Research on Information Fusion for Machine Potential Fault Operation and Maintenance

Wei Xu 1, Yi Wan 2,3,*, Tian-Yu Zuo 1 and Xin-Mei Sha 1

1 School of Mechanical and Electrical Engineering, Sanjiang University, Nanjing 210012, China; xwcontrol@163.com (W.X.); tian_yu_zuo@163.com (T.-Y.Z.); cwaq.sxm@126.com (X.-M.S.)
2 School of Environmental Science, NanJing XiaoZhuang University, Nanjing 211171, China
3 Jiangsu Shentong Valve Co. Ltd., Nantong 226232, China
* Correspondence: wan_sju@163.com

Received: 8 February 2020; Accepted: 14 February 2020; Published: 3 March 2020

Abstract: In recent years, the development of sensor technology in industry has profoundly changed the way of operation and management in manufacturing enterprises. Due to the popularization and promotion of sensors, the maintenance of machines on the production line are also changing from the subjective experience-based machine maintenance to objective data-driven maintenance decision-making. Therefore, more and more data decision model has been developed through AI technology and intelligence algorithms. Equally important, the information fusion between decision results, which got by data decision model, has also received widespread attention. Information fusion is performed on symmetric data structures. The asymmetric data under the symmetric structure leads to the difference in information fusion results. Therefore, fully considering the potential differences of asymmetric data under a symmetric structure is an important content of information fusion. In view of the above, this paper studies how to make information fusion between different decision results through the framework of D-S evidence theory and discusses the deficiency of D-S evidence theory in detail. Based on D-S evidence theory, then a comprehensive evidence method for information fusion is proposed in this paper. We illustrate the rationality and effectiveness of our method through analysis of experiment case. And, this method is applied to a real case from industry. Finally, the irrationality of the traditional D-S method in the comprehensive decision-making results of machine operation and maintenance was solved by our novel method.

Keywords: sensor; information fusion; operation and maintenance; decision-making; D–S evidence theory

1. Introduction

With the promotion of intelligent manufacturing, more and more sensors are being used in the inspection process [1,2]. These collected data are not only conducive to the real-time detection of the machine, but also to the timely detection of potential fault problems of the machine to ensure the safety and reliability of manufacturing. As a new generation of industrial paradigm, industry 4.0 explains the digital competition that industrial enterprises are facing. Vertical integration requires digitalization of the manufacturing process within the enterprise. Horizontal integration requires the enterprise to realize product traceability throughout the supply chain, and end-to-end integration requires seamless connection of the whole industry. Therefore, as the underlying physical technology of digital transformation, sensor technology has achieved unprecedented development in recent years. In particular, it is well known that sensors can bring rich data to manufacturing. Such richness is reflected not only in the quantity of data, but also in the variety of data. Nowadays, through the establishment of a manufacturing execution system (MES) [3], a large number of enterprises have been
able to obtain data regarding machine detection, the manufacturing process, and product quality at the speed of seconds. Furthermore, due to the development of Artificial Intelligence technology, big data analysis has achieved unprecedented progress and expansion [4–6]. AI-based decision-making support [7] is also becoming increasingly popular in the operation and maintenance of manufacturing systems such as production scheduling [8], stock management [9], preventive maintenance [10], etc. Although decision-making support systems are more and more popular, it is worth noting that most enterprise support decisions today are mainly based on the analysis results of a single data source. Such decisions result in a large amount of data not actually being used to maintain the manufacturing systems.

In fact, the key issues here include two points. On one hand, the value of big data is not really mined and utilized. Generally, when enterprises implement AI-based support decisions, they always want to cover as much data as possible by deploying a single AI technology to be able to quickly complete the decision system in only one development. This kind of development approach will reduce the development costs, however, this kind of development mode often mixes different types of data, making the data interpretation poor. In some cases, in order to improve the interpretability of data, the development process chooses to consider only part of the data, which the decision system is unable to use. On the other hand, if we develop multiple models for different data in one decision system, due to the one-sided observation of different data on the manufacturing system, different models may give different results. Information fusion and contradiction neutralization between models will soon become a thorny matter. The solution to the first problem is relatively easy. In practice, different data can be sorted according to the manufacturing process, and then multiple models can be built for different data, so the data value can be mined from multiple aspects. Compared with the first problem, there have been few studies on information fusion and contradiction of multiple models for decision support.

Therefore, based on the consideration of the above problems, this paper proposed a multi-source data information fusion method to support maintenance decisions for potential machine failures. By introducing the Dempster–Shafer (D–S) evidence theory and taking the machine as the problem identification framework, we developed an improved D–S evidence method to resolve the contradiction between the decision results generated from the different AI-based decision models to support the machine’s early potential failure maintenance through our comprehensive evaluation.

