IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 8, 2020, accepted December 31, 2020, date of publication March 2, 2021, date of current version March 11, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3063242*

# Multisensor Data Fusion Based on Modified Belief Entropy in Dempster–Shafer Theory for Smart Environment

**IHSAN ULLAH**[1], **JOOSANG YOUN**[2], **AND YOUN-HEE HAN**[3]
[1]Advanced Technology Research Center, Korea University of Technology and Education, Cheonan 31253, Republic of Korea
[2]Department of Industrial ICT Engineering, Dong-Eui University, Busan 47340, Republic of Korea
[3]Department of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 31253, Republic of Korea

Corresponding author: Youn-Hee Han (yhhan@koreatech.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) through the Basic Science Research Program funded by the Ministry of Education under Grant 2018R1A6A1A03025526, and in part by the Institute for Information and Communications Technology Planning and Evaluation (IITP) through the Korean Government [Ministry of Science and ICT (MSIT)] under Grant 2020-0-00116.

**ABSTRACT** Multisensor data fusion is extensively used to merge data from heterogeneous sensors in a smart environment. However, sensors provide noisy and uncertain information which is a big challenge for researchers. Since uncertainty in the data is a central constraint for data fusion and decision-making systems. Dempster-Shafer's evidence theory is an appropriate method for modeling and fusing uncertain information. In this paper, a novel data fusion scheme is proposed based on the modified belief entropy of the basic probability assignments (BPAs) to quantify the uncertainty in the information and fused them by Dempster-Shafer evidence theory. The proposed DFUDS (data fusion based on measuring uncertainty in Dempster-Shafer) scheme considers the available redundant information in the body of evidence (BoEs). The BoEs obtained from the sensor data are processed by proposed belief entropy, and fuse all pieces of evidence by Dempster's rule of combination to transfer the conflicting data into decision-making results. Extensive computer simulation results show that the proposed scheme outperforms in terms of the degree of uncertainty, evidence, reasoning, and decision accuracy under active contexts of the smart environment.

**INDEX TERMS** Multisensor data fusion, Dempster-Shaffer theory, uncertainty measurements, belief entropy, wireless sensor network.

## I. INTRODUCTION

The smart environment consists of several numbers of tiny and pervasive devices to collect data on surroundings, and extracting information from them [1]. Smart IoT environments are often based on sensing, reasoning, inferencing, and generally exploit a sensory infrastructure [2]. These sensor devices often produce imprecise or corrupt data reducing the inferencing accuracy and energy efficiency of the system. Moreover, data generated in the smart IoT environment are conditionally or unconditionally reliant on each other. In such situation, it is necessary to adopt a method of fusion for heterogeneous sensor data. However, it is very challenging to merge heterogeneous sensor data to infer accurate decisions. In smart IoT environments, quantifying the uncertainty of heterogeneous sensor data is very essential for making the accurate decision. Furthermore, it is also important to maximize the lifetime and performance of the sensor network by operating the system according to the context and condition of sensory data. Hence, a smart IoT environment requires a context-aware operation for better decision-making accuracy with minimal energy consumption and high network load.

Multisensor Data fusion is used to combine data from multiple sources to produce more concise and accurate data [3]–[7]. It can effectively handle the noisy data generated in a dynamic environment and supports data representation in a way that helps the decision-making process based on the available information. Note that the imperfection aspects of the sensor data such as uncertainty, inaccuracy, and inconsistency, can result in false beliefs and inferences about the environment [8], [9]. Multisensor data fusion techniques are broadly classified into three categories concerning the

The associate editor coordinating the review of this manuscript and approving it for publication was Chunsheng Zhu.

employed mathematical method; i) probability-based fusion, ii) artificial intelligence-based fusion, and iii) evidence-based fusion [10]. Dempster–Shafer (DS) evidence theory, as an evidence-based fusion method, is one of the effective approaches for data fusion [11]. It is effective for modeling and processing uncertain information to transfer the conflicting data into decision-making results. However, quantifying the uncertainty of obtained sensor data before applying data fusion is still a challenge and hot topic for the researchers.

Uncertainty often comes from incomplete information and imprecision. Several theories have been developed to deal with the uncertainty of the data, such as Shannon entropy [12], Dempster-Shafer (DS) evidence theory [11], and rough set theory [13]. Some methods e.g. Hohle's confusion measure [14], Yager's dissonance measure [15], the weighted Hartley entropy [16], Klir & Ramer's discord measure [17], Klir & Parviz's strife measure [18], and Deng entropy [19] have been proposed to measure the uncertainty of the evidence in the DS theory. Among these methods, Deng entropy is widely used and has been successfully applied in several applications [20]–[24]. Nevertheless, Deng entropy can be further improved to evaluate the redundant information in the body of evidence (BoE) including different parameters of the sensor data. We adopted previous research [22]–[24] to improve the Deng entropy operation using set theory.

The proposed DFUDS scheme allows to collect data from the IoT environment and quantify the uncertainty in the uncertain information obtained from the sensor before sending them to the fusion node for combining and decision-making. Deng entropy in DS evidence theory is the more efficient and effective way to measure the uncertainty in the sets of BoEs [19]. The proposed scheme extends the Deng entropy for reducing the redundant information in each BoEs by estimating the degree of uncertainty. In our scheme, we used 'union of set theory' to remove the redundant entity from the BoEs before measuring the uncertainty of them. After measuring the degree of uncertainty, all pieces of evidence are combined by DS rule of combination to fuse and get the final result for inferencing and making the decision. Computer simulation has shown that the proposed scheme significantly outperforms the other schemes in terms of quantifying the degree of uncertainty and fusion accuracy. The contributions of the proposed scheme are summarized as follows:

- The proposed entropy is based on Deng entropy to estimate the degree of uncertainty of each sensor measurements fused them. Sensory measurements in the smart environment contain large uncertainty, and decision-making in the uncertain environment is a big challenge. Thus, before combining the evidence for fusing data, proposed belief entropy is applied to the mass value of each sensor measurements.
- Our scheme reduces the redundant information in the BoEs, so that it increases the degree of uncertainty between the BoEs and helps to perform the correct reasoning and improve the inference accuracy based on pieces of evidence from the sensor data of the smart environment.
- DFUDS used DS rule of combination to combined all pieces of evidence from sensors and get the final result for inferencing and making a decision. Such self-optimization and dynamic behaviors able the system to respond to the situation dynamically based on the inferred knowledge and evidence from the fused data of the environments.
- The proposed scheme minimizes energy consumption and network traffic by transferring only the inferred knowledge extracted from the fusion data instead of the entire data.

The rest of the paper is organized as follows: in Section 2 the work related to data fusion for WSN is discussed. The proposed scheme on multisensor data fusion for WSN is presented in Section 3. Section 4 discusses the simulation results, and the conclusion is made in Section 5.

