Computers & Operations Research 36 (2009) 1768 – 1779

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/cor

Failure mode and effects analysis using a group-based evidential reasoning approach

Kwai-Sang Chin[a,*], Ying-Ming Wang[b], Gary Ka Kwai Poon[a], Jian-Bo Yang[c]

[a]Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong
[b]School of Public Administration, Fuzhou University, Fuzhou 350002, PR China
[c]Manchester Business School, The University of Manchester, Manchester M15 6PB, UK

ARTICLE INFO

Available online 14 May 2008

Keywords:
Failure mode and effects analysis
Evidential reasoning approach
Uncertainty modeling
Minimax regret ranking

ABSTRACT

Failure mode and effects analysis (FMEA) is a methodology to evaluate a system, design, process or service for possible ways in which failures (problems, errors, risks and concerns) can occur. It is a group decision function and cannot be done on an individual basis. The FMEA team often demonstrates different opinions and knowledge from one team member to another and produces different types of assessment information such as complete and incomplete, precise and imprecise and known and unknown because of its cross-functional and multidisciplinary nature. These different types of information are very difficult to incorporate into the FMEA by the traditional risk priority number (RPN) model and fuzzy rule-based approximate reasoning methodologies. In this paper we present an FMEA using the evidential reasoning (ER) approach, a newly developed methodology for multiple attribute decision analysis. The proposed FMEA is then illustrated with an application to a fishing vessel. As is illustrated by the numerical example, the proposed FMEA can well capture FMEA team members' diversity opinions and prioritize failure modes under different types of uncertainties.

1. Introduction

Failure mode and effects analysis (FMEA) is an engineering technique used to define, identify and eliminate known and/or potential failures, problems, errors and so on from the system, design, process and/or service before they reach the customer [1–3]. When it is used for a criticality analysis, it is also referred to as failure mode, effects and criticality analysis (FMECA). FMEA has gained wide acceptance and applications in a wide range of industries such as aerospace, nuclear, chemical and manufacturing. A good FMEA can help analysts identify known and potential failure modes and their causes and effects, help them prioritize the identified failure modes and can also help them work out corrective actions for the failure modes. The main objective of FMEA is to allow the analysts to identify and prevent known and potential problems from reaching the customer. To this end, the risks of each identified failure mode need to be evaluated and prioritized so that appropriate corrective actions can be taken for different failure modes. The priority of a failure mode is determined through the risk priority number (RPN), which is defined as the product of the occurrence (O), severity (S) and detection (D) of the failure, namely

$$\text{RPN} = O \times S \times D. \tag{1}$$

The three factors O, S and D are all evaluated using the ratings (also called rankings or scores) from 1 to 10, as described in Tables 1–3. The failures with higher RPNs are assumed to be more important and should be given higher priorities.

FMEA has been proven to be one of the most important early preventative initiatives during the design stage of a system, product, process or service. However, the RPN has been extensively criticized for various reasons [4,5,7–11]:

- Different sets of O, S and D ratings may produce exactly the same value of RPN, but their hidden risk implications may be totally different. For example, two different events with values of 2, 3, 2 and 4, 1, 3 for O, S and D, respectively, will have the same RPN value of 12. However, the hidden risk implications of the two events may be very different because of the different severities of the failure consequence. This may cause a waste of resources and time, or in some cases, a high-risk event being unnoticed.
- The relative importance among O, S and D is not taken into consideration. The three factors are assumed to have the same importance. This may not be the case when considering a practical application of FMEA.

* Corresponding author. Tel.: +852 2788 8306; fax: +852 2788 8423.
E-mail address: mekschin@cityu.edu.hk (K.-S. Chin).

0305-0548/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2008.05.002

- The mathematical formula for calculating RPN is questionable and debatable. There is no rationale as to why *O*, *S* and *D* should be multiplied to produce the RPN.
- The conversion of scores is different for the three factors. For example, a linear conversion is used for *O*, but a nonlinear transformation is employed for *D*.
- RPNs are not continuous with many holes and heavily distributed at the bottom of the scale from 1 to 1000. This causes problems in interpreting the meaning of the differences between different RPNs. For example, is the difference between the neighboring RPNs of 1 and 2 the same or less than the difference between 900 and 1000?
- The RPN considers only three factors mainly in terms of safety. Other important factors such as economical aspects are ignored.
- Small variations in one rating may lead to vastly different effects on the RPN, depending on the values of the other factors. For example, if *O* and *D* are both 10, then a 1-point difference in severity rating results in a 100-point difference in the RPN; if *O* and *D* are equal to 1, then the same 1-point difference results in only a 1-point difference in the RPN; if *O* and *D* are both 4, then a 1-point difference produces a 16-point difference in the RPN.
- The three factors are difficult to precisely determine. Much information in FMEA can be expressed in a linguistic way such as *likely*, *important* or *very high* and so on.

A number of approaches have been suggested in the literature to overcome some of the drawbacks mentioned above. For example, Gilchrist [10] gave a critique of FMEA and proposed an expected cost model. It was formulated as $EC = CnP_fP_d$, where *EC* is the expected cost to the customer, *C* the cost per failure, *n* the items produced per batch or per year, P_f the probability of a failure and P_d the probability of the failure not to be detected. P_f and P_d were assumed to be independent and their product represents the probability that the customer receives a faulty product. The nP_fP_d is the expected number of failures reaching the customer. The expected cost model was claimed to be more rigorous yet practical than the RPN model and to have great benefit of forcing people to think about quality costs.

Ben-Daya and Raouf [7] argued that the probabilities P_f and P_d in the expected cost model were not always independent and very difficult to estimate at the design stage of a product and the severity was completely ignored by the expected cost model. Based on these arguments, they proposed an improved FMEA model which addressed Gilchrist's criticism and gave more importance to the likelihood of occurrence over the likelihood of detection by raising the ratings for the likelihood of occurrence to the power of 2. The improved FMEA model was combined with the expected cost model to

Table 1
Traditional ratings for occurrence of a failure [4–6].

Rating	Probability of occurrence	Possible failure rate
10	Very high: failure is almost inevitable	$\geqslant 1/2$
9		1/3
8	High: repeated failures	1/8
7		1/20
6	Moderate: occasional failures	1/80
5		1/400
4		1/2000
3	Low: relatively few failures	1/15,000
2		1/150,000
1	Remote: failure is unlikely	$\leqslant 1/1,500,000$

Table 2
Traditional ratings for severity of a failure [4–6].

Rating	Effect	Severity of effect
10	Hazardous without warning	Very high severity ranking when a potential failure mode affects safe vehicle operation and/or involves noncompliance with government regulations without warning
9	Hazardous with warning	Very high severity ranking when a potential failure mode affects safe vehicle operation and/or involves noncompliance with government regulations with warning
8	Very high	Vehicle/item inoperable, with loss of primary function
7	High	Vehicle/item operable, but at reduced level of performance. Customer dissatisfied
6	Moderate	Vehicle/item operable, but comfort/convenience item(s) inoperable. Customer experiences discomfort
5	Low	Vehicle/item operable, but comfort/convenience item(s) operable at reduced level of performance. Customer experiences some dissatisfaction
4	Very low	Cosmetic defect in finish, fit and finish/squeak or rattle item that does not conform to specifications. Defect noticed by most customers
3	Minor	Cosmetic defect in finish, fit and finish/squeak or rattle item that does not conform to specifications. Defect noticed by average customer
2	Very minor	Cosmetic defect in finish, fit and finish/squeak or rattle item that does not conform to specifications. Defect noticed by discriminating customers
1	None	No effect

Table 3
Traditional ratings for detection [4–6].

Rating	Detection	Criteria
10	Absolutely impossible	Design control will not and/or cannot detect a potential cause/mechanism and subsequent failure mode; or there is no design control
9	Very remote	Very remote chance the design control will detect a potential cause/mechanism and subsequent failure mode
8	Remote	Remote chance the design control will detect a potential cause/mechanism and subsequent failure mode
7	Very low	Very low chance the design control will detect a potential cause/mechanism and subsequent failure mode
6	Low	Low chance the design control will detect a potential cause/mechanism and subsequent failure mode
5	Moderate	Moderate chance the design control will detect a potential cause/mechanism and subsequent failure mode
4	Moderately high	Moderately high chance the design control will detect a potential cause/mechanism and subsequent failure mode
3	High	High chance the design control will detect a potential cause/mechanism and subsequent failure mode
2	Very high	Very high chance the design control will detect a potential cause/mechanism and subsequent failure mode
1	Almost certain	Design control will almost certainly detect a potential cause/mechanism and subsequent failure mode

provide a quality improvement scheme for the production phases of a product or service in the way that the former was used to identify the critical failures that require immediate remedial action, whereas the later was used in parallel to estimate the cost of failures reaching the customer and to evaluate the impact of the corrective action taken.

