Computational Intelligence in Electrical Engineering
Vol. 12, No. 4, 2022
Research Paper

Introducing a sensor fusion method using an interval decision template for handling uncertainty in sensor data

Mehdi Salkhordeh Haghighi

Faculty of Computer Engineering and Information Technology, Sadjad University, Mashhad, Iran, IEEE Senior Member
Haghighi@sadjad.ac.ir

Abstract :
Data fusion techniques have been used in decision making applications for many years. The methods introduced for data fusion, synergistically combine different types of data received from different sources or sensors. In this research, a fusion method is introduced for data fusion, synergistically combine different types of data received from different sources or sensors. In this research, a fusion method is introduced that can be used in situations in which the data received from different sources or sensors have some level of uncertainty and are represented as intervals. The main challenge in these situations is how the uncertainty may be represented and handled. The contribution of this paper is in three parts. First, a special tool is presented for the representation of the uncertainties that can be used in the fusion methods that use intervals (named interval decision template). Second, a fusion method is introduced to use the tool as a basic structure. And finally, the decision template and fusion method are combined with some known fusion methods for handling uncertainty. The designed experiments indicate how the interval decision template is used for data fusion and also indicate the effectiveness of the presented fusion method. Moreover, some experiments are designed to indicate the effectiveness of using the interval decision template by Dempster Shafer and Bayes methods to handle interval data.

Keywords: Data Fusion, Interval Data, Decision Fusion, Decision Template, Uncertainty, Dempster Shafer**,** Naïve Bayes

1. Introduction

Data fusion (DF) methods have been used in many applications for decades. When uncertainty exists in the sensor measurements, one way to alleviate the uncertainty is using data fusion. By definition, DF is a synergistic combination of data received from different sources (sensors) to increase the accuracy and reliability of decision making. There are some terminologies related to DF in different applications. To be applicable, these terminologies are unified as a model by Joint Directors of Laboratories (JDL) Data Fusion

[1] Submission date:22,01 , 2021
Acceptance date: 01, 05, 2022
Corresponding author: Mehdi Salkhordeh Haghighi **,** Faculty of computer engineering and information technology, Sadjad University, Mashhad, Iran, IEEE Senior Member

10.22108/ISEE.2022.127131.1449

Working Group in 1986 [1]. The JDL model is very general and may be adapted to many application areas.

In this section, a DF model is defined to indicate the way that multi-sensor DF may handle uncertainty in sensor data. In this model, multi-sensor measurements with different uncertainties may be used by a DF system to increase the accuracy of the measurements and reduce the uncertainty. In a simple way, if s_i^j is the i^{th} measurement of the j^{th} sensor (known as a sample) without any uncertainty (zero error $e_i^j = 0$), and there are k measurements of the same sensor, $\{s_i^j, j = 1, \ldots, k\}$, then, the estimated value of $\bar{s^j}$ that is measured by the j^{th} sensor is $\bar{s^j} = f(s_1^j, \ldots, s_k^j)$, that f is the fusion function. If uncertainty exists in s_i^j, indicated by the error value e_i^j, then, each of the k measurements is an interval defined as $I_i^j = s_i^j \mp e_i^j$. In this case, a different fusion function is defined as $\acute{f}(I_1^j, \ldots, I_k^j)$ and used to estimate the measurement of the sensor s^j.

In the model described earlier, each measurement is done by a sensor (s_i^j), with or without uncertainty, and a DF function estimates the measured value. In a complex environment, with different types of sensors, each measurement is composed of different values (known as complex measurement). Each component of a complex measurement is measured by an individual sensor. Therefore, each complex measured value is defined by a vector $\vec{t_i} = (s^1, \ldots, s^n)$ as a test sample, if there are n sensors and each measures a single quantity(s^i). Then, the value of $\vec{t}$ is used to estimate a property of the environment (p) by an estimator $\varphi^k(\vec{t_i})$. A pool of different estimators $\{\varphi^k(\vec{t_i}), k = 1, \ldots, NE\}$ (NE is the number of estimators) with different behaviors is available to estimate the property differently. Each φ^k estimates a different value for the property (p). This redundancy in the estimation of the property (p) is used by a fusion function $f(\varphi^1(\vec{t_i}), \ldots, \varphi^{NE}(\vec{t_i}))$ to estimate the value of p. Diversity in the definition of $\varphi^k. k = 1, \ldots, NE$ should also be considered to improve the efficiency of the fusion function (f) [2].

On the other hand, if each estimator has a level of uncertainty in the estimation of p caused by uncertain sensor measurements $s^i, i = 1, \ldots, n$, the value of each φ^k is an interval $I_i = \varphi^k(\vec{t_i}) \mp e_k, k = 1, \ldots, NE$. Therefore, an interval fusion function $\acute{f}(I_1, \ldots, I_k)$ is used to estimate the value of p. In this manuscript, the contribution is to define the function $\acute{f}$ and some other tools that are necessary for the fusion of interval values. The model of an environment and the DF function described earlier are useful especially in obtaining situational awareness [3] of an environment where different sources of data and sensors are used with different levels of uncertainty. Figure 1 indicates the model of an environment with different types of sensors.

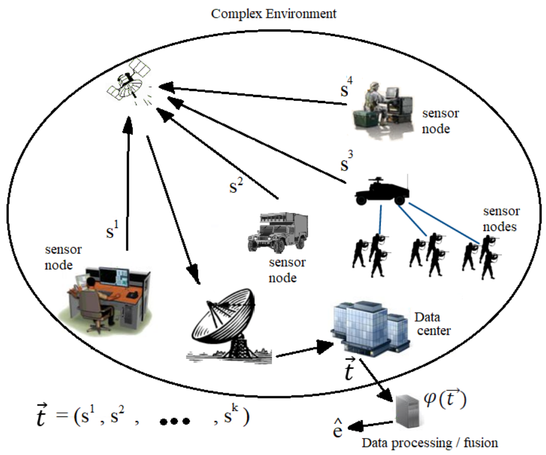


Figure.1.Estimation of the property $\hat{e}$ in an environment with different types of sensor nodes

As indicated in figure 1, each sensor node measures a value that is received by the data center. The data center integrates the measured values and produces a data sample as a vector $(\vec{t})$ that is sent to the data processing server. The sample is mapped by an estimator $\varphi(\vec{t})$ to estimate a property of the environment $(\hat{e})$. The contribution of this manuscript is as follows: first, introducing a DF tool that can be used as a basic structure for DF, second, introducing an interval DF method that can be used as $\varphi(\vec{t})$ in figure 1, and finally, embedding the tool and the method in some of the known fusion methods to make them handle the uncertainty.

The structure of the remaining parts of this manuscript is as follows. In section 2, a review of the DF methods is presented and the aggregation methods used for handling interval data are also reviewed. In section 3, the proposed interval DF method and the tools that are used in the method are introduced. In section 4, details of the experiments are described and the results of the experiments are analyzed.

2. Related works

There are some known DF methods that are used in different applications such as Bayesian, Dempster Shafer [4], Ordered Weighted Averaging [5], and decision template [6]. Some signal level fusion methods are also introduced, e.g. minimum variance estimation, maximum likelihood estimation, random weighting, Kalman filter, particle filter, and random matrix theory based on DF [7] [8]. Some data aggregation methods are also introduced in DF applications [9]. By increasing the use of intelligent systems in decision making applications, and emerging new application areas, more DF methods and tools are also needed.

Recently, multi-sensor DF methods have been used in wide range of applications such as wireless sensor networks [10] [11], internet of things [12] [13], industrial [14], military [15], commercial, biomedical and health [16], navigation and traffic [17] [18] [19].