## II. RELATED WORK

### A. DATA FUSION

The data fusion problem has been studied by several researchers. In [25], the authors presented a smart fusion framework specifically aimed at combining heterogeneous, multimedia, multimodal real-time big data streams from hard and soft smartphone sensors to achieve synergistic convergence leading to better operational intelligence in computer-based decision support systems. In [26], Dynamic Bayesian networks (DBNs) is used to consider the past belief of the system and capturing the dynamicity of the phenomena under observation. The authors adopted DBN to perform adaptive data fusion for various applications such as the new multi-sensor fusion framework which is based on the Dynamic Bayesian Network (DBN) and Convolutional Neural Network (CNN) for Sign Language Recognition (SLR). In [27] a framework is developed which focuses on the privacy-aware data fusion for sensory data and tried to secure user privacy before smart city data are integrated. In [9] a data aggregation and fusion scheme is presented for WSN based on redial basis function neural network (RBFNN). It employed a neural network technique to efficiently aggregate data, and eliminate unnecessary information before classifying them.

A weight-based fuzzy data fusion algorithm is proposed in [28] which enhances the accuracy of the data fusion process by assigning the weight to the clustering head of the WSN. In [29] an adaptive distributed Bayesian approach is proposed for detecting outliers in the collected data. Hou *et al.* [30] proposed an event-driven dynamic clustering scheme and a data fusion algorithm that relies on neural networks, using a dynamic clustering and cluster head selection process based on the severity of the event and the remaining energy of the node. Here, the authors used a back-propagation neural network model to fused and extract large amounts of data. In [31] a multisensor data fusion approach was presented for medical data from body sensor networks, which obtained and fused data from BSNs in the fog computing environment. In [32] an approach is proposed that allows parallel data fusion in WSN which enables a trade-off between different user-defined metrics through the use of a genetic machine learning algorithm. In [33], the authors were proposed a scheme for data fusion and aggregation based on the combination of sensor node scheduling and batch estimation.

In [34], a DS evidence-based algorithm was presented to evaluate the spatial correlation between nodes at different distances. The output from each sensor node is characterized by weighted evidence rather than a crisp value, and the states of adjacent nodes are reasonably fused based on their contribution to detection. In [35], a data fusion technique was proposed that uses DS theory to detect events on Twitter. Two types of data were used in this scheme. The first was an attribute extracted from text using the word collection method, and the second was a visual attribute extracted by applying a scale-invariant attribute transformation. DS theory is applied to combine data from both sources. A multisensor data fusion technique based on DS theory was presented in [36] to detect indoor homes from numerous sources. In [37], a general fusion approach was proposed that uses DS theory to address the uncertainty of sensor readings and capture the characteristics of the environment. This effort is considered an ideal detection model for the existing approach. A two-step technique for building a belief function based on sensor data is described.

### B. DEMPSTER-SHAFER EVIDENCE THEORY

Dempster-Shafer (DS) theory is a proof theory that deals with measures of belief and mass function. It considers a generalizing form of probability theory and is usually used to reason the uncertainty in data from multiple sources [38].

$$\begin{cases} \sum_{A \subseteq \Theta} m(A) = 1 \\ m(\emptyset) = 0 \end{cases} \tag{1}$$

where $\emptyset$ and $m(A)$ represent the null set and basic belief assignment function for subset $A$, respectively. Belief and plausibly for $A$ can be expressed mathematically as

$$bel(A) = \sum_{X \subseteq A} m(x) \tag{2}$$

$$pl(A) = \sum_{X \cap A \neq \emptyset} m(x) = 1 - bel(\bar{A}) \tag{3}$$

where $bel(A)$, $X$, and $pl(A)$ represents the degree of confidence, frame of discernment (FoD), and trustworthiness of $A$ is not false. For any (power set of all pieces of evidence, see Section: III for more details) The relation between belief, unbelief, unknown, $unk(\cdot)$, and plausibility functions for the event, $E$, are described in Figure 1.

$$bel(E) + bel(\bar{E}) \leq 1, \quad \forall E \subseteq \Theta \tag{4}$$

$$bel(E) + bel(\bar{E}) + unk(E) = 1, \quad \forall E \subseteq \Theta \tag{5}$$

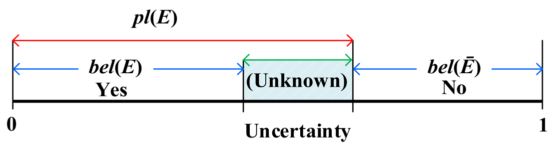



**FIGURE 1.** The relationship between belief, disbelief, unknown, and plausibility function.

$$pl(E) = bel(\Theta) - bel(\bar{E}), \times \forall E \subseteq \Theta \tag{6}$$

$$unk(E) = 1 - \left(bel(E) + bel(\bar{E})\right) \tag{7}$$

### C. UNCERTAINTY MEASURES IN DEMPSTER–SHAFER

Uncertainty measurement of the BoEs in the DS theory is an effective approach to quantify and fuse the data obtained from the sensor. It tries to estimate the level of information in each BoE by using entropy measurements. The proposed scheme extends the Deng entropy to improve the uncertainty measurement and fuse the evidence obtained from BoEs by DS rules of combination. The existing methods for uncertainty measurement in DS are explained in the following [22]–[24].

#### 1) HOHLE's CONFUSION MEASURE

Hohle's confusion measure is one of the earliest methods to measure the uncertainty in DS. Mathematically it can be formulated as follows [14]:

$$C_H(m) = -\sum_{A \subseteq X} m(A) \log_2 \frac{1}{Bel(A)} \tag{8}$$

where $C_H$, $Bel(A)$, $m(A)$, and $X$ denote the Hohle's confusion measure, belief function, mass function of proposition $A$, and frame of discernment (FoD) respectively.

#### 2) YAGER's DISSONANCE MEASURE

Yager contributes to measure the uncertainty and it is called Yager's dissonance measure, $E_Y$, which defined as follows [15]:

$$E_Y(m) = -\sum_{A \subseteq X} m(A) \log_2 Pl(A) \tag{9}$$

where $Pl(A)$ is the plausibility function and $m(A)$ is the mass function of proposition $A$.

#### 3) WEIGHTED HARTLEY ENTROPY

Weighted Hartley entropy was proposed by Dubois and Prade to measure entropy, $E_{DP}$. Using the cardinality, $|A|$, it can be expressed mathematically as follows [16]:

$$E_{DP}(m) = -\sum_{A \subseteq X} m(A) \log_2 |A| \tag{10}$$

#### 4) DISCORD MEASURE

Discord measure, $D_{KR}$, was proposed by Klir and Ramer which describes the belief entropy using the intersection of the focal elements of the FoD as follows [17]:

$$D_{KR}(m) = -\sum_{A \subseteq X} m(A) \log_2 \sum_{B \subseteq X} m(B) \frac{|A \bigcap B|}{|B|} \tag{11}$$

#### 5) STRIFE MEASURE

Strife Measure was presented by Klir and Parviz for belief entropy in DS [18]. It also called Klir and Parviz's strife measure $S_{KP}$, which is more similar to Ramer's method [17].

$$S_{KP}(m) = -\sum_{A \subseteq X} m(A) \log_2 \sum_{B \subseteq X} m(B) \frac{|A \bigcap B|}{|A|} \tag{12}$$

#### 6) GEORGE AND PAL's CONFLICT MEASURE

George and Pal proposed the total conflict measure, $TC_{GP}$,. It can be formulated as follows [39]:

$$TC_{GP}(m) = -\sum_{A \subseteq X} m(A) \sum_{B \subseteq X} m(B) \left(1 - \frac{|A \bigcap B|}{|A \bigcup B|}\right) \tag{13}$$

#### 7) SHANNON ENTROPY

Shannon entropy is usually used to quantify the uncertainty and amount of information contained in a message. It can be formulated as follows [12]:

$$H(p) = -\sum_{i=1}^{n} p_i \log_b p_i \tag{14}$$

where $n$ and $p_i$ represents the number and probability of state $i$, respectively.