Sankar and Prabhu [5] presented a modified approach for prioritization of failures in a system FMEA, which uses the ranks 1–1000 called risk priority ranks (RPRs) to represent the increasing risk of the 1000 possible severity–occurrence–detection combinations. These 1000 possible combinations were tabulated by an expert in order of increasing risk and can be interpreted as 'if–then' rules. The failure having a higher rank was given a higher priority.

Bevilacqua et al. [12] defined RPN as the weighted sum of six parameters which are safety, machine importance for the process, maintenance costs, failure frequency, downtime length and operating conditions, multiplied by the seventh factor, i.e. machine access difficulty, where the relative importance weights of the six parameters were estimated using pairwise comparisons. Monte Carlo simulation was performed as a sensitivity analysis to verify the robustness of the final ranking results.

Braglia [13] developed a multi-attribute failure mode analysis (MAFMA) based on the analytic hierarchy process (AHP) technique, which considers four different factors O, S, D, and expected cost as decision attributes, possible causes of failure as decision alternatives, and the selection of cause of failure as decision goal. The goal, attributes and alternatives formed a three-level hierarchy, where the pairwise comparison matrix was used to estimate attribute weights and the local priorities of the causes with respect to the expected cost attribute, the conventional scores for O, S and D were normalized as the local priorities of the causes with respect to O, S, and D, respectively, and the weight composition technique in the AHP was utilized to synthesize the local priorities into the global priority, based on which the possible causes of failure were ranked. A sensitivity analysis was also conducted to investigate the sensitivity of the priority ranking of the causes to the changes in attribute weights.

Braglia et al. [14] also presented an alternative multi-attribute decision-making approach called fuzzy TOPSIS approach for FMECA, which is a fuzzy version of the technique for order preference by similarity to ideal solution (TOPSIS). The TOPSIS method is a well-known multi-attribute decision-making methodology based on the assumption that the best decision alternative should be as close as possible to the ideal solution and the farthest from the negative-ideal solution. The proposed fuzzy TOPSIS approach allows the risk factors O, S and D and their relative importance to be assessed using triangular fuzzy numbers rather than precise crisp numbers.

Chang et al. [15] used fuzzy sets and gray systems theory for FMEA, where fuzzy linguistic terms such as *very low*, *low*, *moderate*, *high* and *very high* were used to evaluate the degrees of O, S and D, and gray relational analysis was applied to determine the risk priority of potential causes. To carry out the gray relational analysis, fuzzy linguistic assessment information was defuzzified as crisp values, the lowest level of the three factors O, S and D was defined as a standard series, and the assessment information of the three factors for each potential cause was viewed as a comparative series, whose gray relational coefficients and gray relational degree with the standard series were computed in terms of the gray systems theory [16]. Bigger gray relational degree means smaller effect of potential cause. The increasing order of the gray relational degrees represents the risk priority of the potential problems to be improved. In [9], Chang et al. also utilized the gray system theory for FMEA, but the gray relational degrees were computed using the traditional scores 1–10 for the three factors rather than fuzzy linguistic assessment information.

Seyed-Hosseini et al. [17] proposed a method called decision making trial and evaluation laboratory (DEMATEL) for reprioritization of failure modes in FMEA, which prioritizes alternatives based on severity of effect or influence and direct and indirect relationships between them. Direct relationships were a set of connections between alternatives with a set of connection weights representing severity of influence of one alternative on another. An indirect relationship was defined as a relationship that could only move in an indirect path between two alternatives and meant that a failure mode could be the cause of other failure mode(s). Alternatives having more effect on another were assumed to have higher priority and called dispatcher. Those receiving more influence from another were assumed to have lower priority and called receiver.

Bowles and Peláez [18] described a fuzzy logic-based approach for prioritizing failures in a system FMECA, which uses linguistic terms such as *remote*, *low*, *moderate*, *high* and *very high* to describe O, *minor*, *low*, *moderate*, *high* and *very high* for S, *non-detection*, *very low*, *low*, *moderate*, *high* and *very high* for D and *not-important*, *minor*, *low*, *moderate*, *important* and *very important* for the riskiness of failure. The relationships between the riskiness and O, S, D were characterized by a fuzzy if–then rule base which was developed from expert knowledge and expertise. Crisp ratings for O, S and D were fuzzified to match the premise of each possible if–then rule. All the rules that have any truth in their premises were fired to contribute to the fuzzy conclusion. The fuzzy conclusion was then defuzzified by the weighted mean of maximum method (WMoM) as the ranking value of the risk priority. Similar fuzzy inference method also appeared in [6,8,11,19–25].

Fuzzy RPN approaches usually require a large number of rules and it is a tedious task to obtain a full set of rules. The larger the number of rules provided by the users, the better the prediction accuracy of the fuzzy RPN model. Tay and Lim [24] argued that not all the rules were actually required in fuzzy RPN models, eliminating some of the rules did not necessarily lead to a significant change in the model output, but some of the rules might be vitally important and could not be ignored. They thus proposed a guided rules reduction system (GRRS) to simplify the fuzzy logic-based FMEA methodology by reducing the number of rules that need to be provided by FMEA users for fuzzy RPN modeling process.

The above literature review shows that much effort has been paid to the improvement of FMEA by incorporating factor weights, more factors, expert knowledge and/or fuzziness into the analysis, but no or little attention has been paid to the diversity and uncertainty of assessment information. As is known, FMEA is a team function and cannot be performed on an individual basis. In other words, FMEA is a group decision behavior. Different FMEA team members may demonstrate different opinions because of their different expertise and backgrounds [26]. They may provide different assessment information for the same risk factor, some of which may be complete or incomplete, precise or imprecise, known or unknown and certain or uncertain. This diversity and uncertainty of assessment information is sometimes inherent, not easy to eliminate and in need of being considered in FMEA. In this paper we propose a new risk priority model for FMEA using the evidential reasoning (ER) approach. The new model can not only model the diversity and uncertainty of the assessment information in FMEA, but also incorporate the relative importance of risk factors into the determination of risk priority of failure modes in a strict way.

The paper is organized as follows. In Section 2, we develop the risk priority model using the ER approach and incorporate the relative importance weights of risk factors into the determination of risk priority of failure modes. In Section 3, we provide a numerical example to illustrate the potential applications of the new model in FMEA. Section 4 concludes the paper with a summary.

2. Risk priority model using the ER approach

The ER approach was developed for multiple attribute decision analysis (MADA) and has found an increasing number of applications in recent years [27–34]. In this section, we develop a risk priority model for FMEA using the ER approach to model the diversity and uncertainty of the assessment information in FMEA. The new model allows FMEA team members to assess risk factors independently and express their opinions individually. It also allows the risk factors to be aggregated in a rigorous yet nonlinear rather than simple addition or multiplication manner. The model is developed step-by-step as follows.

2.1. Assessment of risk factors using belief structures

O, *S* and *D* are the three major risk factors identified in FMEA. Although more risk factors could be included, the main concern of this paper is not the identification of risk factors, but their assessment and aggregation. The risk priority model to be developed in this paper has no limitation on the number of risk factors and is applicable to any number of risk factors. The three risk factors can be evaluated numerically or linguistically. Both of them have been extensively applied and have their merits and demerits. For example, linguistic terms such as *very low*, *low*, *moderate*, *high*, and *very high* allow FMEA team members to express their opinions in a fuzzy and imprecise way, but the determination of membership functions of the linguistic terms is highly subjective. In this paper, we choose the traditional numerical ratings in Tables 1–3 [4] for the assessment of risk factors, where *O* is rated according to the failure probability which represents the relative number of failures anticipated during the design life of an item, *S* is rated according to the seriousness of the failure mode effect on the next higher assembly, the system or the user and *D* is rated according to the designer's subjective judgment of the likelihood that a failure will be found in subsequent tests. Although linguistic terms can also be used for the model to be developed, the ER algorithm may be different, depending on whether their membership functions intersect or not. If there is no intersection between them as in [28], then the ER algorithm will be the same; otherwise, the fuzzy ER algorithm should be used in the next subsection. The interested reader may refer to [32] for the fuzzy ER algorithm.

As mentioned in the previous section, the numerical ratings in Tables 1–3 were criticized because they were not easy to precisely determine. Such a drawback can be overcome by the ER approach, which allows FMEA team members to provide their subjective judgments in the following flexible ways:

- A precise rating such as 4, which can be written as $\{(4, 100\%)\}$. Such an expression is referred to as a belief structure in the ER approach.
- A distribution such as 4 to 40% and 5 to 60%, which means that a failure mode is assessed with respect to the risk factor under consideration to rating 4 to the degree of 40% and to rating 5 to the degree of 60%. Here the degrees of 40% and 60% represent the confidences (also called belief degrees) of the FMEA team member in his/her subjective judgments and the distribution can be equivalently expressed as $\{(4, 40\%), (5, 60\%)\}$. When all the confidences are summed to one, the distribution is said to be complete; otherwise, it is said to be incomplete. For example, $\{(4, 40\%), (5, 50\%)\}$ is an incomplete distribution or called incomplete assessment, where the missing information of 10% is referred to as local ignorance and could be assigned to any rating between 1 and 10 according to the Dempster–Shafer theory of evidence [35].
- An interval such as 4–5, which means that the rating of a failure mode with respect to the risk factor under evaluation is between 4 and 5. This can be written as $\{(4-5, 100\%)\}$.
- No judgment, which means the FMEA team member is not willing to or cannot provide an assessment for a failure mode with respect to the risk factor under consideration. In other words, the rating by this FMEA team member could be anywhere between 1 and 10 and can be expressed as $\{(1\text{–}10, 100\%)\}$. Such judgments are referred to as total ignorance.

