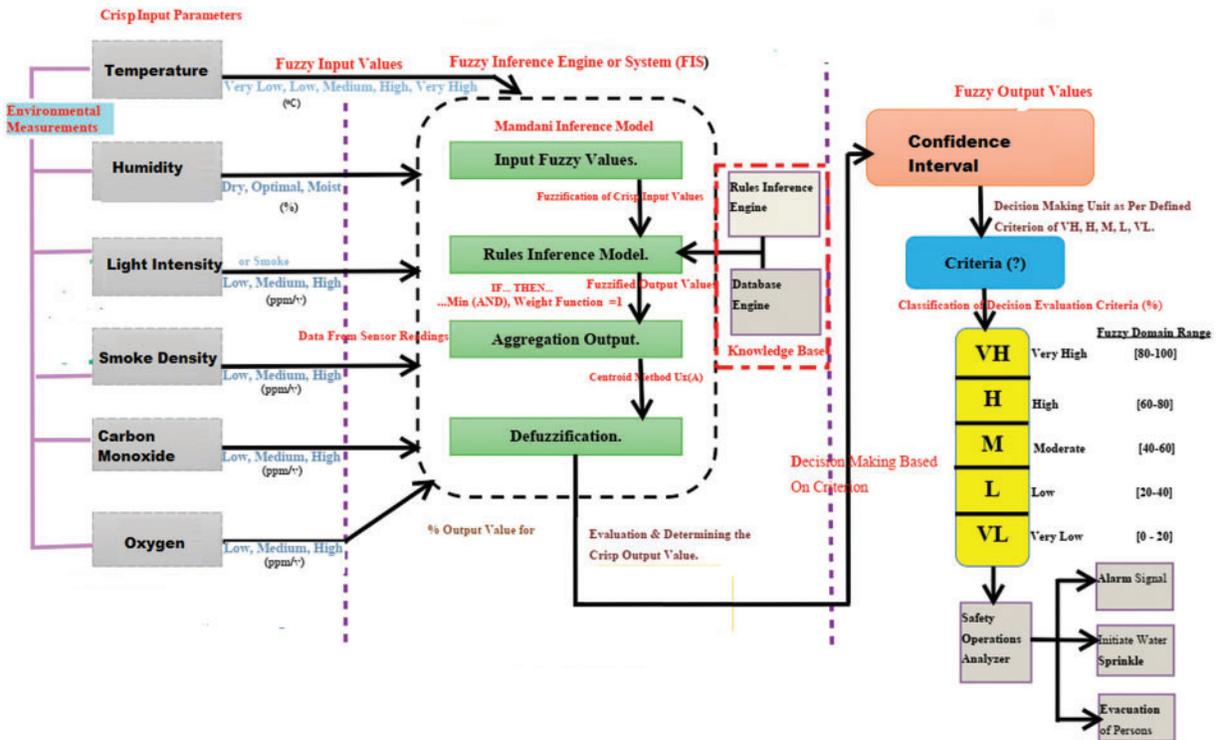


Figure 4: Proposed IOT based fuzzy approximation model

The Center of Gravity (CoG) technique is used for defuzzification, with membership functions of parameters distinct as a triangular shape. Because the output parameter dictates how much load reduction should be made, the defuzzified value specifies how much a value an initialized set-point will alter. The study’s goal is to save energy and maintain the best indoor temperature, however, user comfort was sacrificed.

6 Results and Analysis

Data Fusion Method in Wireless Sensor Networks (DFM) [21] and Data Fusion in Wireless Sensor Networks using Fuzzy Set Theory (DFST) [22] and A New Data Fusion Algorithm for Wireless Sensor Networks Inspired by Hesitant Fuzzy Entropy (HFE) [23] are three methods that are compared in this study (HFE). Fuzzy theory was used in wireless sensor network data fusion, and fire prediction analysis was performed using these three techniques, which is quite similar to what we do. To increase the quality and accuracy of a wireless sensor network’s network service while using less energy, DFM uses a fuzzy theory-based data fusion technique. On the other hand, this approach only differentiates and aggregates the authenticity of gathered data and doesn’t go any farther; hence it falls short of the core performance of the article. For the most part, the DFST method makes use of fuzzy set theory to combine data from numerous sensors on the cluster level. It is adept at dealing with some of the problems associated with reading environmental data, such as inaccuracy and dispute. Without filtering the data collected before processing, we are likely to accept a large amount of incorrect data. The algorithm’s precision cannot be guaranteed at the same time. While this is convenient for processing, it has a major impact on accuracy since it converts the data value from

the same interval to the interval's median value. In HFE, the amount of data acquired by sensor nodes can be reduced, and redundant data can be more effectively used in order to increase the reliability of the gathered info.