#### 8) DENG ENTROPY

Deng entropy, $E_d$, is used to generalize Shannon entropy to measure the uncertainty of DS theory. It degenerates the Shannon entropy to model the information. It can be formulated as [19],

$$E_d(m) = -\sum_{A \subseteq X} m(A) \log_b \frac{m(A)}{2^{|A|} - 1} \tag{15}$$

where $X$ denotes the FoD, $|A|$ represents the cardinality of the proposition $A$. The mass function of proposition $A$, $m(A)$, is treated as the probability of proposition $A$. Mathematically it can be represented as [19]:

$$E_d(m) = -\sum_{A \subseteq X} m(A) \log_b m(A) \tag{16}$$

## III. THE PROPOSED SCHEME

In this section, the proposed DFUDS (data fusion based on measuring uncertainty in Dempster-Shafer) scheme for multisensor data fusion with WSN is presented. It quantifies the uncertainty of the information in the BoE using the proposed

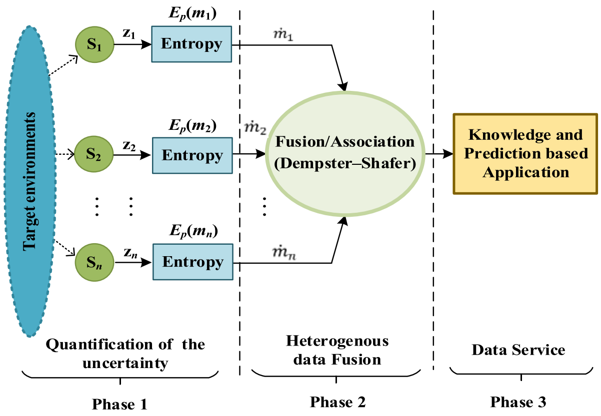



**FIGURE 2.** The three-phase operation of the proposed scheme.

entropy for DS. After that, the individual mass value obtained by mass function from the sensor data, $z_i$, is combined to achieve accurate fusion result using the DS evidence rule of combination. The three-phased operations of the proposed scheme are explained in Figure 2.

In DS evidence theory, mass functions and BoE both are the sources of uncertain information, e.g., elements and the information in a BoE may be different and contain redundant information without changing the mass value. However, Deng entropy only considerers the mass values and the cardinality of the proposition, but the redundant information in the BoE is ignored [22]–[24]. High redundant information indicates a large reduction in uncertainty, while low redundant information indicates a small reduction. On the other hand, 100% redundant information in two BoEs means the BoEs are completely similar. Redundancy and change in the information of the BoE can be evaluated to improve the accuracy of the uncertainty, which is effective for inference and accuracy of the fusion.

Let $\Omega = \{\theta_1, \theta_2, \ldots, \theta_n\}$ denotes a finite set, where $\Omega$ is called the frame of discernment (FoD). The power set $2^{\Omega}$ of $\Omega$, can be written denoted as $\Theta$ is composed of $2^n$ elements as follow:

$$\Theta = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \ldots, \{\theta_n\}, \{\theta_1, \theta_2\}, \ldots, \Omega\}$$

A mass function $m$, of DS theory, is used to mapped power set, $\Theta$, to the interval [0,1], as follows [11]:

$$\begin{cases} \sum_{A \subseteq \Theta} m(A) = 1 \\ m(\emptyset) = 0 \end{cases} \tag{17}$$

here masse function, $m(A)$, indicate how strongly the evidence supports the proposition $A$.

$$(\mathfrak{R}, m) = \{\langle A, (m) \rangle : A \in \Theta, m(A) > 0\}, \tag{18}$$

$$Bel(A) = \sum_{\phi \neq B \subseteq A} m(B) \tag{19}$$

$$Pl(A) = \sum_{B \cap A \neq \phi} m(B) \tag{20}$$

Consider the two BoEs contain different information the mass values are same.
First BoE:

$$\{<\{a,b,c\}, m_1>, <\{d,e,f\}, m_1'>\}\ s.t. m_1(\{a,b,c\}) = 0.4 \text{ and } m_1'(\{d,e,f\}) = 0.6$$

Second BoE:

$$m_2 : m_2\{a,b,c\} = 0.4, m_2'\{c,d,f\} = 0.6$$
$$\{<\{a,b,c\}, m_2>, <\{c,d,f\}, m_2'>\}\ s.t.\ m_2(\{a,b,c\}) = 0.4 \text{ and } m_2'(\{c,d,f\}) = 0.6$$

The mass value of both BoEs are same and Deng entropy measures the same entropy for both BoEs, even the information in the BoEs is different. The first BoE has no redundant information but the second BoE has redundant information, and Deng entropy does not consider the redundant information. Now the proposed scheme tries to handle this problem by reducing the redundant information in the BPAs to quantify the correct uncertainty of the BoEs. Our scheme used 'set theory' for reducing the redundant information to increase the degree of uncertainty of the BoE.

After measuring the uncertainty of the BoE correctly, the evidence obtained from BoEs is then combined by Dempster's rule of combination for fusing purposes. The proposed modified belief entropy, $E_p$ are explained and apply to the second BoEs as follows [22]–[24].
Second BoE:

$$m_2 : m_2\{a,b,c\} = 0.4, \quad m_2'\{c,d,f\} = 0.6$$

By taking the union of the second BoE as:

$$m_2 \cup m_2' : \{a,b,c\} \cup \{c,d,f\}$$
$$m_2 \cup m_2' : \{a,b,c,d,f\}$$

The union of two mass functions represents the reduction of redundant value in the combined mass functions.

$$\text{Let assume,} \quad A' = m_2 \cup m_2'$$

where $A'$ represent the union of the two BoE and its can be input to Deng entropy and the proposed belief entropy $E_p(m)$, based on the Deng Entropy can be formulated as [19].

$$E_p(m) = -\sum_{A\subseteq X} m(A') \log_2 \frac{m(A')}{2^{|A'|}-1} \tag{21}$$

Entropy of $A'$ is measured to get the condensed form of the information in the BPAs. High redundant information indicates a reduction in uncertainty which is hard to differentiate from each other. It increases the uncertainty of the BoEs which differentiates the BPAs in terms of information contained. So, entropy for BPAs is calculated without redundant information to obtain the accurate quantification of the uncertainty for BPAs as explained in Eq. (21).