Obviously, the belief structures in the ER approach provide FMEA team members with an easy-to-use and very flexible way to express their opinions and can better quantify risk factors than the traditional RPN methodology. All failure modes with respect to the three risk factors can be evaluated using belief structures. In the next subsection, we will see how the belief structures of each failure mode with respect to every risk factor provided by FMEA team members individually can be synthesized into a group belief structure and how the group belief structures of each failure mode with respect to the three risk factors can be aggregated into an overall belief structure using a recursive interval ER algorithm.

2.2. Group belief structures and their aggregations

Suppose there are K members $(TM_1, \ldots, TM_K)$ in a FEMA team responsible for the assessment of N failure modes $(FM_1, \ldots, FM_N)$ with respect to L risk factors $(RF_1, \ldots, RF_L)$. Each team member TM_k is given a weight $\lambda_k > 0$ $(k = 1, \ldots, K)$ satisfying $\sum_{k=1}^{K} \lambda_k = 1$ to reflect his/her relative importance in the FMEA team. Each risk factor RF_l is given a weight $w_l > 0$ $(l = 1, \ldots, L)$ satisfying $\sum_{l=1}^{L} w_l = 1$ to reflect its relative importance in the determination of risk priorities of the N failure modes. The two different sets of weights can be determined by using direct rating [36,37], point allocation [36,37], eigenvector method [38], linear programming techniques for multidimensional analysis of preferences (LINMAP) [39], or Delphi method [40], etc. together with the team members' domain knowledge. If there is no sufficient reason or evidence to show the differences among the FMEA team members in their judgment qualities, the team members should be given an equal weight.

Let $\{(H_{ij}, \beta_{ij}^{(k)}(FM_n, RF_l)),\ i = 1, \ldots, 10; j = i, \ldots, 10\}$ be the belief structure provided by TM_k on the assessment of FM_n with respect to RF_l, where H_{ii} for $i = 1\text{–}10$ are the ratings defined for risk assessment, H_{ij} for $i = 1\text{–}9$ and $j = i + 1$ to 10 are the intervals between H_{ii} and H_{jj}, and $\beta_{ij}^{(k)}(FM_n, RF_l)$ are the belief degrees to which FM_n is assessed on RF_l to the intervals H_{ij}. All the ratings H_{ii} for $i = 1\text{–}10$ and the intervals H_{ij} for $i = 1\text{–}9$ and $j = i + 1$ to 10 together form a frame of discernment, which is expressed as

$$H = \left\{ \begin{matrix} H_{11} & H_{12} & \cdots & H_{19} & H_{110} \\ & H_{22} & \cdots & H_{29} & H_{210} \\ & & \vdots & \vdots & \vdots \\ & & & H_{99} & H_{910} \\ & & & & H_{1010} \end{matrix} \right\} = \left\{ \begin{matrix} 1 & 1-2 & \cdots & 1-9 & 1-10 \\ & 2 & \cdots & 2-9 & 2-10 \\ & & \vdots & \vdots & \vdots \\ & & & 9 & 9-10 \\ & & & & 10 \end{matrix} \right\}. \tag{2}$$

The collective assessment of the K team members for each failure mode with respect to each risk factor is also a belief structure, called group or collective belief structure, which is denoted as $\{(H_{ij}, \beta_{ij}), i = 1, \ldots, 10; j = i, \ldots, 10\}$, where β_{ij} is referred to as group or collective belief degree and is determined by

$$\beta_{ij}(FM_n, RF_l) = \sum_{k=1}^{K} \lambda_k \beta_{ij}^{(k)}(FM_n, RF_l),$$
$$i = 1, \ldots, 10; \quad j = i, \ldots, 10;$$
$$n = 1, \ldots N; \quad l = 1, \ldots, L. \tag{3}$$

That is, a group belief degree is the weighted sum of the individual belief degrees corresponding to the same rating or interval. Take the belief structures in Table 6 for severity assessment of failure mode 1 for example. The five individual belief structures are, respectively, as $\{(7, 20\%), (8, 80\%)\}$, $\{(8, 100\%)\}$, $\{(8, 100\%)\}$, $\{(6-7, 50\%), (8-9, 50\%)\}$ and $\{(8, 100\%)\}$, where the ratings and intervals with a zero belief degree are omitted from the belief structures for brevity. The relative importance weights of the five team members are known as 0.3, 0.3, 0.2, 0.1 and 0.1, respectively. By Eq. (3), the group belief structure from the five individual belief structures can be determined as $\{(7, 6\%), (8, 84\%), (6-7, 5\%), (8-9, 5\%)\}$.

The group belief structures for the N failure modes with respect to the L risk factors form a belief decision matrix, which differs from the traditional decision matrix in that the former consists of belief structures while the latter is made up of numerical values. Based on the belief decision matrix, group belief structures on the assessment of each failure mode with respect to the L risk factors can be aggregated into an overall belief structure using a recursive interval ER algorithm.

Different from the traditional RPN which is the simple product of the three risk factors O, S and D, the recursive interval ER algorithm aggregates risk factors in a systematic yet rigorous way, which is neither their simple addition nor their simple multiplication, but a highly nonlinear form of the risk factors. The interested reader may refer to Yang and Xu [34] for the discussion on the nonlinearity of the ER approach. The aggregation is based on the combination rule of the Dempster–Shafer theory of evidence [35] and is detailed below.

Let $\{(H_{ij}, \beta_{ij}(FM_n, RF_l)), i = 1, \ldots, 10; j = i, \ldots, 10\}$ and $\{(H_{ij}, \beta_{ij}(FM_n, RF_p)), i = 1, \ldots, 10; j = i, \ldots, 10\}$ be two group belief structures on the assessment of the failure mode FM_n with respect to the risk factors RF_l and RF_p $(1 \leqslant l, p \leqslant L)$, respectively, and w_l and w_p be the relative importance weights of the two risk factors. The recursive interval ER algorithm first transforms the two group belief structures into basic probability masses by considering the relative importance weights of the two risk factors and using the following equations:

$$m_{ij} = w_l \beta_{ij}(FM_n, RF_l), \quad i = 1, \ldots, 10; \quad j = i, \ldots, 10, \tag{4}$$

$$m_H = 1 - w_l, \tag{5}$$

$$n_{ij} = w_p \beta_{ij}(FM_n, RF_p), \quad i = 1, \ldots, 10; \quad j = i, \ldots, 10, \tag{6}$$

$$n_H = 1 - w_p. \tag{7}$$

The above probability masses are viewed as two pieces of evidence and combined to produce a set of combined probability masses: c_{ij} $(i = 1, \ldots, 10; j = i, \ldots, 10)$ and c_H, which are computed using the following equations [31]:

$$c_{ij} = \frac{1}{1-C}\left[-m_{ij}n_{ij} + \sum_{k=1}^{i}\sum_{l=j}^{10}(m_{kl}n_{ij} + m_{ij}n_{kl}) + \sum_{k=1}^{i-1}\sum_{l=j+1}^{10}(m_{kj}n_{il} + m_{il}n_{kj}) + m_H n_{ij} + m_{ij} n_H\right], \tag{8}$$

$$c_H = \frac{m_H n_H}{1-C}, \tag{9}$$

$$C = \sum_{i=1}^{10}\sum_{j=i}^{10}\sum_{k=1}^{i-1}\sum_{l=k}^{i-1}(m_{kl}n_{ij} + m_{ij}n_{kl}), \tag{10}$$

where the summation process $\sum_{i=i_1}^{i_2} f(i)$ will not be carried out if $i_1 > i_2$. That is, $\sum_{i=i_1}^{i_2} f(i) = 0$ for $i_1 > i_2$. The combined probability masses are then aggregated further with the basic probability masses transformed from the group belief structure on the assessment of the failure mode FM_n with respect to another risk factor. Such an aggregation process is recursively carried out until the L group belief structures on the assessment of the failure mode FM_n are all aggregated. The recursive interval ER algorithm is easy to implement on a microsoft excel worksheet. In Appendix A, we provide Table 11 to show how three pieces of evidence can be recursively combined on a microsoft excel worksheet. If there are more pieces of evidence, they can be recursively combined in the same way.