In a recent work by Yang Dan et al. [20] a Dempster Shafer fusion algorithm is proposed for local track estimation to improve the performance of tracking. In the first step, a filter is used to estimate a local target with unknown measurement noise. Then, a technique is introduced to obtain the estimated tracks of local sensors based on the outputs of the filter. In research by Fuyuan Xiao [21], a method for multi-sensor data fusion is proposed for situations where highly conflicting evidence is fused. The method used a belief divergence measure of evidence and a belief entropy. First, the credibility degree of evidence is obtained to determine the reliability of the evidence. Then, the information volume of the evidence is obtained and the weighted average of the evidence is computed and used by the Dempster combination rule.

In a paper by Claudio M. de Farias et al. [10] a multi-sensor data fusion method is introduced to increase the lifetime in wireless sensor networks. By increasing the number of applications, the data ranges of the sensors overlap and it is more difficult to identify the origin of each data sample to deliver it to the correct application. Their proposed method divides each interval into a set of intervals and assigns each interval to an abstract sensor. Then, correlations among the abstract sensors are identified and used to monitor the behavior of the sensors. In a recent paper, Ronald Yager introduced a multi-criteria aggregation fusion based on ordered weighted averaging aggregations [22]. Some aggregation functions and fuzzy measures are introduced and a multi-criteria decision making by using these measures is also investigated.

In some applications, the data received by the DF system may be defined as intervals. Therefore, some methods are introduced for the fusion of interval data. In a paper by Sergey V. Muravyov et al. [23] a DF method is introduced for interval data. In their method, based on some initial intervals, a fusion method makes an interval including a predefined value with maximum likelihood. Their method produces a ranking over a set of discrete values belonging to the initial intervals. The highest-ranked value is reported as a result of the fusion.

In research by Yanyong Huang et al. [24], a method for the fusion of interval data received from different sources is introduced. The method uses the fuzzy definition of interval data and has the capability of adding new sources of data and removing unnecessary sources. Because multi-source interval datasets were not available in any public databases, they made their synthetic data based on a standard deviation of attributes of each dataset.

In a work by Timothy C. Havens et al. [25] a fuzzy measure is used for information aggregation. They focused on the fuzzy measures that analyze some properties of input data and determine the importance of the input sources. In this paper, a uniqueness measure and additive measure of agreement for interval-valued evidence are used. Then, the fuzzy measure was extended to fuzzy number inputs. Then, interval and fuzzy number evidence are aggregated with the Choquet and Sugeno fuzzy integrals.

In a recent work by Pekala et al. [26] the problem of measuring the degree of inclusion and similarity between interval-valued fuzzy sets is introduced and some similarity measures based on aggregation and uncertainty are used to solve the problem.

In research by Shrinivasan et al. [27] an interval type-2 fuzzy logic decision fusion method is used for air-lane monitoring. In the system, the decision of whether an aircraft is flying along the air-lane or not is determined based on the inputs received from multiple sensors. There is also some other research on fuzzy interval numbers

[27] [28] [29]. Therefore, in the applications where uncertainty is a challenge especially when different sources of data exist, a fusion of interval data may help to design DF systems.

As stated, one of the challenges in designing DF systems is that new applications with different data types and properties have emerged that need new DF methods and tools. Moreover, Uncertainty in the data received from different sensors or sources is another challenge in DF methods. On the other hand, new sources or sensors may be added to the system or removed at any time. To deal with these challenges, the contribution of this paper is introducing a method for the fusion of interval data by using a new DF tool with a wide range of applicability. Details of the method and tool are presented in the next sections.

Many research papers have been published recently in different fields that use multi-sensor fusion methods in sensitive applications. This indicates the importance of multi-sensor fusion. Some examples of these papers are reviewed as follows.

In research by Dash and Jayaraman [30], a probabilistic model for sensor fusion is introduced. Their method is used in a multi-static radar architecture for the detection and tracking of targets. In another research by Zhang et al. [31], a data fusion algorithm is introduced to detect a target using asynchronous data. The data measured by the sensors are changed in the time interval in which a target should be detected because of different sampling rates of the sensors, communication delays, and target movement. The local state estimated information of each sensor and the predicted value of the observation are fused to increase the accuracy of the detection. In the first step, the measured values are used by a weight fusion. Then, global fusion estimation output is produced in a step-by-step filtering fusion process to improve the performance of the fusion.

In research by Sergey Muravyov et al. [32], an Interval Fusion with Preference Aggregation is used to improve the accuracy of measurements in navigation systems. The data from different inertial sensors are used to obtain a single value with minimum uncertainty.

Raphael Voges and Bernardo Wagner [33] in their research used Light Detection and Ranging sensors to provide complementary information about the environment that a mobile robot needs to detect. They used interval analysis to propagate errors from the input sources to the fused information. Initially, camera and Light Detection and Ranging sensors information are changed to a camera coordinate system. Then, visual features, which are detected and tracked, are fused with depth information.

In research by Manogaran et al. [34] the authors used a sensor fusion method in an intelligent transportation system to enhance the results of the fusion method in the connected vehicles. The proposed data Fusion method diminishes errors in assimilating data in different time variations and input sources.

In research by Shen et al. [36] state estimation of a system based on contaminated measurement signals has been investigated. In the research, a survey has been done on the algorithms and fusion methods of multi-rate systems. In another research by Wang et al. [35], the elimination of errors in distributed fusion systems is investigated and a distributed multi-sensor fusion is proposed.

The structure of the remaining parts of the paper is as follows. In section 3, details of the method and the designed tools that are used for the fusion of interval data are presented. In section 4, some experiments are designed to indicate not only the use of the tools and the method for DF but also their effectiveness is compared with some others. Finally, in the conclusion section, the results of this research are summarized and some points for future works are also presented.

3. The proposed method

In this section, a DF method is introduced that can be used in applications that have some form of uncertainty in the received data produced by sensors or other sources. The DF method is presented in four main steps: First, introducing a base method for the fusion of interval data to handle uncertainty. Second, introducing a new decision template named interval decision template (IDT) as a general tool for the fusion of interval data. Third, introducing a new DF method named interval decision template fusion (IDTF) by combining the IDT and the presented method for the fusion of interval data. Fourth, extending some other DF methods to handle this form of uncertainty by using the presented IDT and the interval DF method. In the remaining parts of this section, details of the interval DF method and the structure of the IDT are described and the algorithm of the IDTF is analyzed. Then, the Dempster Shafer and Bayes methods are extended to use the IDTF to handle uncertainty.

Generally, each interval may be defined by a lower and a higher bound indicated by $[L, H]$. Different operations may be defined on interval data as required. One of the basic operations needed for the proposed method is comparing two different interval numbers. To do this, number A is defined by $A = [A_L, A_H]$ and $B = [B_L, B_H]$. Figure 2 indicates the possible cases in which the two numbers are compared. To do this, a similarity measure is introduced by definition 1.