Using DS rule of combination, two mass functions, $m_1$ and $m_2$, can be combined to get collective belief evidence as [11]:

$$m(A) = \left(E_p(m_1) \oplus E_p(m_2)\right)(A) = \frac{1}{1-k}\sum_{B\cap C=A} E_p[m_1(B)]E_p[m_2(C)] \tag{22}$$

where $k$ denotes the standardization constant for the degree of conflict between $m_1$ and $m_2$, and can be formulated as

$$k = \sum_{B\cap C=\emptyset} E_p[m_1(B)]E_p[m_2(C)] \tag{23}$$

Consider the FoD $\Omega = \{\theta_1, \theta_2, \ldots, \theta_n\}$ with $n \geq 2$ and $k$ sensors, for $k \geq 2$. Let $\mathrm{M} = [a_{ij}]$, $1 \leq i \leq k, \leq j \leq n$ be the mass matrix with $k$ rows and $n$ columns. The proposed entropy $E_p$ of the mass $m_i$ from sensor $i$ for the target event, $E$, is denoted as

$$m_i = E_p[m(s_i)]\ \begin{array}{c|cccccc} & \theta_1 & \ldots & \theta_n & \{\theta_1\cup\theta_2\} & \ldots & \Omega \\ s_1 & m_{1\times 1} & \ldots & m_{n\times 1} & \ldots & .. & m_{\Omega\times 1} \\ s_2 & m_{2\times 2} & \ldots & m_{n\times 2} & \ldots & .. & m_{\Omega\times 2} \\ \vdots & \vdots & \ldots & \vdots & .. & .. & \vdots \\ s_k & m_{1\times k} & \ldots & m_{n\times k} & \ldots & .. & m_{\Omega\times k} \end{array} \tag{24}$$

Based on DS theory the individual evidence obtained from the distributed sensors can be fused and the masses obtained from the mass function can be combined by using DS. Suppose, $m_1, m_2, \ldots, m_n$ be $n$ independent values obtained from $n$ sensor nodes $s_1, s_2, \ldots, s_n$ as evidence for an event, $E$, occurred in the smart environment. The rule for combining the evidence from sensor data is shown in Figure 3 and Figure 4.

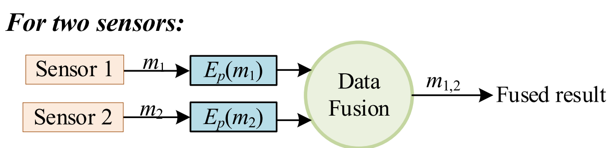



**FIGURE 3.** The Dempster-Shafer masses combination process for 2 sensors.

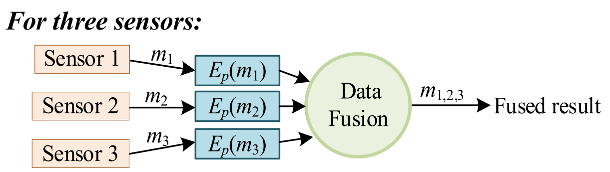



**FIGURE 4.** The Dempster-Shafer masses combination process for 3 sensors.

*For Two Sensors*:

$$m_{1,2}(E) = \frac{\sum_{s_1\cap s_2=E} m_1(s_1)\, m_2(s_2)}{\sum_{s_1\cap s_2\neq\Theta} m_1(s_1)\, m_2(s_2)} \tag{25}$$

**TABLE 1.** The summary of the result obtained from the combination of the evidence by different schemes.

| BoEs | Methods | $m_{1,2,..n}(A)$ | $m_{1,2,..n}(B)$ | $m_{1,2,..n}(C)$ | $m_{1,2,..n}(D)$ | $m_{1,2,..n}(E)$ | $m_{1,2,..n}(F)$ | Target |
|---|---|---|---|---|---|---|---|---|
| $m_1, m_2$ | DS [11] | 0 | 0.5311 | 0.3927 | 0.0239 | 0.0523 | 0 | $B$ |
| | Murphy [42] | 0.4931 | 0.412 | 0.0491 | 0.01921 | 0.0201 | 0.0065 | $A$ |
| | Yager's [15] | 0.5176 | 0.3981 | 0.0165 | 0.0339 | 0.0332 | 0.0085 | $A$ |
| | Deng [19] | 0.5861 | 0.3261 | 0.0156 | 0.0539 | 0.0423 | 0.0035 | $A$ |
| | DFUDS | 0.6161 | 0.1156 | 0.0831 | 0.0739 | 0.0923 | 0.0165 | $A$ |
| $m_1, m_2, m_3$ | DS [11] | 0 | 0.5311 | 0.3927 | 0.0239 | 0.0523 | 0 | $B$ |
| | Murphy [42] | 0.4531 | 0.482 | 0.0491 | 0.01921 | 0.0201 | 0.0065 | $B$ |
| | Yager's [15] | 0.4576 | 0.4181 | 0.0365 | 0.0439 | 0.0232 | 0.0025 | $A$ |
| | Deng [19] | 0.7161 | 0.2131 | 0.0156 | 0.0139 | 0.0323 | 0.0065 | $A$ |
| | DFUDS | 0.8061 | 0.1056 | 0.0131 | 0.0239 | 0.0623 | 0.0165 | $A$ |
| $m_1, m_2, m_3, m_4$ | DS [11] | 0 | 0.3911 | 0.5327 | 0.0239 | 0.0523 | 0 | $C$ |
| | Murphy [42] | 0.5531 | 0.312 | 0.0491 | 0.01921 | 0.0201 | 0.0065 | $A$ |
| | Yager's [15] | 0.6176 | 0.2181 | 0.0965 | 0.0439 | 0.0232 | 0.0025 | $A$ |
| | Deng [19] | 0.7161 | 0.2131 | 0.0156 | 0.0139 | 0.0323 | 0.0065 | $A$ |
| | DFUDS | 0.8061 | 0.1056 | 0.0131 | 0.0239 | 0.0623 | 0.0165 | $A$ |
| $m_1, m_2, m_3, m_4, m_5$ | DS [11] | 0 | 0.3911 | 0.5327 | 0.0239 | 0.0523 | 0 | $C$ |
| | Murphy [42] | 0.5531 | 0.312 | 0.0491 | 0.01921 | 0.0201 | 0.0065 | $A$ |
| | Yager's [15] | 0.6176 | 0.2181 | 0.0965 | 0.0439 | 0.0232 | 0.0025 | $A$ |
| | Deng [19] | 0.7161 | 0.2131 | 0.0156 | 0.0139 | 0.0323 | 0.0065 | $A$ |
| | DFUDS | 0.8061 | 0.1056 | 0.0131 | 0.0239 | 0.0623 | 0.0165 | $A$ |

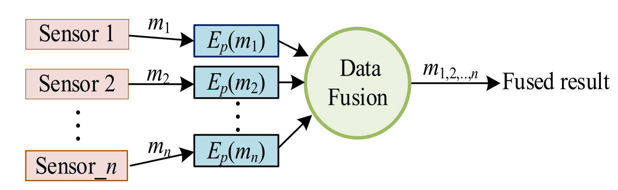



**FIGURE 5.** The Dempster-Shafer masses combination process for *n* sensors.

$$= \frac{\sum_{s_1 \cap s_2 = E} m_1(s_1)\, m_2(s_2)}{1 - \sum_{s_1 \cap s_2 \neq \Theta} m_1(s_1)\, m_2(s_2)} \tag{26}$$

The denominator is a normalization factor, $1 - K$, where $K$ is the conflict factor between the two pieces of evidence and $K < 1$.

$$\begin{aligned} K &= \sum_{s_1 \cap s_2 \neq \Theta} m_1(s_1)\, m_2(s_2) \\ &= \frac{\sum_{s_1 \cap s_2 = E} m_1(s_1)\, m_2(s_2)}{1 - K} \end{aligned} \tag{27}$$

where $m_{1,2}(E)$ is the new evidence for the event, $E$, obtained by the combining individual evidence $m_1(s_1)$ from sensor 1 and $m_2(s_2)$ from sensor 2, $\Theta$ represents the set of all possible evidence from sensor data.