In terms of the time it takes for a network to operate, the three-based method is contrasted. The three strategies are associated to the lengthening of network operation times, as seen in Fig. 5. The rise in network operation times, as shown in Fig. 5, is mostly due to a variation in the number of lasting surviving sensor nodes in the network. The three techniques include a total of 800 sensor nodes, which are divided evenly based on the algorithm's node kinds. The number of sensor nodes lowers to 150 when the algorithm reaches 550 rounds, as seen in Fig. 5. In part, this is due to the fact that the sensor used in this experiment detects smoke concentration as well as temperature and humidity levels in the air, but the far-infrared sensor node, which was primarily used to determine the likelihood of a fire arising, survives the experiment. However, the distributed DFFST, DFM, and HFE networks approach the node that dies before succession before they can be succeeded by another. When it comes to the uniformity of network energy use, our solutions outperform the current two leading algorithms. This effect is achieved by the data fusion algorithm by regularly integrating and passing non-repeated data to the sink nodes. Data transmission times, redundant data transmission, and sensor node energy consumption are all reduced as a result. Data fusion efficiency is improved while redundant data selection is eliminated by sink nodes and base stations. As a result, the network's life cycle and sensor node survival time have both increased.

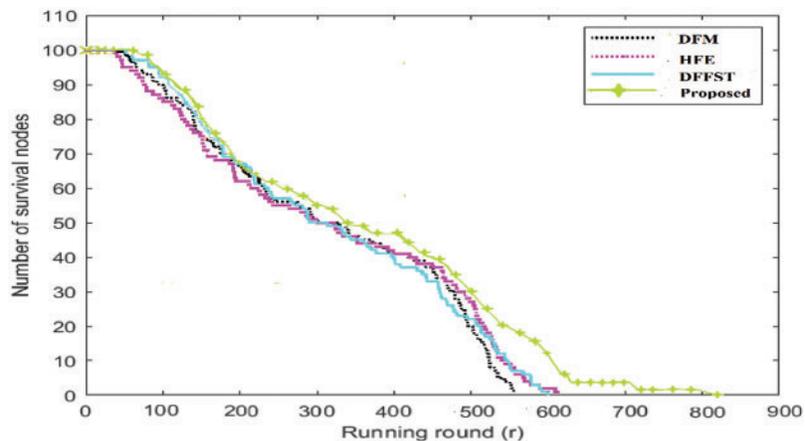


Figure 5: Amount of networks run and surviving sensor nodes

Fig. 5 shows the percentage of correct data gathered by BS, as well as DFM, DFFST, HFE, and proposed, using the proposed approach. In comparison to the previous ways, the anticipated approach delivers a better percentage of valid data, as seen in the figure. Because the suggested method eliminates false data, the true value data is not corrupted throughout the fusion process.

Fig. 6 clearly shows that the suggested data fusion algorithm's total residual energy of sensor nodes is much larger than the other three techniques. Furthermore, the suggested algorithm's overall energy usage is proportional to the network operating times. The proposed method, it is explained, can conserve energy and assure the stability of the entire network's energy usage.

Fig. 7 primarily depicts the link between the network's total residual energy and the number of network runs. Fig. 7 clearly shows that the suggested data fusion algorithm's total residual energy of sensor nodes is much larger than the other three techniques. The suggested algorithm's primary goal

is to decrease duplicate data transmission and thereby reduce energy usage. In addition, the sink node completes data fusion. Because the sink node (cluster head) is fixed in this article, the energy used to pick the cluster head is decreased. Due to the death of sensor nodes in the later stages of network operation, only a few sensor nodes are active, resulting in a reduction in overall energy usage.