The remainder of this paper is organized as follows. Section 2 introduces the current mainstream information fusion methods in industrial scenarios and analyzes their advantages and disadvantages. Then, D–S evidence theory and its improvements are introduced in Section 3 to support subsequent information fusion. Section 4 uses a numerical case to carry out specific calculations and demonstrate the effectiveness of the proposed improved D–S method through comparative analysis. A case of a real industrial operation and maintenance is given in Section 5. Through the analysis of the case, the advantages and feasibility of our method are further explained. Finally, the summary is presented in Section 6.

2. Related Work

The main purpose of information fusion is to integrate multi-source homogeneous or heterogeneous information to obtain a comprehensive information evaluation [11]. At present, there are a variety of information fusion strategies and methods developed by scholars including Bayesian inferences [12], fuzzy reasoning [13,14], D–S evidence theory [15–17], and the neural network method [18]. The comparison between fusion methods is shown in Table 1 as follow.
Bayesian inferences is a method of reasoning using conditional probability. Generally, Bayesian inferences need to obtain the prior probability distribution to reason on the posterior probability, and then makes a decision. However, for industrial scenarios, especially in the maintenance process, it is usually accompanied by the continuous adjustment and maintenance of the machine. Therefore, the data collected by sensors will change due to the adjustment of the machines. Therefore, the decision results often become inaccurate due to changes in the prior probability distribution.

The core of fuzzy reasoning is the extraction of uncertain information in the process of information fusion, then the analysis is carried out by the uncertainty. How to design the rules of fuzzy reasoning is the key issue of fuzzy reasoning. Different manufacturing systems have different structures and settings, so it is generally necessary to design fuzzy reasoning rules according to specific structural and settings, which limits this kind of reasoning method being applied in many heterogeneous manufacturing systems.

The neural network method establishes input–output relationships based on the connection between nodes. In general, the input of the neural network must be homogeneous, and it needs a lot of data and time for training to guarantee the rationality of its reasoning results. The heterogeneous data in industry severely restricts the application of a neural network as an information fusion strategy.

D–S evidence theory is a classic probability-derived method. Compared with other processing methods, D–S evidence theory can effectively deal with uncertain information in unknown environments. Therefore, it is widely used in fault diagnosis, multi-standard decision-making, and other fields. However, the shortcomings of D–S evidence theory are: (1) it is difficult to handle the fusion of extremely contradictory information; (2) the reliability of the evidence itself is not deeply considered; and (3) the lack of consideration of the consistency between the evidences. This paper will focus on these issues in Section 3 and give an improvement strategy to enhance the rationality of D–S evidence theory in order to apply it to practical industrial problems.
3. Information Fusion Based on D–S Evidence

3.1. D–S Evidence Theory

D–S evidence theory was first proposed by Dempster and then developed by Shafer. The basic elements are: (1) define the problem identification framework \( \theta = \{w_1, w_2, \cdots, w_c\} \), and (2) set the basic probability assignment function (BPA) \( m : 2^\theta \rightarrow [0, 1] \), which meets the following conditions:

\[
m(\emptyset) = 0 \quad \text{and} \quad \sum_{A \subseteq \theta} m(A) = 1
\]  

where \( A \) is a subset of \( 2^\theta \), and \( m(A) \) represents the support the degree of evidence for proposition \( A \). Each proposition has a corresponding BPA. The problem probability outside the problem recognition framework is 0, and the sum of the problem probabilities in the framework is 1. Based on the above, for \( \forall A \subseteq \theta, A \neq \emptyset \), the multiple evidence fusion rules are defined as:

\[
m(A) = \frac{1}{K} \sum_{\bigcap A_j = A} \prod_{i=1}^{n} m_i(A_j)
\]

where

\[
K = \sum_{\bigcap A_j \neq \emptyset} \prod_{i=1}^{n} m_i(A_j)
\]

\( K \) is the normalization coefficient, which reflects the contradiction degree between multiple evidences. D–S evidence fusion rules can model uncertain and inaccurate data without prior knowledge of the prior probability [45], and can effectively fuse different evidence from different data sources to obtain a more accurate comprehensive evaluation result.

3.2. Lack of D–S Evidence Theory

D–S evidence theory has at least three serious problems of insufficient evidence fusion. 

1. The error in the fusion of extreme evidence.

Zadeh et al. questioned D–S evidence theory through an example [46]: There are three suspects A, B, and C in a murder, and Witnesses 1 and 2 give the corresponding BPA, as shown in Table 2.

| Witness 1 | Witness 2 | Fusion Result |
|-----------|-----------|---------------|
| A         | 0.99      | 0.00          |
| B         | 0.01      | 0.01          |
| C         | 0.00      | 0.99          |

Table 2. Zadeh’s Paradox.

Table 2 shows the result of information fusion with two witnesses. Through the calculation process of D–S evidence theory, it can be found that if a witness has cast an extreme probability distribution on a certain suspect, no matter how other witnesses make decisions, the overall degree of suspect of this suspect will be very low (even 0). Such conflict cannot be resolved in the classic D–S evidence theory.

2. The reliability of the evidence itself is not considered.

In many decision-making processes, people will have two subjective tendencies to the decision results based on their self-cognition. People who have a good grasp of the decision-making events usually have an obvious tendency to make a probability distribution, while people who only have a
one-sided understanding of the decision-making events usually make a more balanced probability distribution. This leads to differences in reliability between different evidences.