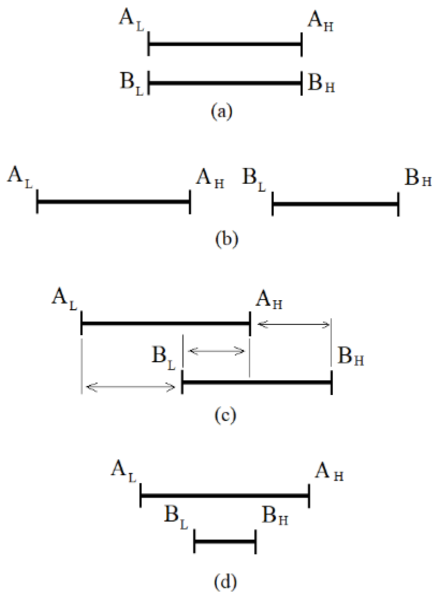


Figure.2. Comparing two different intervals

Definition 1: similarity between two interval numbers (A, B) is defined by a mapping function $S(A, B) \in [0, 1]$ and computed by equation (1). The equation is defined based on the cases indicated in figure 2. When the intervals are completely overlapped, the value of the similarity measure is one (figure 2.a) while, if they have no overlap, this measure is zero (figure 2.b). The similarity of partially overlapped numbers (figures 2.c, 2.d) is computed based on the equation (1).

$$(A, B) = \begin{cases} \dfrac{|A_H - B_L|}{(|B_H - A_H| + |B_L - A_L|)} & ((A_L \leq B_L \; and \; A_H \leq B_H) \; or \\ & (B_L \leq A_L \; and \; B_U \leq A_H)) \; and \\ & (A_L \neq B_L \; or \; A_H \neq B_H) \\ \dfrac{B_H - B_L}{(A_H - B_H + B_L - A_L)} & A_L < B_L \; and \; B_H < A_H \\ \dfrac{A_H - A_L}{(B_H - A_H + A_L - B_L)} & B_L < A_L \; and \; A_H < B_H \\ 0 & A_H \leq B_L \; or \; B_H \leq A_L \\ 1 & A_L = B_L \; and \; A_H = B_H \end{cases} \quad (1)$$

The similarity measure defined by equation (1) is used by the fusion method introduced in this paper and the DT for interval numbers. The next sections describe more details.

3.1. Interval Decision Template Fusion

In this section, a new tool named interval decision template (IDT) is introduced that can be used in the proposed DF method named interval decision template fusion (IDTF). The IDT may also be used in the applications where interval numbers are produced or some form of uncertainty exists in the measurements. Details of the IDTF method are indicated in figure 3.

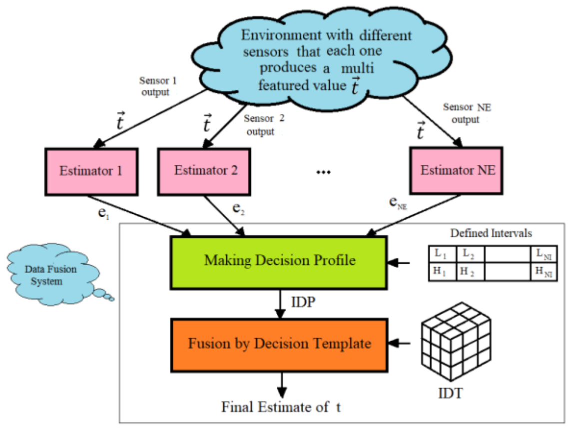


Figure.3. General structure of the DF system

In figure 3, each sensor may be any source of data that produces a real output value indicated as one entry in the vector $\vec{t} = (s^1, s^2, \ldots, s^k)$. All of the entries of the vector $\overrightarrow{IV} = (e_1, e_2, \ldots, e_{NE})$ are different estimates (e_i) of one property of the environment. The fusion of these different estimates produces a more accurate estimate which is the desired output of the IDF system. The IDT in figure 3 is made in the training step based on the training data.

In the IDTF, all entries of the input vector $(\overrightarrow{IV})$ are converted to interval numbers and an interval decision profile (IDP) is produced. Then, a fusion algorithm uses the IDP and the IDT to make a final estimate for the desired quantity. Details of the structure of the IDT and IDP and the algorithm that produces them are indicated in figures 4-7.

In the design of the IDF system, the number of intervals is indicated by NI, and each interval is defined by a lower and a higher bound $[L.H]$ with an interval width (IW) computed by equation (2). The value of NI is determined by the accuracy of sensor outputs. The heuristic method for determining the value of NI is that the more accurate the value of sensor outputs, the higher the value of NI. It is because the higher value of NI produces narrower intervals and increases fusion accuracy. The NE in figure 3 is the number of estimators that produce inputs to the IDF system. The limits of the domain of input values to the IDF system are defined by their minimum and maximum shown by MinI and MaxI, respectively. To define the input intervals, equation (3) is used.

$$IW = (MaxI - MinI)/NI \tag{2}$$

$$interval(i, MinI, MaxI, IW) = \begin{cases} [MinI + (i-1) \times IW, MinI + i \times IW] & 1 \le i < NI \\ [(NI-1) \times IW, MaxI] & i = NI \end{cases} \tag{3}$$

Since each input value received by the IDF system is a real number ($\mathfrak{R}$), a function $INT: (\mathfrak{R}, IW) \to [L, H]$ converts it to an interval number. This function may be defined based on the properties of each input source or sensor (for example, accuracy). In the proposed method, the properties of all the input sources are assumed to be the same and the INT for each value ($\mathfrak{R}$) is defined by equation (4).

$$INT(\mathfrak{R}, IW) = [\mathfrak{R} - \left(\frac{IW}{2}\right), \mathfrak{R} + \left(\frac{IW}{2}\right)] \tag{4}$$

By assuming that the input to each ES is a multi-featured sample$(\vec{t})$, each ES produces an estimated value for the sample and this value is used as an input to the IDF system. If $e_{i,j}$ is the i^{th} estimated value for the j^{th} sample, the input vector to the DF system is defined as $\overrightarrow{IV_j} = (e_{1,j}, \dots, e_{NE,j})$. If predefined intervals are produced for all the ranges of the ES outputs, an interval mapping function $IM: (\mathfrak{R}, Intervals, NI) \to \boldsymbol{N}$ maps each real number $\mathfrak{R}$ produced by each ES to an interval (n). The difference between INT and IM functions is that the former converts a real value to an interval number while the latter maps it into a predefined interval.

As indicated in figure 3, the IDF system uses IDT in the fusion algorithm to produce the final estimate. Making IDT has two steps. In the first step, based on the training samples used for estimators training, an initial decision template (INDT) is made. The structure of the INDT is shown in figure 4.

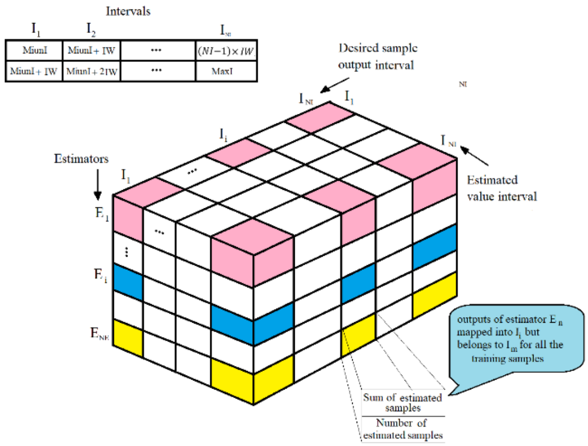


Figure.4. The Structure of INDT

As indicated in figure 4, each entry in the INDT is indicated by the couple (sum, num). The values of 'sum' and 'num' in $INDT(i, j, k)$ are the sum and the number of values produced by E_i for all the training samples and mapped into the interval corresponding to the entry. The intervals in figure 4 $(I_1, \dots, I_{NI})$ are defined by equation (3). In this equation, the values of MinI and MaxI are defined by equation (5). In this equation, NR is the number of training samples.

$$\begin{aligned} MinI &= \min\{e_{i,j} \mid i = 1 \dots NE, j = 1 \dots NR\} \\ MaxI &= \max\{e_{i,j} \mid i = 1 \dots NE, j = 1 \dots NR\} \end{aligned} \tag{5}$$

Therefore, the intervals defined by equation (3) and equations in (5) are shown by 'intervals' that can be used in the algorithm in figure 6 to make INDT.

In the second step, the IDT is made based on the INDT indicated in figure 4. Figure 5 indicates the defined IDT.