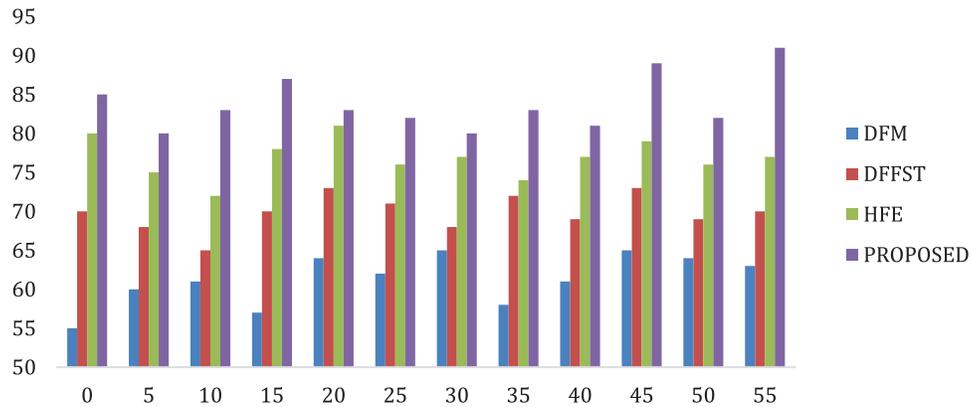


Figure 6: Data collection

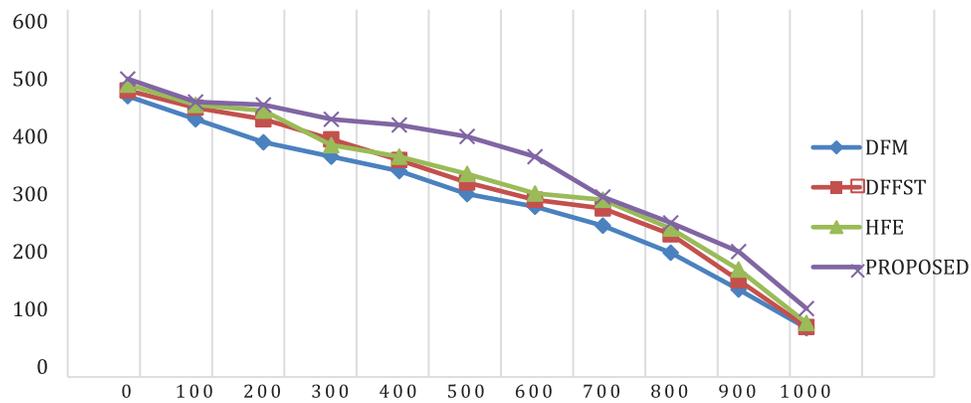


Figure 7: Relationship between total network runs and remaining energy

7 Conclusion

Our work investigates data fusion algorithms in the smart city Internet of things, analyses and discusses related data fusion techniques and protocols, and presents a new fuzzy data fusion method based on event-driven data. Experiments demonstrate that the proposed method is capable of distinguishing and aggregating only the genuine values of the gathered data and eliminating redundant data, which reduces energy consumption and hence increases the network’s life expectancy. The evaluation of the algorithm reveals that the proposed strategy is effective in integrating data from a variety of different sources. A minimum fusion delay way is also built in order to reduce overall delay time of data fusion in the network. When it comes to data collecting, packet size, and energy usage, the suggested method is compared to three corporate approaches. The experimental results suggest that the proposed method outperforms the baseline methods.

Next, this paper's method will be tested on higher-density networks. Complexity and input factors will be evaluated to enhance the inference engine of the technique and to quantify how it improves both the network life cycle and the quality of service. A neural network capable of forecasting environmental conditions, such as predicting the requirements at different times of the day or on different days of the week, is another goal for future study on the system here discussed.