For example, the BPA is presented in Table 3.

| Table 3. Impact of reliability on evidence. |
|-------------------------------------------|
| Witness 1 | Witness 2 |
| A       | 0.60 | 0.60 |
| B       | 0.40 | 0.20 |
| C       | 0.00 | 0.20 |

Witness 1 clearly believes that C cannot be a suspect. Such a BPA helps to narrow the suspects to A and B. If we assume that Witness 1 is a rational witness, such a BPA would be advantageous for the exclusion of suspect C. For the BPA of Witness 2, although A is also considered to be the most suspicious, it is not helpful for the discrimination between B and C. Similarly, it is assumed that Witness 2 is also a rational person. Due to the thinking habits of a rational person, Witness 2 identified that A had a greater possibility of committing the crime, according to their observations. However, to judge B and C, the lack of observations limited them in being unable to make a judgment, so it was considered to give B and C the same BPA. From the perspective of rational man, we think that the judgment from Witness 1 is more valuable because the BPA of Witness 1 was more targeted, which means that they have more confidence in their result, so the reliability of evidence is stronger. However, in the classical D–S evidence theory, the reliability of evidence is not directly considered.

- 3. The relationship between the evidence is not considered in depth.

D–S evidence theory mainly considers the combination of evidence. Such a combination method is often multiplication and accumulation of evidence without distinction. The disadvantage is that the inherent relationship between the evidence cannot be extracted.

For example, see Table 4.

| Table 4. Impact of evidence correlation on evidence. |
|-----------------------------------------------|
| Witness 1 | Witness 2 | Witness 3 | Witness 4 | D–S Fusion Result |
| A       | 0.90     | 0.95     | 0.95     | 0.00     | 0.00     |
| B       | 0.09     | 0.04     | 0.03     | 0.03     | 0.63     |
| C       | 0.01     | 0.01     | 0.02     | 0.97     | 0.37     |

Intuitively, it can be seen that Witness 1, Witness 2, and Witness 3 all mainly pointed the crime at A, but Witness 4 thought that A was unlikely to commit a crime, and C was most likely to commit a crime. From a rational perspective, the result of information fusion should be A with the highest probability, followed by B and C. Since C has only one piece of evidence, and it is obviously inconsistent with other evidence, the credibility of its evidence should be the lowest and should be eliminated during information fusion. Moreover, in order to ensure the reasonableness of the fusion, A should be guaranteed to obtain the highest possible fusion probability so that A can be identified. However, from the results of the current fusion of D–S evidence, it is impossible to avoid the fusion error caused by the inconsistency of Witness 4.

3.3. Improved Evidence Theory

The following assumes that an information fusion problem with \( n \) propositions and \( m \) evidences is considered.
3.3.1. Weight of Evidence Reliability Based on Entropy

To address the issue of evidence reliability, we introduced entropy to characterize. Entropy \cite{47,48} can measure the uncertainty of information effectively. Assume that the proposition set \( A = \{A_1, \ldots, A_i, \ldots, A_n\} \) that needs to be evaluated from evidence set \( Ev = \{Ev_1, \ldots, Ev_k, \ldots, Ev_m\} \), and the corresponding BPA function from \( Ev_k \) is \( m_k(A_1), \ldots, m_k(A_i), \ldots, m_k(A_n) \), then the entropy of \( Ev_k \) is defined as:

\[
\text{Entropy}_k = -\sum_{i=1}^{n} m_k(A_i) \log m_k(A_i) \tag{4}
\]

Equation (4) indicates that if the \( \text{Entropy}_k \) of evidence \( k \) is lower, the dispersion of BPA information is lower, and the reliability of this evidence is higher.

Since entropy is inversely proportional to reliability, Equation (5) can be used to regularize the entropy of all evidence to determine the weights corresponding to \( m_k \) of different evidences,

\[
\omega_k = \frac{\text{Entropy}_k^{-1}}{\sum_{k=1}^{m} \text{Entropy}_k^{-1}} \tag{5}
\]

where \( \omega_k \) is the weight of the evidence \( m_k \).

3.3.2. Weight of Evidence Consistency Based on Evidence Correlation Matrix

Considering the relationship between the evidences, the evidence correlation matrix \( C_{mm} \) is established by Equation (6).

\[
C = \{C_{i,j}\} = \left\{ \begin{array}{c}
\sum_{p=1}^{n} m_i(A_p)m_j(A_p), \ i \neq j \\
1, \ i = j
\end{array} \right\} \tag{6}
\]

By Equation (6), the similarity between different evidences can be obtained and the correlation degree between different evidences can be obtained. Furthermore, the consistency between evidence \( k \) and other evidences (except \( k \)) can be obtained by Equation (7).

\[
\text{cons}_k = \prod_{i=1}^{m} C_{k,i} \tag{7}
\]