Let x_{ij} $(i = 1, \ldots, 10; j = i, \ldots, 10)$ and x_H be the final combined probability masses. The overall assessment of the failure mode FM_n is an overall belief structure, denoted by $\{(H_{ij}, \delta_{ij}(FM_n)), i = 1, \ldots, 10; j = i, \ldots, 10\}$, where $\delta_{ij}(FM_n)$ represents the overall belief degree that the failure mode FM_n is assessed to the interval H_{ij} and is determined by the following equation:

$$\delta_{ij}(FM_n) = \frac{x_{ij}}{1 - x_H}, \quad i = 1, \ldots, 10; \quad j = i, \ldots, 10. \tag{11}$$

In the case that the relative importance weights of risk factors are deterministic, $\delta_{ij}(FM_n)$ for $i = 1, \ldots, 10; j = i, \ldots, 10$ are also deterministic. If the weights themselves are uncertain, say intervals, then $\delta_{ij}(FM_n)$ are also intervals, denoted by $[\delta_{ij}^L(FM_n), \delta_{ij}^U(FM_n)]$, where $\delta_{ij}^L(FM_n)$ and $\delta_{ij}^U(FM_n)$ are determined by the following pair of models:

$$\delta_{ij}^L(FM_n) = \text{Minimize} \quad \frac{x_{ij}}{1 - x_H}$$
$$\text{Subject to} \quad w_l^L \leqslant w_l \leqslant w_l^U, \quad l = 1, \ldots, L,$$
$$\sum_{l=1}^{L} w_l = 1, \tag{12}$$

$$\delta_{ij}^U(FM_n) = \text{Maximize} \quad \frac{x_{ij}}{1 - x_H}$$
$$\text{Subject to} \quad w_l^L \leqslant w_l \leqslant w_l^U, \quad l = 1, \ldots, L,$$
$$\sum_{l=1}^{L} w_l = 1. \tag{13}$$

The overall assessments of the N failure modes can all be obtained in this way. After performing the recursive interval ER algorithm for the N failure modes, we get N overall belief structures, each for one failure mode.

2.3. Expected risk score

The overall belief structure provides for each failure mode a panoramic view which shows the ratings and intervals each failure mode is assessed to and the belief degrees assessed to these ratings and intervals. Such information is helpful for the FMEA team to understand the overall risk of each failure mode. Failure modes assessed to high ratings or intervals with high belief degrees are obviously more risky than those assessed to low ratings or intervals with high belief degrees. In a very small number of cases such as a small number of failure modes, the risk priority of failure modes could be determined by observing and analyzing their overall belief structures. However, in most cases, the risk priority of failure modes cannot be easily determined by analyzing their overall belief structures. For ranking purpose, the overall belief structures need to be converted into expected risk scores, which are defined as below:

$$ERS(FM_n) = \sum_{i=1}^{10}\sum_{j=i}^{10} \delta_{ij}(FM_n)H_{ij}, \quad n = 1, \ldots, N. \tag{14}$$

In the case that there exists $\delta_{ij}(FM_n) \neq 0$ for some $j \neq i$ or the relative importance weights of risk factors are intervals, the expected risk score $ERS(FM_n)$ is an interval, denoted by $[ERS^L(FM_n), ERS^U(FM_n)]$, where $ERS^L(FM_n)$ and $ERS^U(FM_n)$ are determined by

$$\begin{aligned} ERS^L(FM_n) = \text{Minimize} \quad & \sum_{i=1}^{10}\sum_{j=i}^{10} \delta_{ij}(FM_n)H_{ii} \\ \text{Subject to} \quad & w_l^L \leqslant w_l \leqslant w_l^U, \quad l = 1, \ldots, L, \\ & \sum_{l=1}^{L} w_l = 1, \end{aligned} \tag{15}$$

$$\begin{aligned} ERS^U(FM_n) = \text{Maximize} \quad & \sum_{i=1}^{10}\sum_{j=i}^{10} \delta_{ij}(FM_n)H_{jj} \\ \text{Subject to} \quad & w_l^L \leqslant w_l \leqslant w_l^U, \quad l = 1, \ldots, L, \\ & \sum_{l=1}^{L} w_l = 1, \end{aligned} \tag{16}$$

where $\delta_{ij}(FM_n)$ are determined by Eq. (11). Accordingly, the average expected risk score is defined as

$$\overline{ERS}(FM_n) = \tfrac{1}{2}(ERS^L(FM_n) + ERS^U(FM_n)), \quad n = 1, \ldots, N. \tag{17}$$

By solving the above pair of models for each failure mode, the expected risk scores of all the N failure modes can be generated. The bigger the expected risk score, the higher the risk priority. The N failure modes can be prioritized based on their expected risk scores using a minimax regret ranking approach.

2.4. The minimax regret approach for ranking expected risk scores

The minimax regret approach (MRA) developed by Wang et al. [41] is a method for comparing and ranking interval numbers and is briefly summarized below for the sake of application. Let $u_i = [u_i^L, u_i^U] = \langle c_i, d_i \rangle$ $(i = 1, \ldots, N)$ be N intervals, where $c_i = \frac{1}{2}(u_i^L + u_i^U)$ and $d_i = \frac{1}{2}(u_i^U - u_i^L)$ are their midpoints and widths. Without loss of generality, suppose $u_i = [u_i^L, u_i^U]$ is chosen as the biggest interval. Let $v = \max_{j \neq i}(u_j^U)$. Obviously, if $u_i^L < v$, the decision maker (DM) may regret due to the loss of opportunity that other interval numbers might be ranked higher than u_i. The maximum loss the DM may suffer from is given by

$$\text{Max}(r_i) = v - u_i^L = \max_{j \neq i}(u_j^U) - u_i^L.$$

If $u_i^L \geqslant v$, the DM will definitely suffer from no loss of opportunity and thus will not regret. In this situation, the DM's regret is defined as zero, i.e. $r_i = 0$. Combining the above two situations, we have

$$\text{Max}(r_i) = \max\left[\max_{j \neq i}(u_j^U) - u_i^L, 0\right].$$

The minimax regret criterion will choose the interval satisfying the following condition as the best (most desirable) one:

$$\min_i\{\max(r_i)\} = \min_i\left\{\max\left[\max_{j \neq i}(u_j^U) - u_i^L, 0\right]\right\}.$$

Based on the above analysis, Wang et al. [41] gave the following definition for comparing and ranking interval numbers.

Definition 1. Let $u_i = [u_i^L, u_i^U] = \langle c_i, d_i \rangle$ $(i = 1, \ldots, N)$ be N intervals. The maximum regret value (MRV) of each interval u_i is defined as

$$R(u_i) = \max\left[\max_{j \neq i}(u_j^U) - u_i^L, 0\right] = \max\left[\max_{j \neq i}(c_j + d_j) - (c_i - d_i), 0\right], \quad i = 1, \ldots, N. \tag{18}$$

The interval with the smallest MRV should be chosen as the best interval. In order to generate a full ranking for the N intervals, the following eliminating process was suggested by Wang et al.

Step 1. Calculate the MRVs of the N intervals and choose the interval with the smallest MRV as the best one. Suppose u_{i_1} is selected for $1 \leqslant i_1 \leqslant N$.

Step 2. Eliminate u_{i_1} from the further consideration and recalculate the MRVs of the remaining $(N-1)$ intervals, from which choose the one with the smallest MRV as the second best interval. Suppose u_{i_2} is chosen for $1 \leqslant i_2 \leqslant N$, but $i_2 \neq i_1$.

Step 3. Eliminate u_{i_2} from the further consideration and recalculate the MRVs of the remaining $(N-2)$ intervals, from which choose the one with the smallest MRV as the third best interval.

Step 4. Repeat the above elimination process until only one interval u_{i_N} is left. The final ranking is given by $u_{i_1} > u_{i_2} > \cdots > u_{i_N}$.

By means of the above MRA, the N expected risk scores can all be ranked. The ranking will serve as the risk priority of the N failure modes.

3. Application to a fishing vessel

In this section, we study an FMEA problem using the ER approach in a group-based decision-making environment to show its potential applications and benefits. This FMEA example is adapted from [11] and is limited to only a few systems of an ocean going fishing vessel. In other words, not all possible failure modes in a fishing vessel are considered in this example.

The FMEA for the fishing vessel in question investigates four different systems which are structure, propulsion, electrical, and auxiliary systems. Each system is considered for different failure modes that could lead to accidents with undesired consequences. The effects of each failure mode on the system and vessel are studied along with the provisions that are in place or available to mitigate or reduce risks. For each of the failure modes, the systems are investigated for any alarms or condition monitoring arrangements, which are in place. There are 21 failure modes in total which were identified by a FMEA team and are presented together with their effects on the systems and vessel in Table 4.

Suppose the FMEA team is made up of five experts, each playing a different role in the team and given a different weight. The weights for the five members are assumed to be 0.3, 0.3, 0.2, 0.1 and 0.1, respectively. Each team member evaluates the 21 failure modes with respect to three major risk factors O, S and D using the traditional ratings individually. Tables 5–7 present the assessment results of the five team members on the 21 failure modes with respect to the three major risk factors, where incomplete assessments and ignorance information are shaded and highlighted. The relative importance weights of the three risk factors are provided as intervals, i.e. $w_O \in [0.2, 0.35]$, $w_S \in [0.4, 0.5]$ and $w_D \in [0.15, 0.25]$.