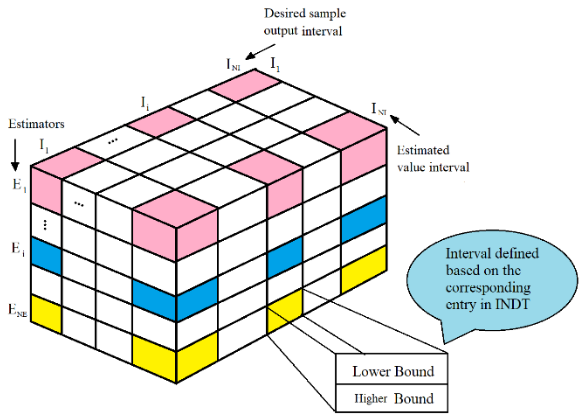


Figure.5. The defined IDT for the fusion of interval data

Each entry in IDT is an interval defined by equation (6) based on the values in the corresponding entry in INDT.

$$IDT(i,j,k) = INT\left(\frac{INDT_{sum(i,j,k)}}{INDT_{num(i,j,k)}}, IW\right) \qquad (6)$$

The algorithm in figure 6 indicates the process of making INDT and IDT.

```
Algorithm making IDT
Inputs: NI, MinI, MaxI, NE, training samples, NR
Outputs: INDT, IDT
// NR is the number of training samples
Define INDT_{NE×NI×NI} ;
Define IDT_{NE×NI×NI} ;
IW = (MaxI − MinI)/NI ;
For i = 1 : NI
    Intervals[i] = interval(i, MinI, MaxI, IW);
End
For i = 1 : NR
    (S_t, d_t) = select(training_samples); //d_t is the
    desired value of sample S_t
     Compute IV_t = (e_{1,t}, …, e_{NE,t}) ;
// this vector is the input to the DF system
  For i = 1 : NE
      j = IM(e_{i,t}, MinI, MaxI, NI);
      k = IM(d_t, MinI, MaxI, NI);
      INDT[i, j, k], sum = INDT[i, j, k], sum + e_{i,t}
;
      INDT[i, j, k], num = INDT[i, j, k], num + 1 ;
    End
 End
 For i = 1 : NE
    For j = 1 : NI
       For k = 1 : NI
          IDT(i, j, k) = INT(INDT.sum(i,j,k)/INDT.num(i,j,k), IW) ;
       End
    End
 End
```

Figure.6. The algorithm of making INDT

By making IDT, a basic tool for the fusion algorithm is produced. The fusion algorithm needs another tool named IDP as indicated in figure 7. By analyzing the instructions in figure 6, it is obvious that the time complexity of the algorithm is $TC = O(NE \times NI^2)$.

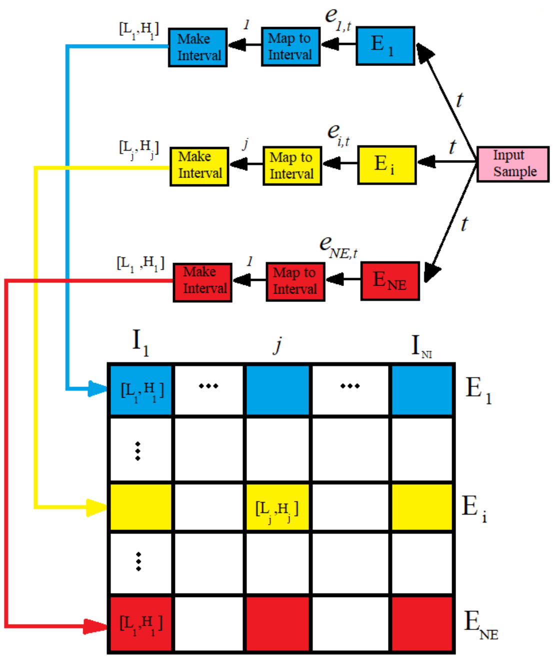


Figure.7: Interval Decision Profile (IDP)

To make IDP for each sample, the inaccuracy of each ES for all the training samples should be computed as $EIA(i)$. A key point in the fusion system is that the estimators should be diverse to increase the effectiveness of the fusion [2]. Therefore, diversity in making the estimators should also be considered. Equation (7) computes the inaccuracy of the i^{th} ES. The value of the inaccuracy of the ES is not only from its training but also from uncertainty in sensor outputs.

$$EIA(i) = 1 - \frac{\sum_{j=1}^{NI} INDT.num(i,j,j)}{\sum_{j=1}^{NI}\sum_{k=1}^{NI} INDT.num(i,j,k)} \qquad (7)$$

After computing equation (7), for each test sample $(\vec{t})$ the outputs of all the estimators are computed as a vector $\overrightarrow{IV_t} = (e_{1,t}, \dots, e_{NE,t})$. Then, each component of $\overrightarrow{IV_t}$ is mapped into an interval by the mapping function $IM(e_{i,t}, intervals, NI) \to j$. In this function, 'intervals' is the list of the defined intervals as indicated in figure 4. By doing this, the value of the entry $IDP(i,j)$ is the interval computed based on the value of $e_{i,t}$ and inaccuracy of the i^{th} ES computed by equation (7).

Basically, the proposed fusion method (IDTF) has four steps for each test sample $\vec{t}$. In the first step, the inaccuracy of the estimators is computed based on equation (7). In the second step, the outputs of all the estimators are computed as a vector $\overrightarrow{IV_t} = (e_{1,t}, \dots, e_{NE,t})$. In the third step, the IDP is computed based on the IV_t. In the last step, the similarity of the IDP with all the values $IDT(1 \dots NE, 1 \dots \mathrm{NI}, k), k = 1 \dots NI$ is computed and the interval corresponding to the highest similarity (lowest dissimilarity) is selected as the fusion output. Equation (8) computes the similarity between different interval matrices $A_{n\times m}$, $B_{n\times m}$ based on equation (1).

$$Sym(A,B) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} S(A(i,j), B(i,j))}{n \times m} \tag{8}$$

Figure 8 indicates details of the IDTF algorithm. In figure 8, the output of the IDTF algorithm is an interval determined in line 16. To have a single value as the output of the algorithm, a mapping function $mapi: [L,H] \to \Re$ is defined to map the interval into a single value, as computed in line 17 in figure 8.

Algorithm IDTF

Inputs: $INDT, IDT, NT, NE, NI, test_{samples}, MaxI, MinI$,*intervals*

Outputs: fi, fo
// fusion interval and fusion output respectively

1- *// NT is the number of test samples*
2- $IW = (MaxI - MinI)/NI$;
2- *// the width of the intervals defined in the training step*
3- *For* $i = 1 : NE$
3- *// lines 3-11 compute IDP*
4- $Inaccuracy(i) = EIA(i, INDT, NI)$; *// computed by equation (7)*
5- *End*
6- *For* $t = 1 : NT$ *// this loop is computed for all the test samples*
7- *Make* IV_t *by producing ES outputs* $e_{i,t}$, $i = 1 \dots NE$;
8- *For* $i = 1 : NE$
9- $j = IM(e_{i,t}, intervals, NI)$;
10- $IDP(i,j) = INT(e_{i,t}, 2 \times inaccuracy(i))$;
11- *End*
12- *For* $k = 1 : NI$
13- $Similarity(k) = Sim(IDT(*,*,k), IDP)$; *// finds similarity*
4- *// based on equation (8)*
14- *End*
15- $i = argmax(similarity(j), j = 1 \dots NI)$; *// find the interval*
5- *// corresponding to maximum similarity*
16- $[fi(t).L, fi(t).H] = [IDT(1,1,i).L, DT,.H]$;
6- *// the interval*
7- *// defined by the fusion*
17- $fo(t) = mapi([fi(t).L, fi(t).H])$; *// mapping the fusion interval*
8- *// into a single value*
18- *End*

Figure.8. Details of the IDTF algorithm

The algorithm in figure 8 is a basic fusion method that uses the IDT and IDP with interval data. However, the IDTF method may be combined with some other fusion methods to adapt them to use interval data. In this case, the effectiveness of the algorithm may also be increased in situations where uncertainty exists in the data received from sensors or other sources, and the intervals are defined to deal with these types of problems. To indicate how this is done, the IDTF is combined with Dempster Shafer (DS) and Bayes methods to introduce new combined fusion methods named Interval DS (IDS) and Interval Bayes (IBS) respectively. More details about these methods are as follows.