For the sake of normalization, the consistency of all evidence should be normalized. For convenience, the normalized result represented by \( \gamma_k \), that is, the weight of evidence consistency.

\[
\gamma_k = \frac{\text{cons}_k}{\sum_{k=1}^{m} \text{cons}_k} \tag{8}
\]

3.3.3. Reference Evidence Generation

According to the weights obtained from Sections 3.3.1 and 3.3.2, the comprehensive reference weight \( \alpha_k \) is defined as follows:

\[
\alpha_k = \omega_k \gamma_k \tag{9}
\]

The reference evidence is defined as

\[
r(A) = \sum_{k=1}^{m} \alpha_k m_k(A) \tag{10}
\]

The evidence reliability weight and the evidence consistency weight are mainly used for the contradictory evidence with low reliability and poor consistency. However, for the fusion of highly consistent and reliable evidence, D-S evidence theory has its advantages. Therefore, we proposed that
D–S evidence theory and reference evidence should be considered together to obtain the comprehensive fusion rule.

3.3.4. Comprehensive Evidence Fusion Rule

In the D–S evidence, the standardization coefficient $K$ ($0 \leq K \leq 1$) is used to judge the degree of contradiction. The contradiction is intense when $K$ approaches 0, and all evidences are consistent when $K$ approaches 1. With this coefficient, we designed the comprehensive evidence fusion rule, which is defined as

$$m(A) = K \times d(A) + (1 - K) \times r(A) \quad (11)$$

where

$$d(A) = \frac{1}{K} \sum_{\bigcap A_i = A} \prod_{k=1}^{m} m_k(A_i) \quad (12)$$

d($A$) represents the fusion result from the D–S evidence theory, and r($A$) represents the reference evidence defined by Equation (10). When $K \to 1$, that is, the contradiction is slight, the fusion result is close to the D–S evidence. When $K \to 0$, that is, when the contradiction is serious, the fusion result is close to the reference evidence.

4. Experiments

We used the data in Table 4 as an example for calculation and analysis.

4.1. Reliability Weight Calculation

First, calculate the reliability entropy corresponding to each evidence by Equation (4), as shown in Table 5.

| $E_1$ | $E_2$ | $E_3$ | $E_4$ |
|------|------|------|------|
| 0.1553 | 0.0971 | 0.1008 | 0.0585 |

Then, the reliability weights can be obtained by Equation (5), as shown in Table 6.

| $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ |
|----------|----------|----------|----------|
| 0.2076   | 0.2547   | 0.2517   | 0.2860   |

4.2. Consistency Weight Calculation

Calculate the evidence correlation matrix $C$ by Equation (6).

$$C = \begin{bmatrix}
1 & 0.8587 & 0.8579 & 0.0124 \\
0.8587 & 1 & 0.9039 & 0.0109 \\
0.8579 & 0.9039 & 1 & 0.0203 \\
0.0124 & 0.0109 & 0.0203 & 1 \\
\end{bmatrix}$$

Then, use Equations (7) and (8) to calculate the consistency weights of the evidences, which is shown in Table 7.
4.3. Reference Evidence Calculation

The reference evidence has two parts, one is the reliability of a single evidence itself, and the other is the consistency between multiple evidences. For subsequent comparative analysis, here we used Equations (9) and (10) to generate three groups of reference evidence, which is shown in Table 8. The first group was generated only by the reliability weights (set all the consistency weights as 1), denoted as \( r_E \). The second group was generated only by the consistency weights (set all the consistency weights as 1), denoted as \( r_C \). The third group was generated by the reliability weights and consistency weights, denoted as \( r \).

Table 8. The calculation results of the three groups of reference evidence.

| Reference Evidence | A   | B   | C   |
|--------------------|-----|-----|-----|
| \( r_E \)          | 0.6680 | 0.0450 | 0.2870 |
| \( r_C \)          | 0.9362 | 0.0490 | 0.0148 |
| \( r \)            | 0.9381 | 0.0469 | 0.0150 |

4.4. Comprehensive Evidence Fusion Generation

The final fusion result can be obtained by comprehensive evidence fusion using Equation (11), as shown in Table 9. For comparative analysis, Table 9 shows the results of D–S evidence, the results generated based on reliability weights, the results generated based on consistency weights, and the results generated based on reliability weights and consistency weights.

Table 9. Comprehensive evidence fusion results.

| Evidence Fusion Result | A   | B   | C   |
|------------------------|-----|-----|-----|
| \( D - S \)            | 0.0000 | 0.6255 | 0.3745 |
| \( D - S - r_E \)      | 0.6680 | 0.0450 | 0.2870 |
| \( D - S - r_C \)      | 0.9362 | 0.0490 | 0.0148 |
| \( D - S - r \)        | 0.9381 | 0.0469 | 0.0150 |