Table 4
FMEA for a fishing vessel [11].

Item	Description	Component	Failure mode	Failure effect on system	Failure effect on vessel	Alarm	Provision
1	Structure	Rudder bearing	Seizure	Rudder jam	No steering ctrl	No	Stop vessel
2	Structure	Rudder bearing	Breakage	Rudder loose	Reduced steering ctrl	No	Stop vessel
3	Structure	Rudder structure	Structural failure	Function loss	Reduced steering	No	Use beams
4	Propulsion	Main engine	Loss of output	Loss of thrust	Loss of speed	Yes	None
5	Propulsion	Main engine	Auto shutdown	M/E stops	Loss of speed	Yes	Anchor
6	Propulsion	Shaft and propeller	Shaft breakage	Loss of thrust	Loss of speed	No	Anchor
7	Propulsion	Shaft and propeller	Shaft seizure	Loss of thrust	Loss of speed	Yes	Anchor
8	Propulsion	Shaft and propeller	Gearbox seizure	Loss of thrust	Loss of speed	Yes	Anchor
9	Propulsion	Shaft and propeller	Hydraulic failure	Cannot reduce thrust	Cannot reduce speed	No	Anchor
10	Propulsion	Shaft and propeller	Prop. blade failure	Loss of thrust	Loss of speed	No	Slow steaming
11	Air services	Air receiver	No start air pressure	Cannot start M/E	No propulsion	Yes	Recharge receiver
12	Electrical system	Power generation	Generator fail	No electric power	Some system failures	Yes	Use stand by generators
13	Electrical system	Main switch board	Complete loss	Loss of main supply	No battery charging	Yes	Use emergency 24 V
14	Electrical system	Emergency S/B	Complete loss	Loss of emergency supply	No emergency supply	No	Use normal supply
15	Electrical system	Main batteries	Loss of output	Loss of main 24 V	Loss of main low volt	Yes	Use emergency 24 V
16	Electrical system	Emergency batteries	Loss of output	Loss of emergency supply	No emergency supply	No	Use normal supply
17	Auxiliary system	Fuel system	Contamination	M/E and generation stop	Vessels stops	Yes	Anchor
18	Auxiliary system	Fuel system	No fuel to M/E	M/E stops	Vessel stops	No	Anchor
19	Auxiliary system	Water system	No cooling water	Engine overheat	M/E auto cut-out	Yes	Use stand-by pump
20	Auxiliary system	Hydraulic	System loss	No steering	No steering	Yes	Stop vessel
21	Auxiliary system	Lube oil system	Loss of pressure	Low pressure cut-off	M/E stops	Yes	Use stand-by pump

Table 5
Occurrence assessment by FMEA team members.

Failure mode	FMEA team member				
	1	2	3	4	5
1	1	1	[redacted]	1	1
2	1: 50%, 2: 50%	1	1	1	1
3	2	2: 90%	2	2	2
4	8	8	8: 80%, 9: 20%	8	8
5	6	6	6	6	6
6	2	2	2	2	2–3
7	2	2	2	2	2
8	1	1: 75% 2: 25%	1	1	1
9	[redacted]	3	3	3	3
10	1: 80% 2: 20%	1	1	1	1–2: 85% 3: 15%
11	4	4	4	3–4: 75% 5: 25%	4
12	9	9	9	9	9
13	8	8: 80%	8	8	8
14	3	3	4	3	3
15	3	3	3	3	3: 70% 4: 30%
16	1	1	1	1	1
17	3–5: 90% 6: 10%	4	4	4	4
18	2	2	2: 90%	2	2
19	7	7	7	7	7: 80%
20	9	9	9	7: 30% 8–9: 70%	9
21	9	8–9	9	9	9

Obviously, there is no existing FMEA method that can be used, without making some kind of assumptions, to deal with the assessment information in Tables 5–7, which is different from one team member to another and also includes incomplete assessments and ignorance. To carry out a priority analysis, we first use belief structures to express the FMEA team members' individual assessments and synthesize them into group belief structures by using Eq. (3), as presented in Table 8. The group belief structures are then aggregated into overall belief structures using the recursive interval ER algorithm described in Section 2. The results are presented in Table 9, where the overall belief degrees, which are intervals, are determined by models (12)–(13). The overall belief structures in Table 9 are finally converted into expected risk scores by solving models (15)–(16) for each failure mode. Table 10 presents the expected risk scores of the 21 failure modes, which are visualized in Fig. 1, and their average expected risk scores are computed by Eq. (17). The expected risk scores are ranked using the MRA. The risk priority ranking of the 21 failure modes is presented in the last column of Table 10.

From the overall belief structures in Table 9, it is observed that the overall belief degrees are all intervals. This is because the relative importance weights of the three risk factors are intervals and uncertain. Such a type of uncertainty is referred to as interval uncertainty [30,31]. Take the failure mode FM_{16} for example. The five FMEA team members unanimously evaluate it to the ratings 1, 8, and 3 on O, S and D three risk factors, respectively. The final overall assessment,

Table 6
Severity assessment by FMEA team members.

Failure mode	FMEA team member				
	1	2	3	4	5
1	7: 20%, 8: 80%	8	8	6–7: 50%, 8–9: 50%	8
2	8	8	8	8	[redacted]
3	7-9: 90%	8	6–8	8	8
4	8	8	8	8	7-9: 80%
5	8	7-9: 90%	8	8	8
6	8	8	8	6–8	8
7	9: 75%, 8: 25%	9	9	9	9
8	4	4	4	4: 50%, 5: 50%	3–5: 75%, 6–7: 25%
9	2	2	2	2	2
10	2	2	1–2: 60%, 3–4: 40%	2	2
11	2	2–3	2	2	2
12	3	3	3: 60%, 4: 40%	3	3
13	2–3: 80%, 3–4: 20%	3	3	3	3
14	7	8	7	7	7
15	3	3	3	3	3
16	8	8	8	8	8
17	8	4	8	8	8
18	7	7	7	7	7
19	2	1–2: 75%, 2–3: 25%	2	2	2
20	8	8	[redacted]	8	8
21	3	3	3	3	3

Table 7
Detectability assessment by FMEA team members.

Failure mode	FMEA team member				
	1	2	3	4	5
1	3	3	3	3	3: 90%
2	3	3	3	3	3
3	4	4	4	4	3–4: 80%, 5–6: 20%
4	5	5	5	5	5
5	6	6	6: 85%, 7: 15%	6	6
6	1	1: 85%, 2: 15%	2	[redacted]	1
7	3	2	2	1–2: 75%, 3–4: 25%	2
8	3	3	3: 80%, 4: 20%	3	3
9	3	3	3–4: 60%, 5: 40%	3	3
10	4	4	4	4	4
11	3: 70%, 5: 30%	3	3	3	[redacted]
12	7	7	7	7	7
13	6	6	6	5–7	6
14	4	4	4	4	4
15	4	4: 95%	4	4	4
16	3	3	3	3	3
17	5	5	5	5	5
18	7	6–8	7	7	7
19	4	4	4	8-9: 90%	4
20	4: 60%	9	9	9	9
21	6	6	4–6	6	4: 25%, 5–7: 75%

however, is $\{(1, 20–35\%), (3, 10.29–21.67\%), (8, 43.33–60\%)\}$, which means that the overall belief degree of FM_{16} being assessed to rating 1 is between 20% and 35%, to rating 3 between 10.29% and 21.67%, and to rating 8 between 43.33% and 60%, depending on what values the three risk factor weights take within their intervals. Generally speaking, as long as the original assessment information including

Table 8
Group assessment of the FMEA team on the 21 failure modes with respect to the three major risk factors.