In DS theory, a belief degree is assigned to each element such that the total value of the degrees for all the members is 1. For this assignment, the upper and lower bounds of the probability interval can be defined as interval [B, Pl] where B and P1 are lower and higher values of the probabilities. In the proposed method, interval [L, H] indicates a range of possible values for sensor inputs that may be used as a range of possible probabilities for sensor inputs that can be used in the DS theory. The way that the interval values may be used in the DS method is described in detail here.

To describe IDS, the DS method should be changed to use the IDT as a tool for DF. In the IDS method, $IDT_j^i(1 \dots NI)$ is a vector indicating all the values of i^{th} row of the IDT (produced by i^{th} ES) related to the j^{th} output interval. Moreover, the value of $DP^i(1 \dots NI)$ is the i^{th} row of the DP produced by i^{th} ES. Therefore, equation (9) computes the degree of support

(DOS_j) for some input value (x) to the j^{th} interval.

$$DOS_j(x) = k \times \prod_{i=1}^{NE} belief_j(E_i) \tag{9}$$

In equation (9) the value of $belief(E_i)$ is the belief degree of the i^{th} ES for the j^{th} interval computed by equation (10).

$$belief_j(E_i) = \frac{\varphi_j^i(x)\prod_{k\neq j}(1-\varphi_k^i(x))}{1-\varphi_j^i(x)\left[1-\prod_{k\neq j}(1-\varphi_k^i(x))\right]} \tag{10}$$

In equation (10), the value of φ_j^i is the proximity of the i^{th} row of the IDT (produced for the j^{th} output interval) with the i^{th} row of the IDP produced for x. Therefore, the value of $\varphi_j^i(x)$ is computed by equation (11).

$$\varphi_j^i(x) = \frac{(1+\sum_{k=1}^{NI}[1-sim(IDT(i,k,j),IDP(i,k)])^{-1}}{\sum_{l=1}^{NI}(1+\sum_{k=1}^{NI}[1-sim(IDT(i,k,l),IDP(i,k)])^{-1}} \tag{11}$$

By computing equation (9) for the sample (x) and each interval (j), the fusion output for the x is computed by equation (12).

$$fo(x) = mapi(interval(argmax\left(DOS_j(x)\right)) \tag{12}$$

By using equation (9), for each test sample (x), a degree of support is computed as a similarity measure, and equation (12) determines the interval with a maximum similarity that can be used as the fusion output after the mapping function (mapi) maps the interval into a single real value. Figure 9 describes details of the algorithm that produces the fusion output for a test sample (x).

Algorithm IDS // interval Dempster Shafer
Inputs: NI, NE, IDT, INDT, x
Output: fo //estimated output for the sample (x)
1- *Define DOS[NI] to be the degree of support of x to each interval defined by INDT* 2- *Compute IDP for the input sample* 3- *Set all entries of DOS to 1* 4- *For $j = 1:NI$* 5- *Compute proximity of i^{th} row of IDT for the j^{th} interval to the i^{th} row of IDP for x by equation (11)* 6- *Compute belief degree for the i^{th} ES for the j^{th} interval* 7- $DOS[j] = DOS[j] \times k$ 8- *For* $i = 1{:}NE$ 9- $DOS[j] = DOS[j] \times belief(i)$ 10- *End* 11- *End* 12- $m = argmax(DOS[n]).n = 1{:}NI$ 13- $[fi.L, fi.H] = [IDT(1,1,m).L, IDT(1,1,m).H]$ 14- $fo = mapi([fi.L, fi.H])$

Figure 9. Interval Dempster Shafer (IDS) algorithm

The Bayes method is also another fusion method that may be adapted for interval data. To do this, the INDT used in the IDTF is also used by the Bayes method to define the interval Bayes (IBS) method. To compute the $DOS_j(x)$ which is the degree of support of the test sample (x) for the j^{th} interval (or the posterior probability of j^{th} interval), equation (13) is used.

$$DOS_j(x) = \left(\frac{\sum_{k=1}^{NE}\sum_{l=1}^{NI} INDT(k,l,j).num}{\sum_{k=1}^{NE}\sum_{l=1}^{NI}\sum_{m=1}^{NI} INDT(k,l,m).num}\right)^{NE} \times \prod_{i=1}^{NE} INDT(i, INum(x, intervals), j).num \tag{13}$$

By using equation (13) for every interval (j), equation (12) is also used to compute the fusion output. In equation (13), INum is a function that represents the interval that contains the input x. In the next section, the designed experiments use the equation 9 thru 13 to implement the IDS and IBS methods for comparison. The algorithm in figure 10 indicates details of the interval Bayes algorithm.

Algorithm IBS //Interval Bayes
Inputs: NI, NE, IDT, INDT, x
Output: fo // estimated value for the test sample (x)
1- *s=INum(x,intervals);* 2- *N=sum of all entries of INDT[1..NE,1..NI,1..NI].num* 3- *For i=1:NI* 4- *Pk=sum of all entries of* $INDT[1..NE,1..NI,i].num$ 5- $mu[i] = pk/N$ 6- *For j=1..NI* 7- $mu[i] = mu[i] \times INDT[j,s,i].num$ 8- $mi = argmax(mu[i], i = 1..NI)$; 9- *End* 10- *End* 11- $[fi.L, fi.H] = [IDT(1,1,mi).L, IDT(1,1,mi).H]$ 12- $fo = mapi([fi.L, fi.H])$

Figure.10.Interval Bayes (IBS) algorithm

To analyze the time complexity (TC) of the IDTF algorithm, three different parameters should be considered: NE, the number of estimators, NI, the number of intervals, and NT, the number of test samples. In figure 8, the TC for lines 3-5 is $TC_{35} = O(NE \times NI^2)$. The TC for lines 8-11 is $TC_{811} = O(NE \times NI)$. The TC for lines 12-14 is $TC_{1214} = O(NE \times NI^2)$. None of the other lines in figure 8 have a time complexity more than these values. Therefore, equation (14) computes the total time complexity of the algorithm in figure 7.

$$TC = O(NE \times NI^2 + NI \times (NE \times NI + NE \times NI^2)) = O(NE \times NI^3) \quad (14)$$

4. Experiments and analysis

In this section, some experiments are designed to not only indicate how the IDT may be used as a tool for DF but also indicate the effectiveness of using it with other DF methods. In section 3.1 details of the datasets used for the experiments are described and the evaluation method is determined. The methods that are selected for comparison are also introduced. Details of the experiments, comparisons, and analysis of the results are presented in section 3.2.

4.1. Introducing datasets and evaluation method

As indicated in figure 2, the estimators receive sensor outputs and estimate a property of the environment. In order to simulate the scenario for the experiments, the features of each sample in each dataset are defined as sensor outputs (indicated as vector $\vec{t}$) received by the ESs used in the DF (figure 3). The output of each ES is the estimated output value for the sample. Therefore, inputs to the IDTF algorithm are the values estimated by all the ESs (indicated as vector $\overrightarrow{IV}$) and the output value is the final estimation ($\hat{e}$).