4.5. Results Analysis

From Figure 1, we can clearly see that for the D–S method, since Witness4 assigned an extreme probability of 0 to \( A \), regardless of the probability that other people allocate, the criminal suspicion of \( A \) will always be excluded. This is contrary to the normal inference logic, which is obviously unreasonable. The final fusion results given in Table 9 show that \( D - S - r_E \), \( D - S - r_C \), and \( D - S - r \) all overcame the problems of D–S evidence theory under extreme contradiction, and they all gave \( A \) a larger evaluation weight. Compared with \( D S - r_E \), \( D S - r_C \) and \( D S - r \), it can be seen that \( D S - r_E \) assigned a larger evaluation weight to \( A \). Additionally, \( D S - r_E \) gave a large weight to \( C \) compared to \( B \). This is obviously not reasonable. \( D S - r_C \) and \( D S - r \) both performed better. Both gave \( A \) with an absolutely large weight, and in the comparison between \( B \) and \( C \), they were more inclined to suppress the distribution changes caused by abnormal evidence. As a result, their probability assignments tended to support more consistent evidence. In addition, \( D S - r \) assigned a larger weight to \( A \) than \( D S - r_C \), and assigned a smaller weight to \( C \), which shows that this comprehensive method can increase the dispersion of the weight value, producing results that have a greater tendency in discrimination. In particular, for decision makers who do not understand the internal process of fusion, they can make decisions more confidently based on the fusion results.
5. Application and Verification

5.1. Problem Description

Industrial manufacturing lines need to face many important operation and management decisions, among which machine maintenance decision is an important part of product quality assurance in the manufacturing process. Since around 2000, a large number of manufacturing enterprises realized the importance of manufacturing information, so have carried out the transformation of manufacturing system information and digitization. Therefore, more machine information and manufacturing execution information can be quickly captured. In the last decade, the popular AI method has been promoted to provide intelligent methods for the analysis of manufacturing information to improve the effect of decision support. Sun et al. [49] presented the machine importance information and machine alarm information of an electronic manufacturing enterprise analyzed by intelligent methods for 10 different machines. Specifically, the objective alarm data were clustered and mapped to the potential failure probability of the machine, and the machine importance was analyzed by the fuzzy method in [49]. Next, we will conduct information fusion for these 10 different machines with our improved D–S evidence and analyze the fusion results.

The specific methods are as follows:

5.2. Establish Problem Identification Framework

Ten machines in a production line were used as the identification framework of D–S evidence theory, as follows.

\[ \theta = \{ \text{machine}_1, \text{machine}_2, \text{machine}_3, \text{machine}_4, \text{machine}_5, \]
\[ \text{machine}_6, \text{machine}_7, \text{machine}_8, \text{machine}_9, \text{machine}_{10} \]

Table 10 shows the suspected failure probability of 10 machines derived from machine importance information, and Table 11 shows the suspected failure probability of machines derived from the machine alarm.
According to the probability of machine suspected fault obtained from the machine importance and machine alarm, information fusion is needed to guide the resource allocation under the limited operation and maintenance resources.

### 5.3. Suspected Failure Probability Of Machines

According to the comprehensive evidence fusion method proposed in this paper, the suspected failure probability of the comprehensive machine was calculated, as shown in Table 12.

**Table 12.** The suspected failure probability of machines derived from the comprehensive evidence fusion method.

| Machine_Order | Probability | Machine_Order | Probability |
|---------------|-------------|---------------|-------------|
| 1             | 0.2958      | 6             | 0.0303      |
| 2             | 0.2385      | 7             | 0.0158      |
| 3             | 0.1530      | 8             | 0.0755      |
| 4             | 0.0438      | 9             | 0.0085      |
| 5             | 0.1175      | 10            | 0.0214      |

According to the D–S evidence, the suspected failure probability of the machines is shown in Table 13.

**Table 13.** The suspected failure probability of machines derived from D-S evidence method.

| Machine_Order | Probability | Machine_Order | Probability |
|---------------|-------------|---------------|-------------|
| 1             | 0.5149      | 6             | 0.005       |
| 2             | 0.2826      | 7             | 0.001       |
| 3             | 0.1377      | 8             | 0.008       |
| 4             | 0.0010      | 9             | 0.000       |
| 5             | 0.0486      | 10            | 0.000       |

### 5.4. Comparative Analysis

From the perspective of Figure 2, we can see that the first two machines, machine 1 and machine 2, both achieved a large probability in both methods. Therefore, both the D–S evidence and the comprehensive evidence fusion method defined by us could recognize machine 1 and machine 2. However, in the D-S evidence method, there was a great difference between machine 1 and machine 2 with the comprehensive probability values. In absolute terms, the probability of machine 3 was even
closer to that of machine 4. For the D–S evidence, the probability value of machine 1 was particularly large. Therefore, from the perspective of maintenance, it will tend to invest more resources on machine 1, which will have an adverse impact on the maintenance of other machines.