Failure mode	Occurrence	Severity	Detectability
1	{(1, 80%), (1–10, 20%)}	{(7, 6%), (8, 84%), (6–7, 5%), (8–9, 5%)}	{(3, 99%), (1–10, 1%)}
2	{(1, 85%), (2, 15%)}	{(8, 90%), (7–9, 9%), (1–10, 1%)}	{(3, 100%)}
3	{(2, 97%), (1–10, 3%)}	{(6–8, 20%), (8, 50%), (1–10, 30%)}	{(4, 90%), (3–4, 8%), (5–6, 2%)}
4	{(8, 96%), (9, 4%)}	{(8, 90%), (7–9, 8%), (1–10, 2%)}	{(5, 100%)}
5	{(6, 100%)}	{(8, 70%), (7–9, 27%), (1–10, 3%)}	{(6, 97%), (7, 3%)}
6	{(2, 90%), (2–3, 10%)}	{(8, 90%), (6–8, 10%)}	{(1, 65.5%), (2, 24.5%), (1–10, 10%)}
7	{(2, 100%)}	{(8, 7.5%), (9, 92.9%)}	{(2, 60%), (3, 30%), (1–2, 7.5%), (3–4, 2.5%)}
8	{(1, 92.5%), (2, 7.5%)}	{(4, 85%), (5, 5%), (3–5, 7.5%), (6–7, 2.5%)}	{(3, 96%), (4, 4%)}
9	{(3, 70%), (1–10, 30%)}	{(2, 100%)}	{(3, 80%), (5, 8%), (3–4, 12%)}
10	{(1, 84%), (2, 6%), (3, 1.5%), (1–2, 8.5%)}	{(2, 80%), (1–2, 12%), (3–4, 8%)}	{(4, 100%)}
11	{(3–4, 7.5%), (4, 90%), (5, 2.5%)}	{(2, 70%), (2–3, 30%)}	{(3, 81%), (5, 9%), (1–10, 10%)}
12	{(9, 100%)}	{(3, 92%), (4, 8%)}	{(7, 100%)}
13	{(8, 94%), (1–10, 6%)}	{(3, 70%), (2–3, 24%), (3–4, 6%)}	{(6, 90%), (5–7, 10%)}
14	{(3, 80%), (4, 20%)}	{(7, 70%), (8, 30%)}	{(4, 100%)}
15	{(3, 97%), (4, 3%)}	{(3, 100%)}	{(4, 98.5%), (1–10, 1.5%)}
16	{(1, 100%)}	{(8, 100%)}	{(3, 100%)}
17	{(4, 70%), (6, 3%), (3–5, 27%)}	{(4, 30%), (8, 70%)}	{(5, 100%)}
18	{(2, 98%), (1–10, 2%)}	{(7, 100%)}	{(7, 70%), (6–8, 30%)}
19	{(7, 98%), (1–10, 2%)}	{(2, 70%), (1–2, 22.5%), (2–3, 7.5%)}	{(4, 96%), (1–10, 4%)}
20	{(7, 3%), (9, 90%), (8–9, 7%)}	{(8, 80%), (1–10, 20%)}	{(9, 70%), (8–9, 27%), (1–10, 3%)}
21	{(9, 70%), (8–9, 30%)}	{(3, 100%)}	{(4, 2.5%), (6, 70%), (4–6, 20%), (5–7, 7.5%)}

Table 9
Overall assessment of the 21 failure modes.

Failure mode	Overall belief structure for each failure mode
1	{(1, 15.2–26.19%), (3, 10.49–22.15%), (7, 2.69–3.65%), (8, 37.72–51.16%), (6–7, 2.25–3.05%), (8–9, 2.25–3.05%), (1–10, 3.99–6.75%)}
2	{(1, 17.1–29.84%), (2, 3.02–5.27%), (3, 10.35–21.73%), (8, 38.85–53.78%), (7–9, 3.89–5.38%), (1–10, 0.43–0.6%)}
3	{(2, 22.37–36.28%), (4, 10.79–21.17%), (6, 0.04–0.07%), (8, 19.6–26.87%), (3–4, 0.96–1.89%), (5–6, 0.24–0.47%), (6–8, 7.84–10.75%), (1–10, 12.7–16.71%)}
4	{(5, 8.04–17.88%), (8, 77.16–86.43%), (9, 0.74–1.22%), (7–9, 2.82–4.01%), (1–10, 0.71–1%)}
5	{(6, 43.73–60.89%), (7, 0.38–0.72%), (8, 26.89–38.88%), (7–9, 10.37–15%), (1–10, 1.15–1.67%)}
6	{(1, 6.55–13.46%), (2, 24.1–38.61%), (8, 38.21–53.5%), (2–3, 1.98–3.43%), (6–8, 4.25–5.94%), (1–10, 1–2.05%)}
7	{(2, 34.85–51.71%), (3, 2.97–6.02%), (8, 3.01–4.3%), (9, 37.26–53.22%), (1–2, 0.74–1.5%), (3–4, 0.25–0.5%)}
8	{(1, 18.11–31.87%), (2, 1.47–2.58%), (3, 10.5–21.5%), (4, 37.64–51.42%), (5, 2.13–2.94%), (3–5, 3.2–4.41%), (6–7, 1.07–1.47%)}
9	{(2, 42.65–59.17%), (3, 30.62–44.19%), (5, 0.83–1.71%), (3–4, 1.25–2.56%), (1–10, 5.34–8.9%)}
10	{(1, 17.7–29.77%), (2, 37.17–48.83%), (3, 0.3–0.52%), (4, 10.27–21.4%), (1–2, 7.89–9.23%), (3–4, 3.25–4.52%)}
11	{(2, 29.97–40.77%), (3, 10.54–20.17%), (4, 16.97–29.92%), (5, 1.62–2.65%), (2–3, 15.34–19.56%), (1–10, 0.96–1.99%)}
12	{(3, 39.87–55.2%), (4, 3.47–4.8%), (7, 10.29–21.67%), (9, 20–35%)}
13	{(3, 30.67–42.17%), (6, 9.35–19.72%), (8, 18.5–32.22%), (2–3, 10.52–14.46%), (3–4, 2.63–3.61%), (5–7, 1.04–2.19%), (1–10, 1.18–2.06%)}
14	{((3, 15.79–27.36%), (4, 17.48–30.29%), (7, 29.64–41.45%), (8, 12.7–17.76%)}
15	{(3, 81.32–91.29%), (4, 8.59–18.42%), (1–10, 0.12–0.26%)}
16	{(1, 20–35%), (3, 10.29–21.67%), (8, 43.33–60%)}
17	{(4, 35.14–43.94%), (5, 10.65–22.57%), (6, 0.56–0.96%), (8, 27.59–39.03%), (3–5, 5.02–8.6%)}
18	{(2, 16.24–29.75%), (7, 63.95–78.42%), (6–8, 2.81–5.7%), (1–10, 0.33–0.61%)}
19	{(2, 30.54–42.17%), (4, 9.85–20.66%), (7, 19.55–34.16%), (1–2, 9.82–13.55%), (2–3, 3.27–4.52%), (1–10, 1.03–1.55%)}
20	{(7, 0.6–0.96%), (8, 31.58–44.36%), (9, 37.76–52.23%), (8–9, 5.13–7.73%), (1–10, 7.5–10.44%)}
21	{(3, 39.63–56.18%), (4, 0.24–0.49%), (6, 6.89–13.73%), (9, 14.18–24.4%), (4–6, 1.96–3.92%), (5–7, 0.73–1.47%), (8–9, 11.05–15.76%), (1–10, 0.29–0.59%)}

Table 10
Expected risk scores and risk priority rankings of the 21 failure models.

Failure mode	Expected risk score (ERS)		Average ERS	Risk priority ranking
	Minimum	Maximum		
1	4.5144	5.9795	5.2470	10
2	4.4401	5.6816	5.0609	12
3	3.8215	6.0675	4.9445	15
4	7.3854	7.8229	7.6042	2
5	6.5909	7.3014	6.9462	3
6	4.3073	5.6143	4.9608	14
7	4.8387	6.0498	5.4443	9
8	2.7878	3.3429	3.0654	18
9	2.2616	3.2560	2.7588	20
10	1.8879	2.3227	2.1053	21
11	2.5819	3.1918	2.8869	19
12	5.0480	6.0013	5.5247	7
13	4.3393	5.4605	4.8999	13
14	5.1238	5.7961	5.4600	6
15	3.0835	3.2026	3.1431	17
16	4.4667	5.6000	5.0334	11
17	5.2625	5.8249	5.5437	5
18	5.4192	6.2478	5.8335	4
19	3.2105	4.2780	3.7443	16
20	7.6406	8.7400	8.1903	1
21	4.8481	6.0443	5.4462	8

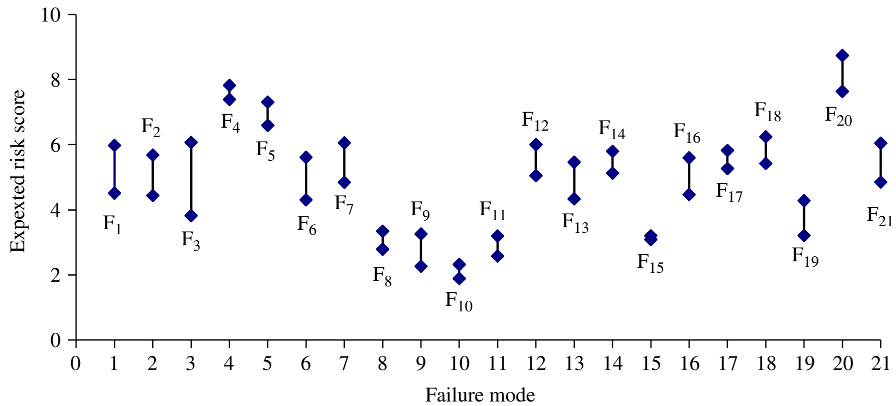

Fig. 1. Visualization of the expected risk scores of the 21 failure modes.

weight information contains uncertainty such as incomplete assessment, interval and/or ignorance, the final overall assessment will be uncertain.