In the IDTF method (figure 3), for the simulated scenario used for the experiments, the structure of the estimators may be any nonlinear regression component, fuzzy inference system, neural network or any other trainable estimator. A key point that should be considered is that the estimators should be diverse not only in structure but also in behavior. The effectiveness of any fusion system depends on the diverse operation of its components. If the components (estimators) have the same behavior, the fusion of their outputs may not improve the performance of the system. Therefore, the diversity of the components should be guaranteed [2].

In the designed experiments, the ESs are neural networks, each with one or two hidden layers with a different number of neurons. To have more diverse ESs, bagging is used to select training data. Therefore, each ES is trained by randomly selecting 40% of the bag data by replacement. These assumptions guarantee the diversity required for the fusion system.

In the experiments, the IDTF is compared with the IDS and IBS methods. It should be noted that using intervals for sensor outputs is when some level of uncertainty exists in the values. Therefore, the methods that use interval data may not be compared with the methods that use non-interval data. It is because IDS and IBS methods are used for the comparisons.

In this section, mean square error (MSE) is used as a measure to compare the selected fusion methods. If d_i is the desired output value for the test sample x, the MSE is computed by equation (15).

$$MSE = \frac{1}{NT}\sum_{i=1}^{NT}(fo(i) - d_i)^2 \quad (15)$$

In equation (15), the value of $fo(i)$ is the fusion output for the i^{th} test sample produced by different fusion methods and NT is the number of test samples.

The datasets used for the experiments in section 3.2 are selected from KEEL dataset repository [37] and UCI repository of machine learning as described in table 1.

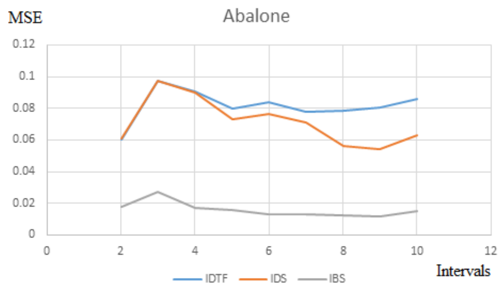
Table (1): Datasets description

No	Name	#Features	#Samples	Dataset
1	Abalone	8	4177	Keel
2	Concrete	8	1030	UCI
3	Fireman	5	1200	Keel
4	Electric Grid Stability	12	10000	UCI
5	Laser	4	993	Keel
6	Stock	9	950	Keel
7	Wine_red	11	1599	UCI
8	Wine_white	11	4898	UCI

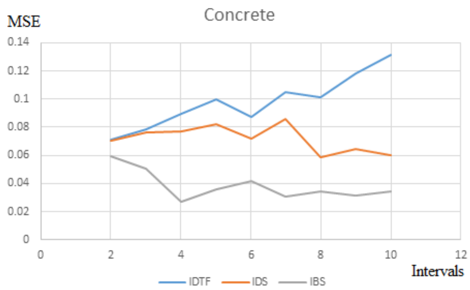
In the designed experiments, 70% of the samples in each dataset are selected randomly for training and the remaining 30% for testing. All the datasets are normalized and 4-fold cross-validation is used for all the experiments. Details of the experiments are described in the next section.

4.2 Experiments details

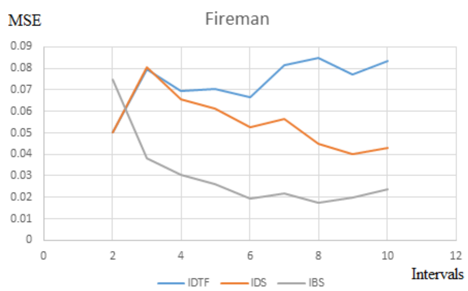
In the methods that use IDT as a tool for data fusion, the number of intervals and the number of ESs are the two parameters that influence the effectiveness of the methods. To determine the effect of the number of intervals, the next set of experiments is designed. Figure 11 compares MSE for the selected interval methods when the number of intervals increases.



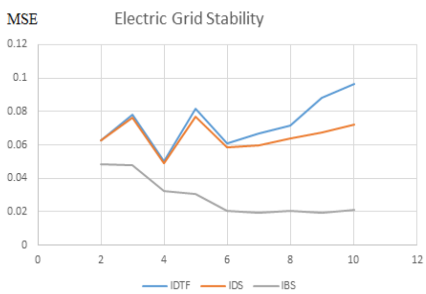
(a)



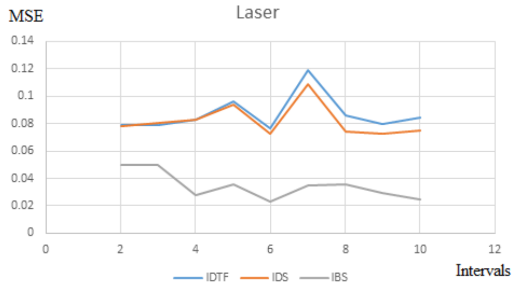
(b)



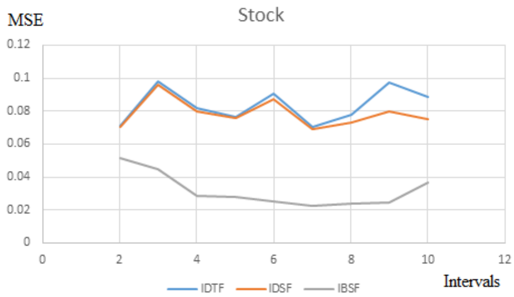
(c)



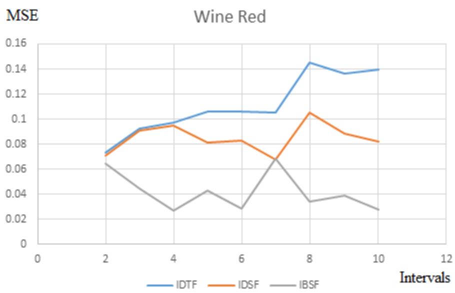
(d)



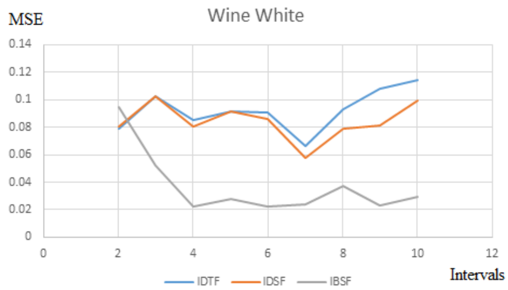
(e)



(f)



(g)



(h)

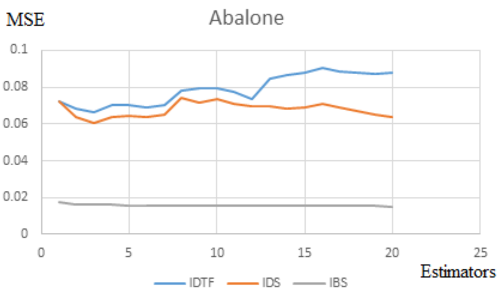
Figure 11. The effect of the number of intervals on total error of the four outputs

By analyzing the results in figure 11, some points should be considered. As a general rule, it is expected that by increasing the number of intervals, the uncertainty in ES outputs decreases, and the error also decreases. In the figure, increasing the number of intervals causes a decrease in error to some point. In some intervals, an increase in the error reveals the fact that the decisions for the samples near the borders may cause more errors, especially for the IDTF that only compares intervals, while for the IDS and IBS that probability is also considered, the negative effect of border data also decreases as indicated in figures 10(a, b, c).