*For Three Sensors*:

$$m_{1,2,3}(E) = \frac{\sum_{s_1 \cap s_2 \cap s_3 = E} m_1(s_1)\, m_2(s_2)\, m_3(s_3)}{\sum_{s_1 \cap s_2 \cap s_3 \neq \Theta} m_1(s_1)\, m_2(s_2)\, m_3(s_3)} \tag{28}$$

$$= \frac{\sum_{s_1 \cap s_2 \cap s_3 = E} m_1(s_1)\, m_2(s_2)\, m_3(s_3)}{1 - \sum_{s_1 \cap s_2 \cap s_3 \neq \Theta} m_1(s_1)\, m_2(s_2)\, m_3(s_3)} \tag{29}$$

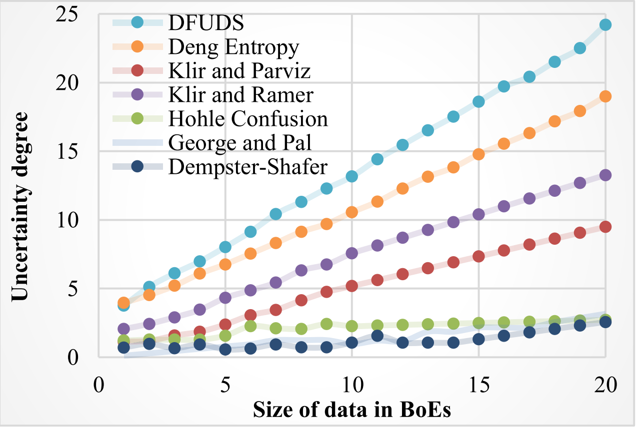



**FIGURE 6.** The uncertainty degree of BoEs based on size.

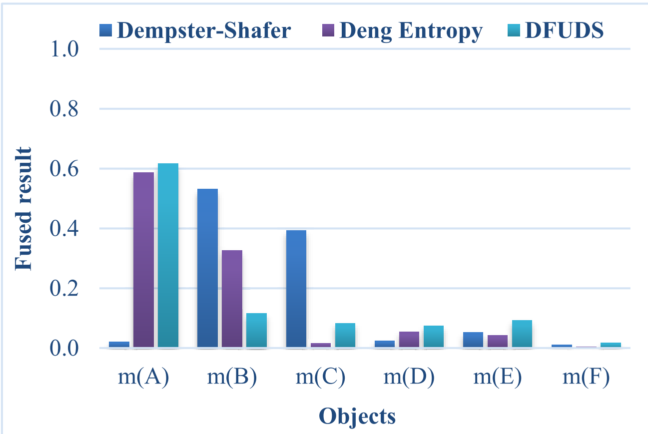
$$= \frac{\sum_{s_1 \cap s_2 \cap s_3 = E} m_1(s_1)\, m_2(s_2)\, m_3(s_3)}{1 - K} \tag{30}$$

where $m_{1,2,3}(E)$ is the new evidence about event, $E$, obtained by combining individual evidence $m_1(E)$ from sensor 1, $m_2(E)$ from sensor 2, and $m_3(E)$ from sensor 3.

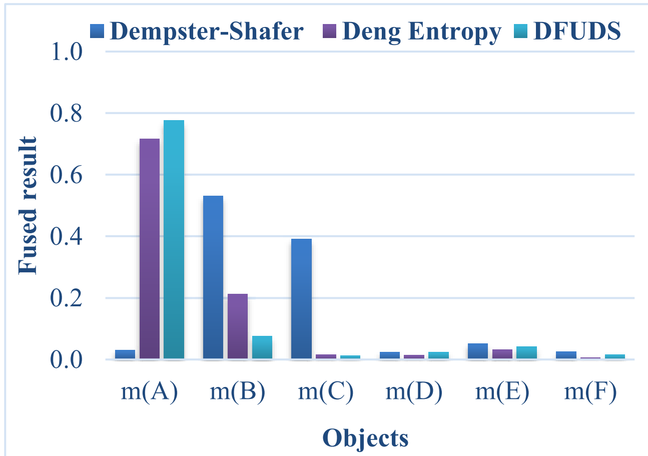
The DS theory can merge all the pieces of evidence from multiple sources. The generalized form for the combination of evidence from the $n$ number of sources (sensor nodes) is defined as shown in Figure 5.

*For n Sensors*:

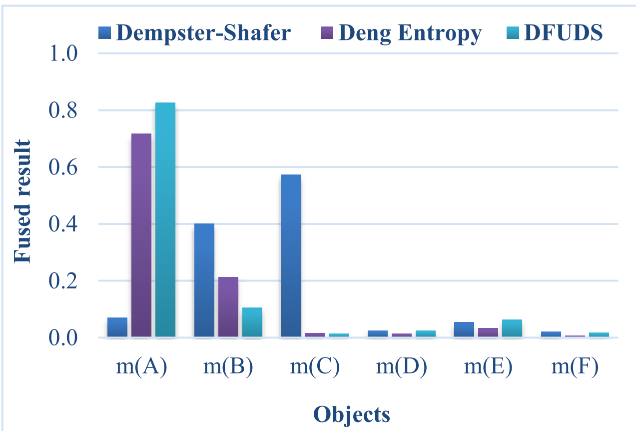
$$m_{1,2,\ldots,n}(E) = \frac{1}{1-K} \sum_{\cap_i s_i = E} \left( \prod_{1 \le i \le n} m_i(s_i) \right) \tag{31}$$



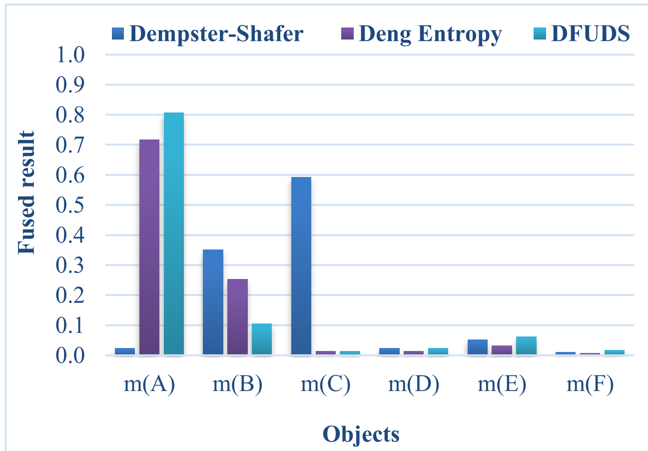

(a) Fused result by $m_1, m_2$




(b) Fused result by $m_1, m_2, m_3$




(c) Fused result by $m_1, m_2, m_3, m_4$




(d) Fused result by $m_1, m_2, m_3, m_4$

**FIGURE 7.** The comparison of fused results based on sensor mass values.

$$K = \sum_{\cap s_i = \Theta} \left( \prod_{1 \le i \le n} m_i(s_i) \right) \tag{32}$$