![Figure 2. Probability comparison before and after information fusion.](image)

Moreover, it can be observed in Figure 2 that the machine alarm probability of machine 5 exceeded 20%, but the probability of machine importance was less than 1%. This contradiction led to the D–S evidence giving a particularly small probability to machine 5. Since there were only two pieces of evidence in this case, the result of using the comprehensive evidence fusion method to calculate the evidence correlation will be that both evidences were 0.5. That is, the importance of the two pieces of evidence cannot be distinguished. Therefore, the comprehensive evidence fusion method tends to guarantee the rationality of the fusion result by means of probability compromise. For the case given in Section 4, the effect of the last piece of evidence was minimized because there were multiple pieces of evidence that contradicted the last piece of evidence. Therefore, it can be seen that our comprehensive evidence fusion method can automatically adjust in the actual calculation, according to the amount of evidence and the contradiction of evidence.

In summary, in this practical case, the comprehensive evidence fusion method overcomes the extreme probability distribution of D–S. Moreover, it fully considers the relationship between multiple evidences.

**Table 14.** Fusion result comparison with [49].

| Machine_Order | Comprehensive Evidence Fusion | Result From [49] | Normalization Result Form [49] |
|---------------|-------------------------------|------------------|--------------------------------|
| 1             | 0.3349                        | 0.2890           | 0.3349                         |
| 2             | 0.2413                        | 0.2083           | 0.2413                         |
| 3             | 0.1488                        | 0.1285           | 0.1488                         |
| 4             | 0.0332                        | 0.0286           | 0.0332                         |
| 5             | 0.1100                        | 0.0950           | 0.1100                         |
| 6             | 0.0250                        | 0.0216           | 0.0250                         |
| 7             | 0.0135                        | 0.0116           | 0.0135                         |
| 8             | 0.0674                        | 0.0582           | 0.0674                         |
| 9             | 0.0072                        | 0.0062           | 0.0072                         |
| 10            | 0.0187                        | 0.0162           | 0.0187                         |
In Table 14, we also compare our fusion results with those in [49]. First, the sum of the probabilities of the 10 machines in [49] was not 1, but 0.8631. Therefore, we also give the normalization results of [49].

In view of the results of non-normalization, it will lead to a low overall probability, which is not good for intuitive visualization. Additionally, if we look at machine 3, the two probabilities that machine 3 needs to fuse are 0.1713 and 0.1315 (see Tables 10 and 11), respectively. As it is only the fusion of two pieces of evidence, it is impossible to judge the quality of the evidence itself by the correlation between the pieces of evidence, so a reasonable probability result should be between these two probability values. However, the probability given by [49] was 0.1285 (see Table 14), which is lower than the two evidences, which we think is unreasonable.

Considering the results of normalization, there is a similar problem. The two probabilities that observations from machine 1 need to fuse were 0.3309 and 0.2547, respectively. The fusion probability results from normalization currently given was 0.3349, which is higher than the two evidences, which is unreasonable.

As we can see from our comprehensive fusion method, as shown in Figure 2, the overall probability sum was normalized, which is good for intuitive visualization. Furthermore, the probability from our fusion method is between the probability values of the two that need to be fused, which is reasonable.

6. Conclusions

This paper compares the advantages and disadvantages of popular information fusion technology in the manufacturing industry, and holds that D–S evidence theory has a broad prospect in manufacturing operation and management. Through the discussion of the shortcomings of D–S evidence theory, an information fusion method was proposed based on D–S evidence theory. Then, through the analysis and comparison of numerical cases, it was verified that the improved method proposed in this paper had a better information fusion effect. Moreover, this paper also applied this method to the real machine operation and maintenance case in the industry, and illustrates the feasibility and practicability of this method through comparative analysis. We noticed that the manufacturing industry is facing the explosive growth of data at present, and manufacturing will also be data driven in the near future. Effective and fast information fusion will be a key work in decision support systems, which deserves to be studied in depth. We also noted that in many industrial cases, information perception exists at different levels of industrial manufacturing and it is difficult to directly integrate. Therefore, information fusion should be carried out at different levels, and it is more necessary to develop corresponding methods to link the analysis results at all levels. Combining the Analytic Hierarchy Process (AHP) with information fusion and improving the interpretability of information will be the focus of future research work.

Author Contributions: All authors contributed equally and significantly in writing this paper. All authors have read and agreed to the published version of the manuscript.

Funding: The author would like to thank the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant Nos. 19KJB460006, 17KJB460011) and the China Postdoctoral Science Foundation (Grant No. 2019M651642) for the funding support to this research.