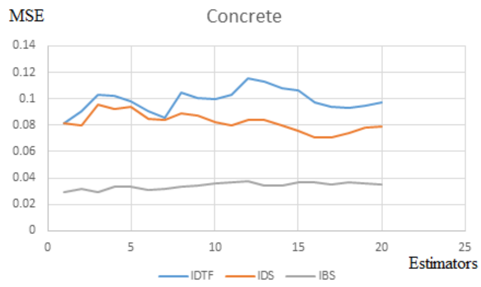
By comparing the results obtained by IDTF and IDS in figure 11, it is obvious that IDS produces fewer errors than IDTS. The reason is that in IDTS only intervals are compared for the similarity measure, while in IDS this interval comparison is combined with probability in the form of proximity and belief. Fewer errors produced by IBS in figure 10 with respect to the others are mainly because for computing the degree of support, no interval comparisons are used, instead, only the number of training samples in each interval is used for the computations by equation (13), therefore less uncertainty exists in the computations.

In a conclusion, the experiments in figure 11 indicate the sensitivity of the methods to the number of intervals. Moreover, they indicate that by combining the IDTF method with other fusion methods, more accurate results may be obtained while the flexibility of IDTF for combining with other fusion methods is also proved.

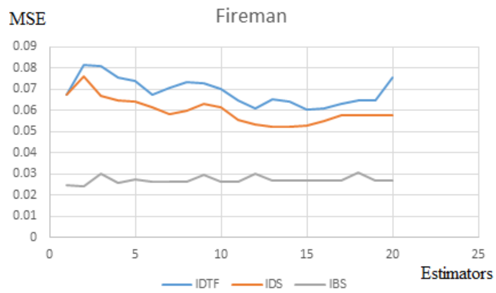
The next set of experiments compares the performance of the selected methods by increasing the number of ESs as indicated in figure 12. In these experiments, the number of intervals is the same for all the datasets and methods. The number of intervals is selected based on the results obtained in figure 11 by the fact that with a mean value of 5 intervals, moderate errors are obtained in most cases.



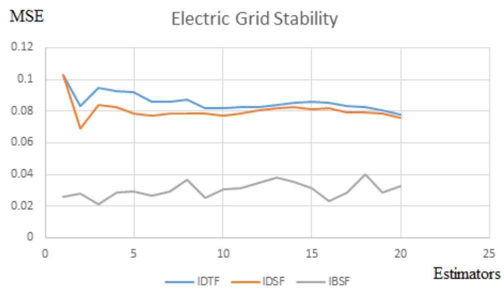
(a)



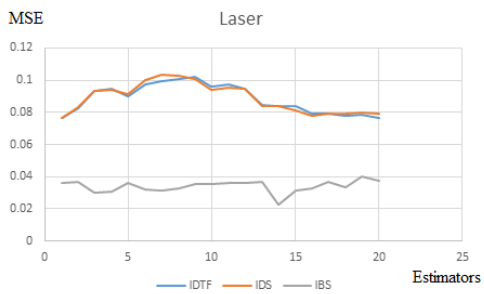
(b)



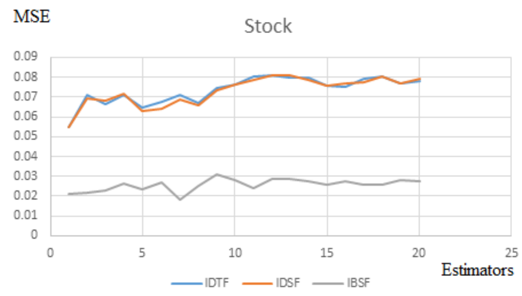
(c)



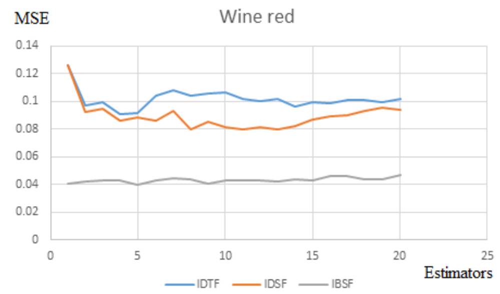
(d)



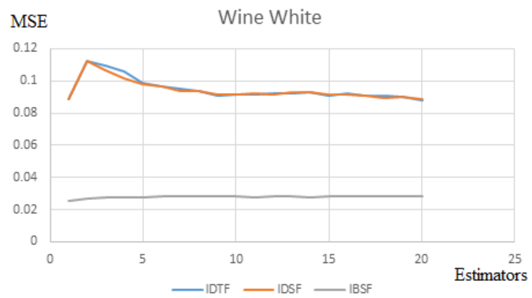
(e)



(f)



(g)



(h)

Figure 12: Comparing the effect of the number of ESs on the performance of the methods

A key point in figure 12 is that the ESs used in the experiments have different structures and are trained with different sets of training data. Therefore, their operations are diverse and they are randomly selected for these experiments. As shown in figure 12, IDS and IDTF methods are more sensitive to the number of ES than IBS. Moreover, the minimum error is obtained in different ranges of the number of ESs for different datasets. Basically, in the fusion algorithms, it is expected that by using more ESs, a smooth decrease in error is obtained. The results indicate that increasing the number of ESs makes the fusion system behave more smoothly although in all the cases, this increase may not decrease the value of error. This may be because of not only diversity but also the random selection of ESs. This figure also indicates that the error of IBS is less than the two others. As described in figure 11, this is mainly because, in IBS, no interval comparison is used as a source of uncertainty. Instead, mapping samples into intervals is used to compute the degree of support.

5. Conclusion

In this research, a fusion method named IDTF is introduced based on the interval definition of values. The IDTF uses interval decision template (IDT) and interval decision profile (IDP) as fusion tools. The fusion method is also useful in situations where uncertainty exists in the sensor outputs. The IDTF uses not only IDT and IDP as fusion tools but also a method for interval data comparison. Moreover, the IDT, IDP, and IDTF can be combined with other fusion methods to make new ones. In this research, they are used with Demster Shafer and Bayes methods to make interval Dempster Shafer (IDS) and interval Bayes (IBS) fusion methods. In the experiments, two main parameters of IDTF were considered: the number of intervals that define the range of input values, and the number of ESs used for the fusion. Because the methods that use intervals introduce some level of uncertainty in the data, they cannot be compared with the methods that don’t use intervals. The designed experiments compare the effects of changing the two parameters in IDTF, IDS, and IBS methods. In the first set of experiments, the effect of increasing the number of intervals is investigated and in the second set, the effect of the number of ESs is considered. The experiments indicate the effectiveness of using the interval fusion method with Dempster Shafer and Bayes methods. More research is needed to indicate how the proposed interval fusion method can be used with other methods in different application areas.

As the fusion methods are becoming more important and widely used in different application areas, the need for introducing more fusion methods also becomes more important. In future works, embedding the IDT method in a fuzzy integral algorithm to overcome the uncertainty in interval-valued applications may also be considered.