Plausibility $pl(E)$ is the sum of all the masses of the sets $s_i$ for events $E$. The plausibility can be derived from support or belief as follows if $bel(\Theta) = 1$

$$pl(E) = 1 - bel(\bar{E}), \quad \forall E \subseteq \Theta \tag{33}$$

Unknown mass $unk(E)$ is defined as a wrong prediction rate about the event and can be express as

$$unk(E) = 1 - \left(bel(E) + bel(\bar{E})\right) \tag{34}$$

## IV. PERFORMANCE EVALUATION

In this section, the proposed scheme is evaluated by computer simulation. Python Library, PYUDS, (python library for measuring uncertainty in Dempster-Shafer) is used to evaluate the effectiveness of the proposed scheme in terms of entropy information in the BoEs, degree of uncertainty, the accuracy of the fused result [40], [41]. Here, the DFUDS scheme measures the entropy of the BoEs using modified Deng Entropy, and fused the evidence by the DS theory. It quantifies the degree of uncertainty in the BoEs to get accurate evidence from the sensor data. The mass values obtained by the mass function are combined using Dempster's rule of combination to fused and infer the final result.

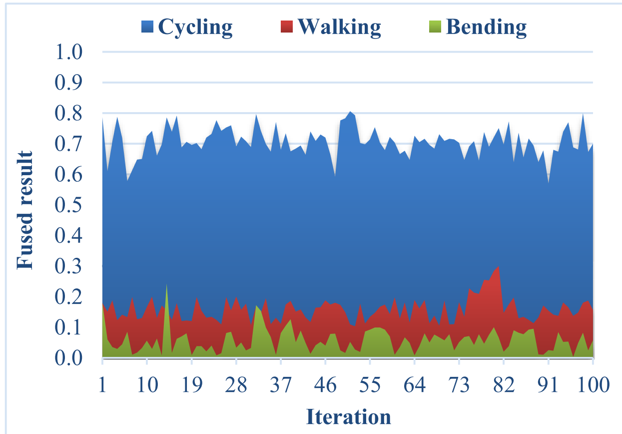
Fours representative uncertainty measurement techniques are compared with the proposed scheme, which DS [11], Murphy [42], Yager [15], and Deng [19]. The simulation consists of two phases: 1) quantifying the uncertainty and 2) data fusion with DS. In the uncertainty quantification phase, the redundant information in the BoEs is reduced, and then the degree of uncertainty between the BoEs is quantified. In the second phase, the combination rule of Dempster is applied to fuse all the pieces of evidence obtained from distributed sensors. Several metrics are examined to evaluate the performance of different uncertainty measurement and fusion schemes, which are the degree of uncertainty, the accuracy of the fused result, and the probability of detection of human activities based on fused data. The simulation results are explained in the following. Table 1 reports the summary of the fused results of the four schemes compared with the proposed scheme.

Figure. 6 shows the comparison of the degree of uncertainty measured by existing methods and the proposed scheme. The results of the figure show that the uncertainty measured by George and Pal, Klir and Parviz, Hohle's Confusion, Klir and Ramer, Deng Entropy, DS, and DFUDS increases as the number of elements in BoE increases. Similarly, Hohle's confusion uncertainty measures, George & Pal, and DS have reflected the same result in terms of degree uncertainty. It confirms that these three uncertainty methods fail to detect the change in the element of BoEs. The existing methods of entropy do not consider the redundant information in the BoEs which reduces the uncertainty between the BoEs. The proposed scheme reduces the redundant information to correctly quantify the uncertainty between the BoEs. Therefore, the proposed belief entropy is effective for uncertainty measurement in the Dempster-Shafer framework.

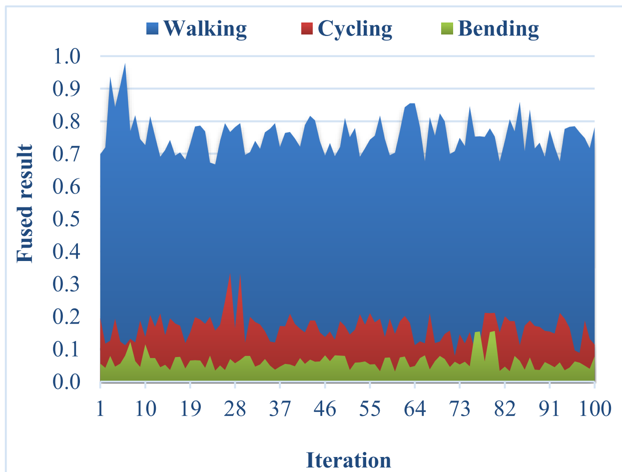
Figure 7 shows the fusion results for six different objects by combining BoEs from different sensors. Believes are distributed for different objects. The proposed entropy and Deng entropy have a strong belief for object '$A$' than the Ds entropy. The belief of the proposed scheme for other objects is very close to the Deng entropy in the case of two BoEs, $m_1$ and $m_2$. Likewise, belief results by other groups of BoEs are evaluated as shown by Figure 7 (b), (c), and (d). The fusion results with DS and Deng methods have the weak belief for all objects with all evidence and the proposed scheme has the strong and correct decision for target '$A$'. Due to conflicting evidence, the DS combination rules indicate weak evidence for the target objects, $B$, $C$, $D$, $E$, and $F$, however, the evidence results of DFUDS and Deng entropy are strong for target '$A$'. It is because the uncertainty measurement of the DFUDS scheme is more effectual and give more accurate result than the other schemes even the BoEs have same mass value. The proposed belief entropy measured the redundant information from the BoEs to contribute efficiently in sensor data fusion, especially in the presence of conflict in the evidence. The effects and benefits of new entropy convince better performance of the DFUDS scheme.

Figure 8 shows the fusion result of the three sensors by the proposed scheme for three different activities, cycling in Figure 8(a), walking in Figure 8(b), and bending in Figure 8(c). The detection result of the three sensors is fused to recognized human activities. The fusion result obtained by the proposed scheme is recognized, about 0.75 percent for cycling activity while 0.18 for walking, and 0.1 percent for bending. Similarly, the fusion result is evaluated for other activities as shown in Figure 8(b) and 8(c), respectively. The simulation result of the three schemes shows that DFUDS gives more strong and correct fused result than the other methods.