Acknowledgments: The authors acknowledge Sanjiang University and Nanjing Xiaozhuang University.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Mandroli, S.S.; Shrivastava, A.K.; Ding, Y. A survey of inspection strategy and sensor distribution studies in discrete-part manufacturing processes. *IIE Trans.* 2006, 38, 309–328. [CrossRef]
2. Sun, S.; Lin, H.; Ma, J.; Li, X. Multi-sensor distributed fusion estimation with applications in networked systems: A review paper. *Inf. Fusion* 2017, 38, 122–134. [CrossRef]
3. Kletti, J. (Ed.) Manufacturing Execution System-MES; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
4. Russell, S.; Norvig, P. Artificial Intelligence: A Modern Approach; Pearson Education Limited: London, UK, 2016.
5. Chen, M.; Mao, S.; Liu, Y. Big data: A survey. Mob. Netw. Appl. 2014, 19, 171–209. [CrossRef]
6. Russom, P. Big data analytics. TDWI Best Pract. Rep. Fourth Quart. 2011, 19, 1–34.
7. Acknson, S.; Brownie, J. AI-based decision support tool for strategic decision making in the factory of the future. Comput. Integr. Manuf. Syst. 1992, 5, 83–90. [CrossRef]
8. Kusiak, A.; Ahn, J. Intelligent scheduling of automated machining systems. Comput. Integr. Manuf. Syst. 1992, 5, 3–14. [CrossRef]
9. Dağlı, C.H. Knowledge-based systems for cutting stock problems. Eur. Journ. Oper. Res. 1990, 44, 160–166. [CrossRef]
10. Confalonieri, M.; Barni, A.; Valente, A.; Cinus, M.; Pedrazzoli, P. An AI based decision support system for preventive maintenance and production optimization in energy intensive manufacturing plants. In Proceedings of the 2015 IEEE International Conference on Engineering, Technology and Innovation/International Technology Management Conference (ICE/ITMC), Belfast, UK, 22–24 June 2015; pp. 1–8.
11. Tang, Y.; Zhou, D.; Xu, S.; He, Z. A Weighted Belief Entropy-Based Uncertainty Measure for Multi-Sensor Data Fusion. Sensors 2017, 17, 928. [CrossRef]
12. Box, G.E.; Tiao, G.C. Bayesian Inference in Statistical Analysis; John Wiley & Sons: Hoboken, NJ, USA, 2011.
13. Takagi, H.; Hayashi, I. NN-driven fuzzy reasoning. Int. Journ. Approx. Reason. 1991, 5, 191–212. [CrossRef]
14. Mizumoto, M.; Zimmermann, H. Comparison of fuzzy reasoning methods. Fuzzy Sets Syst. 1982, 8, 253–283. [CrossRef]
15. Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. In Classic Works of the Dempster-Shafer Theory of Belief Functions; Springer: Berlin/Heidelberg, Germany, 2008; pp. 57–72.
16. Shafer, G. A Mathematical Theory of Evidence; Princeton University Press: Princeton, NJ, USA, 1976.
17. Barnett, A. Computational methods for a mathematical theory of evidence. In Classic Works of the Dempster-Shafer Theory of Belief Functions; Springer: Berlin/Heidelberg, Germany, 2008; pp. 197–216.
18. Carpenter, G.A.; Martens, S.; Ogas, O. Self-organizing information fusion and hierarchical knowledge discovery: a new framework using ARTMAP neural networks. Neural Netw. 2005, 18, 287–295. [CrossRef]
19. Yuan, T.; Kuo, W. Spatial defect pattern recognition on semiconductor wafers using model-based clustering and Bayesian inference. Eur. Journ. Oper. Res. 2008, 190, 228–240. [CrossRef]
20. Zheng, Y.; Qin, S.; Wang, F. PLS-based similarity analysis for mode identification in multimode manufacturing processes. IFAC-Pap. 2015, 48, 777–782. [CrossRef]
21. Khakifirooz, M.; Chien, C.F.; Chen, Y. Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower industry 4.0. Appl. Soft Comput. 2018, 68, 990–999. [CrossRef]
22. Kozan, E.; Köksoy, O. A bayesian parameter estimation approach to response surface optimization in quality engineering. Sak. Universitesi Fen Bilimleri Enstitüısı Derg. 2019, 23, 767–774. [CrossRef]
23. Aminzadeh, M.; Kurfess, T.R. Online quality inspection using Bayesian classification in powder-bed additive manufacturing from high-resolution visual camera images. Journ. Intell. Manuf. 2019, 30, 2505–2523. [CrossRef]
24. Ji, W.; Li, Y.; AbouRizk, S.M. Integrated data-driven approach for analyzing pipe welding operator-quality performance. Autom. Constr. 2019, 106, 102814. [CrossRef]
25. Lin, L.; Lin, C.L. The use of grey-fuzzy logic for the optimization of the manufacturing process. Journ. Mater. Process. Technol. 2005, 160, 9–14. [CrossRef]
26. Scarpetta, H.; Gomide, F. Fuzzy reasoning and fuzzy Petri nets in manufacturing systems modeling. Journ. Intell. Fuzzy Syst. 1993, 1, 225–241. [CrossRef]