References

[1] D. L. Hall, M. E. Liggins, and J. Llinas, Handbook of Multisensor Data Fusion, CRC Press, 2009.
[2] M. Salkhordeh Haghighi, A. Vahedian and H. Sadooghi Yazdi, "Creating and measuring diversity

in multiple classifier systems using support vector data description," Applied Soft Computing, Vol. 11, No. 8, pp. 4931-4942, 2011.
[3] S. Teixeira, B. Alves Agrizzi, J. G. Pereira Filho, S. Rossetto, I. S. Araújo Pereira, P. Dockhorn Costa, A. F. B. Branco, and R. R. Martinelli, "LAURA architecture: Towards a simpler way of building Situation-Aware and Business-Aware IoT Applications," The Journal of Systems & Software, 2019.
[4] R. R. Yager, "On the Dempster-Shafer Framework and New Combination Rules," Information Sciences, Vol. 41, pp. 93-137, 1987.
[5] M. Salkhordeh Haghighi and S. Kahrobaee, "Applying a fuzzy interval ordered weighted averaging aggregation fusion to the nondestructive determination of retained austenite phase in D2 tool steel," NDT and E International, Vol. 103, pp. 39-47, 2019.
[6] M. Salkhordeh Haghighi, A. Vahedian, and H. Sadooghi Yazdi, "Extended decision template presentation for combining classifiers," Expert Systems with Applications, Vol. 38, No. 7, pp. 8414-8418, 2011.
[7] H. Chen, X. Zhang, Q. Wang, and Y. Bai, "Efficient Data Fusion Using Random Matrix Theory," IEEE Signal Processing Letters, 2018.
[8] D. Hall and S. McMullen, Mathematical Techniques in Multisensor Data Fusion, Norwood: Artech House Publishers, 2004.
[9] P. J. Jesus, C. Baquero, and P. S. Almei, "A Survey of Distributed Data Aggregation Algorithms," IEEE Communications Surveys & Tutorials, Vol. 17, No. 1, 2015.
[10] C. M. de Farias, L. Pirmez, and G. Fortini, "A multi-sensor data fusion technique using data correlations among multiple applications," Future Generation Computer Systems, 2018.
[11] X. Deng, Y. Jiang, L. T. Yang, M. Lin, L. Yi, and M. Wang, "Data Fusion Based Coverage Optimization in Heterogeneous Sensor Networks: A Survey," Information Fusion, 2018.
[12] N. Nesa and I. Banerjee, "IoT-Based Sensor Data Fusion for Occupancy Sensing Using Dempster–Shafer Evidence Theory for Smart Buildings," IEEE Internet of Things Journal, Vol. 4, No. 5, pp. 1563 - 1570, 2017.
[13] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A Survey on Data Fusion in the Internet of Things: Towards Secure and Privacy-Preserving Fusion," Information Fusion, Vol. 51, pp. 129-144, 2019.
[14] S. Kahrobaee, M. Salkhordeh Haghighi, and I. Ahadi Akhlaghi, "Improving nondestructive characterization of dual-phase steels using," Journal of Magnetism and Magnetic Materials, Vol. 458, pp. 317-326, 2018.
[15] D. Janczak and M. Sankowski, "Data fusion for ballistic targets tracking using least squares," AEU - International Journal of Electronics and Communications, Vol. 66, No. 6, pp. 512-519, 2012.
[16] C. Brüser, J. M. Kortelainen, S. Winter, M. Tenhunen, J. Pärkkä, and S. Leonhardt, "Improvement of Force-Sensor-Based Heart Rate Estimation Using Multichannel Data Fusion," IEEE Journal of Biomedical and Health Informatics, Vol. 19, No. 1, pp. 227 - 235, 2015.
[17] S. Khankalantary, S. Rafatnia, and H. Mohammadkhani, "An adaptive constrained type-2 fuzzy Hammerstein neural network data fusion scheme for low-cost SINS/GNSS navigation system," Applied Soft Computing Journal, Vol. 86, 9 11 2020.
[18] S. Mil and M. Piantanakulchai, "Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions," Applied Soft Computing, Vol. 72, pp. 65-78, 2018.
[19] M. K. Al-Sharman, B. J. Emran, M. A. Jaradat, H. Najjaran, R. Al-Husari, and Y. Zweiri, "Precision landing using an adaptive fuzzy multi-sensor data fusion architecture," Applied Soft Computing, Vol. 69, pp. 149–164, 2018.
[20] Y. Dan, J. Hongbing, and G. Yongchan, "A robust D–S fusion algorithm for multi-target multi-sensor with higher reliability," Information Fusion, Vol. 47, pp. 32-44, 2019.
[21] F. Xiao, "Multi-sensor data fusion based on the belief divergence measure of evidence and the belief entropy," Information Fusion, 2018.
[22] R. R. Yager, "On the fusion of multiple multi-criteria aggregation functions with a focus on the fusion of OWA aggregations," Knowledge-Based Systems, 2019.
[23] M. Sergey V, L. I. Khudonogova, and E. Y. Emelyanova, "Interval data fusion with preference aggregation," Measurement, 2017.
[24] Y. Huang, T. Li, C. Luo, H. Fujita, and S.-j. Horng, "Dynamic Fusion of Multi-source Interval-valued Data by Fuzzy Granulation," IEEE Transactions on Fuzzy Systems, pp. 1063-6706, 2018.
[25] T. C. Havens, D. T. Anderson, and C. Wagner, "Data-Informed Fuzzy Measures for Fuzzy Integration of Intervals and Fuzzy Numbers," IEEE Transactions on Fuzzy Systems, Vol. 23, No. 5, 2015.
[26] B. Pe̜kala, K. Dyczkowski, P. Grzegorzewski, and U. Bentkowska, "Inclusion and similarity measures for interval-valued fuzzy sets based on aggregation and uncertainty assessment," Information Sciences, Vol. 547, pp. 1182-1200, 2021.
[27] L. Shrinivasan and J. R. Raol, "Interval type-2

fuzzy-logic-based decision fusion system for air-lane monitoring," IET Intelligent Transport Systems, Vol. 12, No. 8, pp. 860-867, 2018.
[28] J. Liu, X. Deng, D. Wei, Y. Li, and Y. Deng, "Multi-attribute Decision-making Method Based on Interval-valued Intuitionistic Fuzzy Sets and D-S Theory of Evidence," in 24th Chinese Control and Decision Conference (CCDC), 2012.
[29] B. Hai and H. Yao, "An Approach to Information Fusion and Group Decision Making Based on Choquet Integral Operators for Interval-Valued Intuitionistic Fuzzy Data," International Conference on Machine Learning and Cybernetics (ICMLC), 2018.
[30] D. Dash and V. Jayaraman, "A Probabilistic Model for Sensor Fusion Using Range-Only Measurements in Multistatic Radar," IEEE Sensors Letters, Vol. 4, No. 6, 2020.
[31] K. Zhang, Z. Wang, L. Guo, Y. Peng, and Z. Zheng, "An Asynchronous Data Fusion Algorithm for Target Detection Based on Multi-Sensor Networks," IEEE Access, 2020.
[32] S. Muravyov, P. Baranov, L. Khudonogova, and M. Dai Ho, "Inertial MEMS Sensors Accuracy Improvement by Interval Fusion with Preference Aggregation," in IEEE Sensors, 2020.
[33R. Voges, and B. Wagner, "Interval -based Visual - LiDAR Sensor Fusion," IEEE Robotics and Automation Letters, Vol. 6, No. 2, 2021.
[34] G. Manogaran, V. Balasubramanian, B. S. Rawal, V. Saravanan, . C. E. Montenegro-Marin, V. Ramachandran and P. Malarvizhi Kumar, "Multi-Variate Data Fusion Technique for Reducing Sensor Errors in Intelligent Transportation Systems," IEEE Sensors, Vol. 21, No. 14, 2020.
[35] Y. Shen, Z. Wang, H. Dong, and H. Liu, "Multi-sensor multi-rate fusion estimation for networked systems: Advances and perspectives," Information Fusion, Vol. 82, pp. 19-27, 2022.
[36] Y. Wang, X. Li, J. Chen, and Y. Liu, "A condition-based maintenance policy for multi-component systems subject to stochastic and economic dependencies," Reliability Engineering and System Safety, Vol. 219, 2022.
[37] J. Alcalá-Fdez,, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, " KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework," Journal of Multiple-Valued Logic and Soft Computing, Vol. 17, No. 2-3, pp. 255-287, 2011.