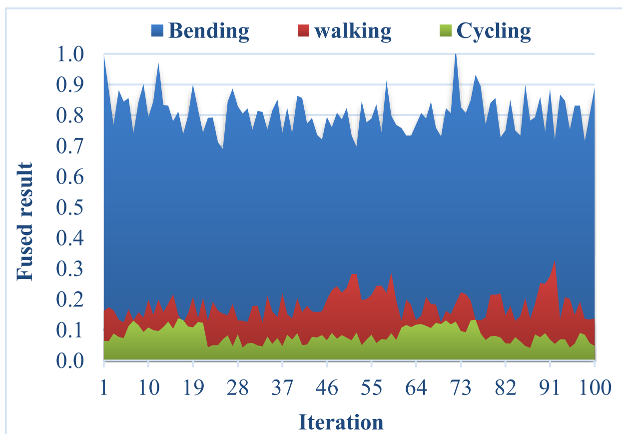
Figure 9 shows the accuracy of the proposed scheme after applying the modified belief entropy to the BoEs in the DS evidence theory. The proposed scheme provides a more accurate fused result than the Deng entropy and typical




(a) Cycling




(b) Walking




(c) Bending

**FIGURE 8. Fusion result of three activities cycling, walking and bending.**

DS evidence theory. Deng entropy and DS evidence theory do not detect the redundant values in the BoEs. The proposed scheme used the modified belief entropy to the BoEs of the sensors data before combining them. As shown in

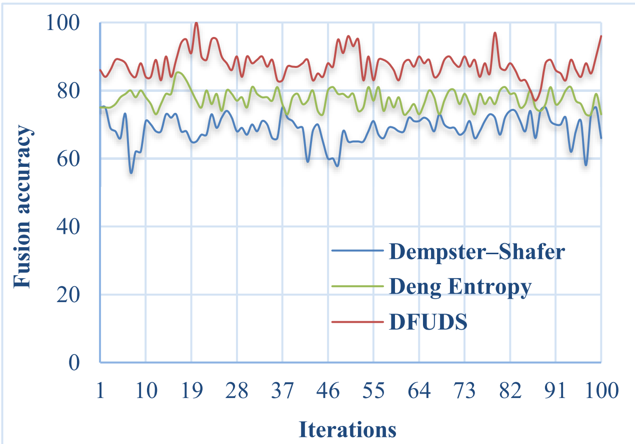



**FIGURE 9.** The comparison of fusion accuracy.

the figure if the accuracy of fusion result given by typical DS evidence theory is around 65 percent, the proposed scheme gives 25 percent better result than the DS theory and around 20 percent better result than the Deng entropy. By applying the modified belief entropy, the proposed scheme gives the efficient result as shown in the figure. It seems that the performance of the proposed method is better than the Deng entropy and DS scheme. It improves the efficiency of the system in terms of measuring the entropy of the BoEs.

## V. CONCLUSION

In this article, a new method called DFUDS is proposed for data fusion. We use belief entropy for individual sensor data as a mass-function in DS theory to quantify uncertainty and merge multi-sensor data. The modified entropy adopted by Deng entropy is used for the mass value of the sensor data and the DS theory is used to fuse it. As a result, the inference accuracy is improved and the energy consumption and network traffic can be significantly reduced. Simulation results show that the proposed method performs significantly better in terms of quantify uncertainty, inference accuracy, data fusion, and reduce redundant information.

The accuracy of the DS evidence theory for data fusion in the IoT environment can be improved if the detrimental influence of the redundant information is reduced by quantifying the uncertainty of the BoEs. In the future, the performance of the proposed scheme will be further enhanced by multisensor heterogeneous data fusion using feature selection, extraction algorithm of the machine learning approaches to select the best feature for the fusion based on the target. The proposed scheme will also be extended using artificial intelligence inference to improve inference accuracy obtained by multisensor fusion in the smart IoT environment.

## REFERENCES

[1] T. Qiu, C.-W. Tsai, C. Cecati, J. J. P. C. Rodrigues, P. Rawat, H. Song, F. Piccialli, and M. Hodkiewicz, "IEEE access special section editorial: Intelligent systems for the Internet of Things," *IEEE Access*, vol. 7, pp. 146342–146347, 2019.

[2] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How can heterogeneous Internet of Things build our future: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2011–2027, 3rd Quart., 2018.

[3] G. Liu and F. Xiao, "A data-driven dynamic data fusion method based on visibility graph and evidence theory," *IEEE Access*, vol. 7, pp. 104443–104452, 2019.

[4] I. Ullah and H. Y. Youn, "A novel data aggregation scheme based on self-organized map for WSN," *J. Supercomput.*, vol. 75, no. 7, pp. 3975–3996, 2019.

[5] I. Ullah and H. Y. Youn, "Efficient data aggregation with node clustering and extreme learning machine for WSN," *J. Supercomput.*, vol. 7, pp. 1–27, Mar. 2020.

[6] I. Ullah and H. Y. Youn, "Intelligent data fusion for smart IoT environment: A survey," *Wirel. Pers. Commun.*, vol. 114, no. 1, pp. 409–430, 2020.

[7] X. He, T. Wang, W. Liu, and T. Luo, "Measurement data fusion based on optimized weighted least-squares algorithm for multi-target tracking," *IEEE Access*, vol. 7, pp. 13901–13916, 2019.

[8] Y. Song and Y. Deng, "Divergence measure of belief function and its application in data fusion," *IEEE Access*, vol. 7, pp. 107465–107472, 2019.

[9] I. Ullah, H. Y. Youn, and Y.-H. Han, "An efficient data aggregation and outlier detection scheme based on radial basis function neural network for WSN," *J. Ambient Intell. Humanized Comput.*, pp. 1–7, Jan. 2021.

[10] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Inf. Fusion*, vol. 35, pp. 68–80, May 2017.

[11] G. Shafer, *A Mathematical Theory of Evidence*, vol. 42. Princeton, NJ, USA: Princeton Univ. Press, 1976.

[12] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, 2001.

[13] B. Sun, W. Ma, and H. Zhao, "An approach to emergency decision making based on decision-theoretic rough set over two universes," *Soft Comput.*, vol. 20, no. 9, pp. 3617–3628, Sep. 2016.

[14] U. Hohle, "Entropy with respect to plausibility measures," in *Proc. 12th IEEE Int. Symp. Multiple Valued Log.*, Paris, France, 1982, pp. 167–169.

[15] R. R. Yager, "Entropy and specificity in a mathematical theory of evidence," *Int. J. Gen. Syst.*, vol. 9, no. 4, pp. 249–260, Jan. 1983.

[16] D. Dubois and H. Prade, "A note on measures of specificity for fuzzy sets," *Int. J. Gen. Syst.*, vol. 10, no. 4, pp. 279–283, Feb. 1985.

[17] G. J. Klir and A. Ramer, "Uncertainty in the dempster-shafer theory: A critical re-examination," *Int. J. Gen. Syst.*, vol. 18, no. 2, pp. 155–166, Dec. 1990.

[18] G. J. Klir and B. Parviz, "A note on the measure of discord," in *Uncertainty in Artificial Intelligence*. Amsterdam, The Netherlands: Elsevier, 1992, pp. 138–141.

[19] Y. Deng, "Deng entropy," *Chaos, Solitons Fractals*, vol. 91, pp. 549–553, Oct. 2016.

[20] W. Jiang, B. Wei, C. Xie, and D. Zhou, "An evidential sensor fusion method in fault diagnosis," *Adv. Mech. Eng.*, vol. 8, no. 3, 2016, Art. no. 1687814016641820.

[21] K. Yuan, F. Xiao, L. Fei, B. Kang, and Y. Deng, "Conflict management based on belief function entropy in sensor fusion," *SpringerPlus*, vol. 5, no. 1, p. 638, Dec. 2016.

[22] D. Wang, J. Gao, and D. Wei, "A new belief entropy based on deng entropy," *Entropy*, vol. 21, no. 10, p. 987, Oct. 2019.