27. Liu, N.M.; Horng, T.; Chiang, K.T. The method of grey-fuzzy logic for optimizing multi-response problems during the manufacturing process: a case study of the light guide plate printing process. Int. Journ. Adv. Manuf. Technol. 2009, 41, 200–210. [CrossRef]
28. Zhou, X.; Peng, T. Application of multi-sensor fuzzy information fusion algorithm in industrial safety monitoring system. Sof. Sci. 2020, 122, 104531. [CrossRef]
29. Seiti, H.; Hafezalkotob, A.; Martinez, L. R-sets, Comprehensive Fuzzy Sets Risk Modeling for Risk-based Information Fusion and Decision-making. *IEEE Trans. Fuzzy Syst.* 2019. [CrossRef]
30. Yuqin, J.; Peixia, W.; Yue, L. Study of manufacturing system based on neural network multi-sensor data fusion and its application. In Proceedings of the IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, Changsha, China, 8–13 October 2003; Volume 2, pp. 1022–1026.
31. Goebel, K.; Badami, V.; Perera, A. Diagnostic information fusion for manufacturing processes. In Proceedings of the Second International Conference on Information Fusion, Sunnyvale, CA, USA; 1999; Volume 99, pp. 331–336.
32. Chen, S.L.; Jen, Y.W. Data fusion neural network for tool condition monitoring in CNC milling machining. *Int. J. Mach. Tools Manuf.* 2000, 40, 381–400. [CrossRef]
33. Li, Y.; Xiao, Z.; Zhen, X.; Cao, X. Attentional information fusion networks for cross-scene power line detection. *IEEE Geosci. Remote Sens. Lett.* 2019, 16, 1635–1639. [CrossRef]
34. Liu, Y.; Wang, Z.; Ma, L.; Alsaaedi, F.E. A partial-nodes-based information fusion approach to state estimation for discrete-time delayed stochastic complex networks. *Inf. Fusion* 2019, 49, 240–248. [CrossRef]
35. Jian, X.; Li, W.; Guo, X.; Wang, R. Fault diagnosis of motor bearings based on a one-dimensional fusion neural network. *Sensors* 2019, 19, 122. [CrossRef]
36. Yuan, K.; Xiao, F.; Fei, L.; Kang, B.; Deng, Y. Modeling Sensor Reliability in Fault Diagnosis Based on Evidence Theory. *Sensors* 2016, 16, 113. [CrossRef]
37. Gong, Y.; Su, X.; Qian, H.; Yang, N. Research on fault diagnosis methods for the reactor coolant system of nuclear power plant based on DS evidence theory. *Ann. Nucl. Energy* 2018, 112, 395–399. [CrossRef]
38. Qingdong, D.; Lingyu, X.; Hai, Z. DS evidence theory applied to fault diagnosis of generator based on embedded sensors. In Proceedings of the Third International Conference on Information Fusion, Paris, France, 10–13 July 2000; Volume 1.
39. Chorabae, M.K.; Zavadskas, E.K.; Amiri, M.; Antucheviciene, J. A New Method of Assessment Based On Fuzzy Ranking and Aggregated Weights (Afraw) For Mcdm Problems Under Type-2 Fuzzy Environment. *Econ. Comput. Econ. Cybern. Stud. Res.* 2017, 50, 39–68.
40. Wang, H.; Lin, D.; Qiu, J.; Ao, L.; Du, Z.; He, B. Research on multiobjective group decision-making in condition-based maintenance for transmission and transformation equipment based on DS evidence theory. *IEEE Trans. Smart Grid* 2015, 6, 1035–1045. [CrossRef]
41. Zhang, Z. Data Fusion Optimization Analysis of Wireless Sensor Networks Based on oint DS Evidence Theory and Matrix Analysis. In Proceedings of the 2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCE), Hohhot, China, 24–26 October 2019; pp. 689–6894.
42. Shi, J.; Luo, L.; Fan, C. Research on Fusion Method of Fault Diagnosis Based on DBN and Correlation Model for Optimized DS Evidence Theory. In Proceedings of the 2019 Prognostics and System Health Management Conference (PHM-Paris), Paris, France, 2–5 May 2019; pp. 356–361.
43. Jiang, S.; Lian, M.; Lu, C.; Ruan, S.; Wang, Z.; Chen, B. SVM-DS fusion based soft fault detection and diagnosis in solar water heaters. *Energy Explor. Exploit.* 2019, 37, 1125–1146. [CrossRef]
44. Tang, X.; Gu, X.; Wang, J.; He, Q.; Zhang, F.; Lu, J. A Bearing Fault Diagnosis Method Based On Feature Selection Feedback Network and Improved DS Evidence Fusion. *IEEE Access.* 2020. [CrossRef]
45. Su, X.; Mahadevan, S.; Xu, P.; Deng, Y. Handling of dependence in Dempster–Shafer theory. *Int. Journal Intell. Syst.* 2015, 30, 441–467. [CrossRef]
46. Zadeh, L.A. Review of A Mathematical Theory of Evidence. *Ai Mag.* 1984, 5, 235–247.
47. Wehrl, A. General properties of entropy. *Rev. Mod. Phys.* 1978, 50, 221. [CrossRef]
48. Gray, R.M. *Entropy and Information Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011.
49. Sun, Z.H.; Liu, R.; Ming, X. A Fault Diagnosis and Maintenance Decision System for Production Line Based on Human-machine Multi-Information Fusion. In Proceedings of the 2018 Artificial Intelligence and Cloud Computing Conference, Tokyo, Japan, 21–23 December 2018; pp. 151–156. [CrossRef]