As is clear from Table 10, failure mode 20 has the biggest minimum and maximum expected risk scores and is therefore given a top risk priority, followed by failure modes 4, 5 and 18. On the contrary, failure mode 10 has the smallest minimum and maximum expected risk scores among the 21 failure modes and is thus ranked at the bottom. The 21 failure modes are all completely ranked and distinguished from each other and there is no tie between their risk priority rankings. This is one of the benefits of the use of the ER approach for FMEA.

It is also observed from Fig. 1 that failure mode 15 has the smallest uncertainty and failure mode 3 has the biggest uncertainty (i.e. the widest ERS interval). This is because failure mode 15 is only assessed to two ratings 3 and 4 plus a very small amount of ignorance (missing) information (0.12–0.26%), while failure mode 3 is assessed to multiple ratings 2, 4, 6 and 8 and multiple intervals 3–4, 5–6 and 6–8 plus a significant amount of ignorance information (12.7–16.71%). Therefore, reducing uncertainty in original assessment information can significantly reduce the uncertainty in the final assessment results. This can help the FMEA team to build a more stable risk priority ranking. In another word, in order to build a stable risk priority ranking, FMEA team members should do their best to make their judgments as accurately as possible. For example, if FMEA team member 1 (TM_1) can provide some information rather than ignorance about the assessment of FM_3 with respect to the severity of the failure, say, 6–8, then the expected risk score of FM_3 will narrow down from $[3.8215, 6.0675]$ to $[4.5931, 6.0634]$. If TM_1 provides a more precise judgment, say, $\{(7, 30\%), (8, 70\%)\}$, then $ERS(FM_3)$ will be further improved to $[4.8140, 6.0095]$. Instead of providing an interval 6–8, if TM_3 can also provide a precise assessment, say, $\{(6, 10\%), (7, 90\%)\}$, then $ERS(FM_3)$ will become $[4.8918, 5.8772]$. If the uncertainty involved in the relative importance weights of the three risk factors can be reduced as well, say, from $w_O \in [0.2, 0.35]$, $w_S \in [0.4, 0.5]$ and $w_D \in [0.15, 0.25]$ to $w_O \in [0.25, 0.35]$, $w_S \in [0.4, 0.45]$ and $w_D \in [0.2, 0.25]$, then $ERS(FM_3)$ will be further improved to $[4.8918, 5.4529]$. When the factor weights can be precisely determined, say, $w_O = 0.35$, $w_S = 0.4$ and $w_D = 0.25$,$ERS(FM_3)$ will become $[4.8918, 5.0071]$, which is much narrower and more precise than the original interval $[3.8215, 6.0675]$.

In the above analysis, it is assumed that the five FMEA team members do their evaluations separately and their results are then tabulated and synthesized with their relative importance weights. This is, however, not the only way for the five members to do their assessments. They can also do their evaluations together in a room. The proposed FMEA approach has no special requirements on how the FMEA team members do their evaluations, separately or together. When doing their assessments together in a room, the FMEA team members may try to reach a consensus on every evaluation to be conducted. If they can reach a consensus on all the evaluations, the FMEA team will be viewed as a whole and there will be no need to assign any weight to any member of the team. The team's opinions in this situation are directly modeled as group belief structures. If the FMEA team members cannot reach a consensus on the evaluations they conduct, then the team members' individual opinions should be weighted and synthesized into group belief structures just as they do their evaluations separately. Obviously, no matter how the FMEA team members do their assessments, the proposed FMEA approach is always applicable. This is one of the advantages of the use of the proposed FMEA.

It may be argued that when the FMEA team members do their assessment together in a room, individual belief structures (opinions) may be biased due to the pressure of the group or group thinking. In particular, an inexperienced engineer may not be expected to contribute as much as that by an experienced one and a young engineer could be further influenced away from his/her original opinions. This phenomenon does happen in some methods such as majority rule which requires the minority to be subordinate to the majority. The proposed FMEA approach, however, has no such a phenomenon and does not follow the majority rule. The FMEA team members can express their opinions independently and freely.

4. Conclusions

Considering the fact that FMEA is a group decision function and cannot be done on an individual basis and different FMEA team members may provide different assessment information, we proposed in this paper an FMEA using the group-based ER approach, which can capture FMEA team members' diversity opinions and prioritize failure modes under different types of uncertainties such as incomplete assessment, ignorance and intervals. The core of the proposed FMEA was the development of the risk priority model using the group-based ER approach, which includes assessing risk factors using belief structures, synthesizing individual belief structures into group belief structures and aggregating the group belief structures into overall belief structures, converting the overall belief structures into expected risk scores and ranking the expected risk scores using the MRA. The proposed FMEA was examined with an illustrative application to a fishing vessel and proved to be useful and practical.

Table 11
Recursive combination of three pieces of evidence.

$m \oplus n$			Evidence: m								
			H_{11}	…	H_{1N}	H_{22}	…	H_{2N}	…	H_{NN}	H
			m_{11}	…	m_{1N}	m_{22}	…	m_{2N}	…	m_{NN}	m_H
Evidence: n	H_{11}	n_{11}	$n_{11}m_{11}$ $\{H_{11}\}$	…	$n_{11}m_{1N}$ $\{H_{11}\}$	$n_{11}m_{22}$ $\{\Phi\}$	…	$n_{11}m_{2N}$ $\{\Phi\}$	…	$n_{11}m_{NN}$ $\{\Phi\}$	$n_{11}m_H$ $\{H_{11}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{1N}	n_{1N}	$n_{1N}m_{11}$ $\{H_{11}\}$	…	$n_{1N}m_{1N}$ $\{H_{1N}\}$	$n_{1N}m_{22}$ $\{H_{22}\}$	…	$n_{1N}m_{2N}$ $\{H_{2N}\}$	…	$n_{1N}m_{NN}$ $\{H_{NN}\}$	$n_{1N}m_H$ $\{H_{1N}\}$
	H_{22}	n_{22}	$n_{22}m_{11}$ $\{\Phi\}$	…	$n_{22}m_{1N}$ $\{H_{22}\}$	$n_{22}m_{22}$ $\{H_{22}\}$	…	$n_{22}m_{2N}$ $\{H_{22}\}$	…	$n_{22}m_{NN}$ $\{\Phi\}$	$n_{22}m_H$ $\{H_{11}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{2N}	n_{2N}	$n_{2N}m_{11}$ $\{\Phi\}$	…	$n_{2N}m_{1N}$ $\{H_{2N}\}$	$n_{2N}m_{22}$ $\{H_{22}\}$	…	$n_{2N}m_{2N}$ $\{H_{2N}\}$	…	$n_{2N}m_{NN}$ $\{H_{NN}\}$	$n_{2N}m_H$ $\{H_{2N}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{NN}	n_{NN}	$n_{NN}m_{11}$ $\{\Phi\}$	…	$n_{NN}m_{1N}$ $\{H_{NN}\}$	$n_{NN}m_{22}$ $\{\Phi\}$	…	$n_{NN}m_{2N}$ $\{H_{NN}\}$	…	$n_{NN}m_{NN}$ $\{H_{NN}\}$	$n_{NN}m_H$ $\{H_{NN}\}$
	H	n_H	n_Hm_{11} $\{H_{11}\}$	…	n_Hm_{1N} $\{H_{1N}\}$	n_Hm_{22} $\{H_{22}\}$	…	n_Hm_{2N} $\{H_{2N}\}$	…	n_Hm_{NN} $\{H_{NN}\}$	n_Hm_{NN} $\{H\}$
Nonnormalized probability masses			Sum for $\{H_{11}\}$	…	Sum for $\{H_{1N}\}$	Sum for $\{H_{22}\}$	…	Sum for $\{H_{2N}\}$	…	Sum for $\{H_{NN}\}$	n_Hm_{NN} $\{H\}$
Normalized probability masses			c_{11}	…	c_{1N}	c_{22}	…	c_{2N}	…	c_{NN}	c_H
Evidence: s	H_{11}	s_{11}	$s_{11}c_{11}$ $\{H_{11}\}$	…	$s_{11}c_{1N}$ $\{H_{11}\}$	$s_{11}c_{22}$ $\{\Phi\}$	…	$s_{11}c_{2N}$ $\{\Phi\}$	…	$s_{11}c_{NN}$ $\{\Phi\}$	$s_{11}c_H$ $\{H_{11}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{1N}	s_{1N}	$s_{1N}c_{11}$ $\{H_{11}\}$	…	$s_{1N}c_{1N}$ $\{H_{1N}\}$	$s_{1N}c_{22}$ $\{H_{22}\}$	…	$s_{1N}c_{2N}$ $\{H_{2N}\}$	…	$s_{1N}c_{NN}$ $\{H_{NN}\}$	$s_{1N}c_H$ $\{H_{1N}\}$
	H_{22}	s_{22}	$s_{22}c_{11}$ $\{\Phi\}$	…	$s_{22}c_{1N}$ $\{H_{22}\}$	$s_{22}c_{22}$ $\{H_{22}\}$	…	$s_{22}c_{2N}$ $\{H_{22}\}$	…	$s_{22}c_{NN}$ $\{\Phi\}$	$s_{22}c_H$ $\{H_{22}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{2N}	s_{2N}	$s_{2N}c_{11}$ $\{\Phi\}$	…	$s_{2N}c_{1N}$ $\{H_{2N}\}$	$s_{2N}c_{22}$ $\{H_{22}\}$	…	$s_{2N}c_{2N}$ $\{H_{2N}\}$	…	$s_{2N}c_{NN}$ $\{H_{NN}\}$	$s_{2N}c_H$ $\{H_{2N}\}$
	⋮	⋮	⋮	…	⋮	⋮	…	⋮	…	⋮	⋮
	H_{NN}	s_{NN}	$s_{NN}c_{11}$ $\{\Phi\}$	…	$s_{NN}c_{1N}$ $\{H_{NN}\}$	$s_{NN}c_{22}$ $\{\Phi\}$	…	$s_{NN}c_{2N}$ $\{H_{NN}\}$	…	$s_{NN}c_{NN}$ $\{H_{NN}\}$	$s_{NN}c_H$ $\{H_{NN}\}$
	H	s_H	s_Hc_{11} $\{H_{11}\}$	…	s_Hc_{1N} $\{H_{1N}\}$	s_Hc_{22} $\{H_{22}\}$	…	s_Hc_{2N} $\{H_{2N}\}$	…	s_Hc_{NN} $\{H_{NN}\}$	s_Hc_{NN} $\{H\}$
Nonnormalized probability masses			Sum for $\{H_{11}\}$	…	Sum for $\{H_{1N}\}$	Sum for $\{H_{22}\}$	…	Sum for $\{H_{2N}\}$	…	Sum for $\{H_{NN}\}$	$n_Hm_{NN}\{H\}$
Normalized probability masses			x_{11}	…	x_{1N}	x_{22}	…	x_{2N}	…	x_{NN}	x_H
Aggregated belief degrees δ_{ij}			$\frac{x_{11}}{1-x_H}$		$\frac{x_{1N}}{1-x_H}$	$\frac{x_{22}}{1-x_H}$		$\frac{x_{2N}}{1-x_H}$		$\frac{x_{NN}}{1-x_H}$	–