[23] D. Zhou, Y. Tang, and W. Jiang, "A modified belief entropy in Dempster-Shafer framework," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0176832.

[24] Y. Tang, D. Zhou, S. Xu, and Z. He, "A weighted belief entropy-based uncertainty measure for multi-sensor data fusion," *Sensors*, vol. 17, no. 4, p. 928, Apr. 2017.

[25] G. Chetty and M. Yamin, "A distributed smart fusion framework based on hard and soft sensors," *Int. J. Inf. Technol.*, vol. 9, no. 1, pp. 19–31, Mar. 2017.

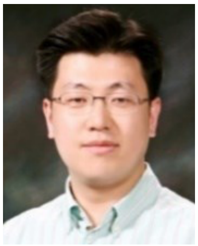
[26] Q. Xiao, Y. Zhao, and W. Huan, "Multi-sensor data fusion for sign language recognition based on dynamic Bayesian network and convolutional neural network," *Multimedia Tools Appl.*, vol. 78, no. 11, pp. 15335–15352, Jun. 2019.

[27] L. Qi, C. Hu, X. Zhang, and M. R. Khosravi, "Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment," *IEEE Trans. Ind. Informat.*, early access, Jul. 28, 2020, doi: 10.1109/TII.2020.3012157.

[28] Q. Wang, H. Liao, K. Wang, and Y. Sang, "A variable weight based fuzzy data fusion algorithm for WSN," in *Proc. Int. Conf. Ubiquitous Intell. Comput.*, 2011, pp. 490–502.
[29] A. De Paola, S. Gaglio, G. L. Re, F. Milazzo, and M. Ortolani, "Adaptive distributed outlier detection for WSNs," *IEEE Trans. Cybern.*, vol. 45, no. 5, pp. 902–913, May 2015.
[30] X. Hou, D. Zhang, and M. Zhong, "Data aggregation of wireless sensor network based on event-driven and neural network," *Chin. J. Sens. Actuators*, vol. 27, no. 1, pp. 142–148, 2014.
[31] M. Muzammal, R. Talat, A. H. Sodhro, and S. Pirbhulal, "A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks," *Inf. Fusion*, vol. 53, pp. 155–164, Jan. 2020.
[32] A. R. Pinto, C. Montez, G. Araájo, F. Vasques, and P. Portugal, "An approach to implement data fusion techniques in wireless sensor networks using genetic machine learning algorithms," *Inf. Fusion*, vol. 15, pp. 90–101, Jan. 2014.
[33] X. Liu, R. Zhu, A. Anjum, J. Wang, H. Zhang, and M. Ma, "Intelligent data fusion algorithm based on hybrid delay-aware adaptive clustering in wireless sensor networks," *Future Gener. Comput. Syst.*, vol. 104, pp. 1–14, Mar. 2020.
[34] K. Liu, T. Yang, J. Ma, and Z. Cheng, "Fault-tolerant event detection in wireless sensor networks using evidence theory," *KSII Trans. Internet Inf. Syst.*, vol. 9, no. 10, pp. 3965–3982, 2015.
[35] S. M. Alqhtani, S. Luo, and B. Regan, "Multimedia data fusion for event detection in Twitter by using Dempster-Shafer evidence theory," *Int. J. Comput. Electr. Autom. Control Inf. Eng. World Acad. Sci. Eng. Technol.*, vol. 9, no. 12, pp. 2234–2238, 2015.
[36] N. Nesa and I. Banerjee, "IoT-based sensor data fusion for occupancy sensing using Dempster–Shafer evidence theory for smart buildings," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1563–1570, Oct. 2017.
[37] M. R. Senouci, A. Mellouk, N. Aitsaadi, and L. Oukhellou, "Fusion-based surveillance WSN deployment using Dempster–Shafer theory," *J. Netw. Comput. Appl.*, vol. 64, pp. 154–166, Apr. 2016.
[38] D. Zhang, J. Cao, J. Zhou, and M. Guo, "Extended dempster-shafer theory in context reasoning for ubiquitous computing environments," in *Proc. Int. Conf. Comput. Sci. Eng.*, 2009, pp. 205–212.
[39] T. George and N. R. Pal, "Quantification of conflict in Dempster-Shafer framework: A new approach," *Int. J. Gen. Syst.*, vol. 24, no. 4, pp. 407–423, Mar. 1996.
[40] S. H. Hussein. (2016). *Python Library for Measuring Uncertainty in Dempster-Shafer (Pyuds). Contribute to Angyjoe/Pyuds Development*. [Online]. Available: https://github.com/angyjoe/pyuds
[41] F. Palumbo, C. Gallicchio, R. Pucci, and A. Micheli, "Human activity recognition using multisensor data fusion based on reservoir computing," *J. Ambient Intell. Smart Environments*, vol. 8, no. 2, pp. 87–107, Mar. 2016.
[42] C. K. Murphy, "Combining belief functions when evidence conflicts," *Decis. Support Syst.*, vol. 29, no. 1, pp. 1–9, Jul. 2000.

**IHSAN ULLAH** received the B.S. and M.S. degrees in computer science from the University of Peshawar, Pakistan, in 2001 and 2004, respectively, and the Ph.D. degree in computer engineering from Sungkyunkwan University, Suwon, South Korea, in 2019. From September 2019 to August 2020, he was a Postdoctoral Research Fellow with the Ubiquitous Computing Technology Research Institute (UTRI), Sungkyunkwan University. Since 2020, he has been a Research Professor with the School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan, South Korea. His research interests include data aggregation, data fusion, virtual network embedding, network slicing (5G), the IoT (Internet of Things), artificial intelligence, machine learning, cloud computing, and wireless sensor networks.

**JOOSANG YOUN** received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, Korea University, Seoul, South Korea, in 2001, 2003, and 2008, respectively. In 2002, he was a Researcher with the Next Generation Protocol Team, Electronics and Telecommunications Research Institute. Since 2008, he has been a Professor with the Department of Industrial ICT Engineering, Dong-Eui University, Busan, South Korea. His primary research interest includes theory and application for the IoT, including protocol design and performance analysis. Since 2010, his research interests have been focusing on the Internet of Things, edge intelligence, edge-cloud computing, and future network architecture. He has also made several contributions to ITU-T, ISO/JTC 1, and IETF standardization.



**YOUN-HEE HAN** received the B.S. degree in mathematics, and the M.S. and Ph.D. degrees in computer science and engineering from Korea University, Seoul, South Korea, in 1996, 1998, and 2002, respectively. From 2002 to 2006, he was a Senior Researcher with the Next Generation Network Group, Samsung Advanced Institute of Technology. Since 2006, he has been a Professor with the School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan, South Korea. Since 2002, his activities have been focusing on mobility management, media independent handover, and cross-layer optimization for efficient mobility support. He has published approximately 200 research articles on the theory and application of mobile computing, and holds 30 patents on information and communication technology domain. His current research interests include theory and application of mobile computing, including protocol design and mathematical analysis, mobile sensor/actuator networks, social network analysis, machine learning, and deep learning. He has made several contributions in IETF and IEEE standardization, and served as the Co-Chair of the working group at the Korea TTA IPv6 Project Group. Since 2011, he has been serving as an Editor for the *Journal of Information Processing Systems*.