Acknowledgement: We sincerely acknowledge and thanks IIC University of Technology to give us the opportunity for this research work. Also, would like to thank my Dr. Muhammad Amir, and Prof. Dr. Noor Zaman Jhanjhi for their kind support and help.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] D. Wenxiu, X. Jing, Z. Yan and L. T. Yang, "A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion," *Information Fusion*, vol. 51, no. 1, pp. 129–144, 2019.
- [2] Z. Chunsheng, H. Wang, X. Liu, L. Shu, L. T. Yang *et al.*, "A novel sensory data processing framework to integrate sensor networks with mobile cloud," *IEEE Systems Journal*, vol. 10, no. 3, pp. 1125–1136, 2014.
- [3] C. Yongfeng, Y. Ma, Z. Zhao, Y. Li, W. Liu *et al.*, "Research on data fusion algorithm and anti-collision algorithm based on internet of things," *Future Generation Computer Systems*, vol. 85, pp. 107–115, 2018.
- [4] C. Ningyuan, S. B. Nasir, S. Sen and A. Raychowdhury, "Self-optimizing IoT wireless video sensor node with in-situ data analytics and context-driven energy-aware real-time adaptation," *IEEE Transactions on Circuits and Systems I*, vol. 64, no. 9, pp. 2470–2480, 2017.
- [5] A. Ullah, N. Ishaq, M. Azeem, H. Ashraf, N. Jhanjhi *et al.*, "A survey on continuous object tracking and boundary detection schemes in IoT assisted wireless sensor networks," *IEEE Access*, vol. 9, no. 1, pp. 126324–126336, 2021.
- [6] L. Shancang, L. D. Xu, and S. Zhao, "The internet of things: A survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.
- [7] P. Ferrer-Cid, J. M. Barcelo-Ordinas, J. Garcia-Vidal, A. Ripoll and M. Viana, "Multisensor data fusion calibration in IoT air pollution platforms," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3124–3132, 2020. <https://doi.org/10.1109/JIOT.2020.2965283>.
- [8] B. Henrik, S. F. Andler, M. Brohede, R. Johansson, A. Karlsson *et al.*, "On the definition of information fusion as a field of research," 2007.
- [9] S. Yajuan, H. Wang, K. Zhang, and X. Yang, "Associated clustering strategy for wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 10, no. 4, pp. 817234, 2014.
- [10] J. W. Sung, K. W. Lim, Y. B. Ko and S. J. Park, "Efficient clustering-based data aggregation techniques for wireless sensor networks," *Wireless Networks*, vol. 17, no. 5, pp. 1387–1400, 2011.
- [11] M. Shafiq, H. Ashraf, A. Ullah, M. Masud, M. Azeem *et al.*, "Robust cluster-based routing protocol for IoT-assisted smart devices in WSN," *Computers, Materials and Continua*, vol. 67, no. 3, pp. 3505–3521, 2021.
- [12] P. J. E. Ambrosio and N. Mort, "Hybrid Kalman filter-fuzzy logic adaptive multisensor data fusion architectures," in *Proc. IEEE Int. Conf. on Decision and Control (IEEE Cat. No.03CH37475)*, Maui, HI, USA, pp. 5215–5220, 2003.
- [13] V. Subramanian, T. F. Burks and W. E. Dixon, "Sensor fusion using fuzzy logic enhanced kalman filter for autonomous vehicle guidance in citrus Groves," *Transactions of the ASABE*, vol. 52, no. 5, pp. 1411–1422, 2009.

- [14] S. Hongge, Y. Cao, S. He, and X. Yan, "Application of fuzzy data fusion in multi-sensor environment monitor," in *Proc. Int. Conf. on Computational Intelligence and Security*, Beijing, China, pp. 550–553, 2009.
- [15] J. Anees, H. -C. Zhang, S. Baig, B. G. Lougou and T. G. R. Bona, "Hesitant fuzzy entropy-based opportunistic clustering and data fusion algorithm for heterogeneous wireless sensor networks," *Sensors*, vol. 20, no. 3, pp. 913, 2020.
- [16] P. Sarwesh, N. S. V. Shet and K. Chandrasekaran, "Effective integration of reliable routing mechanism and energy efficient node placement technique for low power IoT networks," *International Journal of Grid and High Performance Computing (IJGHPC)*, vol. 9, no. 4, pp. 6–35, 2017.
- [17] F. C. C. Garcia, C. M. C. Creayla and E. Q. B. Macabebe, "Development of an intelligent system for smart home energy disaggregation using stacked denoising autoencoders," *Procedia Computer Science*, vol. 105, no. 1, pp. 248–255, 2017.
- [18] T. M. Behera, U. C. Samal and S. K. Mohapatra, "Energy efficient modified LEACH protocol for IoT application," *IET Wireless Sensor System*, vol. 8, no. 5, pp. 223–228, 2018.
- [19] W. Zhang, J. Yang, H. Su, M. Kumar and Y. Mao., "Medical data fusion algorithm based on internet of things," *Personal and Ubiquitous Computing*, vol. 22, no. 5, pp. 895–902, 2018.
- [20] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—II," *Information Sciences*, vol. 8, no. 4, pp. 301–357, 1975.
- [21] D. Izadi, J. H. Abawajy, S. Ghanavati and T. Herawan, "A data fusion method in wireless sensor networks," *Sensors*, vol. 15, no. 2, pp. 2964–2979, 2015.
- [22] M. Collotta, G. Pau and A. V. Bobovich, "A fuzzy data fusion solution to enhance the QoS and the energy consumption in wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 17, no. 1, pp. 3418284, 2017.
- [23] J. Wang, O. T. Tawose, L. Jiang and D. Zhao, "A new data fusion algorithm for wireless sensor networks inspired by hesitant fuzzy entropy," *Sensors*, vol. 19, no. 4, pp. 784, 2019.