Note: N is the number of ratings, Φ is the empty/null set and the *sum for* $\{H_{ij}\}$ represents the sum of all the probability masses assigned to the rating/interval H_{ij}, seeing $\{H_{11}\}$ for example, whose probability masses are highlighted and shaded.

In comparison with the traditional RPN and its variants, the proposed FMEA has the following advantages:

- The relative importance weights of risk factors are considered. They can not only be deterministic but also uncertain such as intervals or preference order.
- Risk factors are aggregated in a highly nonlinear manner which is neither the simple addition nor the simple product of the risk factors.
- The diversity and uncertainty of FMEA team members' assessment information can be well reflected and modeled using belief structures.
- Failure modes can be fully ranked and well distinguished from each other unless some of them are assessed to be the same.
- Expected risk score is a continuous number from 1 to 10 without any holes, which is either a crisp number or an interval.
- More risk factors can be included if necessary. The proposed FMEA is not limited to *O*, *S* and *D*, but applicable to any number of risk factors.
- There is no need to build any rule bases which are highly subjective. Different experts may make distinct judgments, leading to different rules.

Acknowledgments

The authors would like to thank one anonymous reviewer for his/her constructive comments and suggestions which are very helpful in improving the paper. The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China, project no. CityU 111906 and also partially supported by the Natural Science Foundation of Fujian Province of China under the project no. A0710005.

Appendix A. Recursive combination manner of multiple pieces of evidence

For the convenience of the readers to understand and implement the interval ER algorithm, we provide Table 11 to show how three pieces of evidence can be combined on a microsoft excel worksheet in a recursive way. If there are more pieces of evidence to be combined, they can be recursively combined in the same way.

References

[1] Linton JD. Facing the challenges of service automation: an enabler for e-commerce and productivity gain in traditional services. IEEE Transactions on Engineering Management 2003;50(4):478–84.
[2] Stamatis DH. Failure Mode and Effect Analysis: FMEA from Theory to Execution. Milwaukee, Wisconsin: ASQC Quality Press; 1995.
[3] Vliegen HJW, van Mal HH. Rational decision making: structuring of design meetings. IEEE Transactions on Engineering Management 1990;37(3):185–91.
[4] Bowles JB. An assessment of PRN prioritization in a failure modes effects and criticality analysis. Journal of the IEST 2004;47:51–6.
[5] Sankar NR, Prabhu BS. Modified approach for prioritization of failures in a system failure mode and effects analysis. International Journal of Quality & Reliability Management 2001;18(3):324–35.
[6] Xu K, Tang LC, Xie M, Ho SL, Zhu ML. Fuzzy assessment of FMEA for engine systems. Reliability Engineering & System Safety 2002;75:17–29.
[7] Ben-Daya M, Raouf A. A revised failure mode and effects analysis model. International Journal of Quality & Reliability Management 1996;13(1):43–7.
[8] Braglia M, Frosolini M, Montanari R. Fuzzy criticality assessment model for failure modes and effects analysis. International Journal of Quality & Reliability Management 2003;20(4):503–24.
[9] Chang CL, Liu PH, Wei CC. Failure mode and effects analysis using gray theory. Integrated Manufacturing Systems 2001;12(3):211–6.
[10] Gilchrist W. Modelling failure modes and effects analysis. International Journal of Quality & Reliability Management 1993;10(5):16–23.
[11] Pillay A, Wang J. Modified failure mode and effects analysis using approximate reasoning. Reliability Engineering & System Safety 2003;79:69–85.
[12] Bevilacqua M, Braglia M, Gabbrielli R. Monte Carlo simulation approach for a modified FMECA in a power plant. Quality and Reliability Engineering International 2000;16:313–24.
[13] Braglia M. MAFMA: multi-attribute failure mode analysis. International Journal of Quality & Reliability Management 2000;17(9):1017–33.
[14] Braglia M, Frosolini M, Montanari R. Fuzzy TOPSIS approach for failure mode, effects and criticality analysis. Quality and Reliability Engineering International 2003;19:425–43.
[15] Chang CL, Wei CC, Lee YH. Failure mode and effects analysis using fuzzy method and gray theory. Kybernetes 1999;28:1072–80.
[16] Deng JL. Control problems of gray systems. Systems & Control Letters 1982;1(5):211–5.
[17] Seyed-Hosseini SM, Safaei N, Asgharpour MJ. Reprioritization of failures in a system failure mode and effects analysis by decision making trial and evaluation laboratory technique. Reliability Engineering & System Safety 2006;91:872–81.
[18] Bowles JB, Peláez CE. Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. Reliability Engineering & System Safety 1995;50:203–13.
[19] Guimarães ACF, Lapa CMF. Fuzzy inference to risk assessment on nuclear engineering systems. Applied Soft Computing 2007;7:17–28.
[20] Guimarães ACF, Lapa CMF. Effects analysis fuzzy inference system in nuclear problems using approximate reasoning. Annals of Nuclear Energy 2004;31(1):107–15.
[21] Puente J, Pino R, Priore P, de la Fuente D. A decision support system for applying failure mode and effects analysis. International Journal of Quality & Reliability Management 2002;19(2):137–50.
[22] Sharma RK, Kumar D, Kumar P. Systematic failure mode effect analysis (FMEA) using fuzzy linguistic modeling. International Journal of Quality & Reliability Management 2005;22(9):986–1004.
[23] Tay KM, Lim CP. Application of fuzzy inference techniques to FMEA. In: Abraham A, de Baets B, Köppen M, Nickolay B, editors. Applied soft computing technologies: the challenge of complexity. Berlin, Heidelberg: Springer; 2006.
[24] Tay KM, Lim CP. Fuzzy FMEA with a guided rules reduction system for prioritization of failures. International Journal of Quality & Reliability Management 2006;23(8):1047–66.
[25] Zafiropoulos EP, Dialynas EN. Reliability prediction and failure mode effects and criticality analysis (FMECA) of electronic devices using fuzzy logic. International Journal of Quality & Reliability Management 2005;22(2):183–200.
[26] Naude P, Lockett G, Holmes K. A case study of strategic engineering decision making using judgmental modeling and psychological profiling. IEEE Transactions on Engineering Management 1997;44(3):237–44.
[27] Guo M, Yang JB, Chin KS, Wang H. Evidential reasoning based preference programming for multiple attribute decision analysis under uncertainty. European Journal of Operational Research 2007;182(3):1294–312.
[28] Li Y, Liao X. Decision support for risk analysis on dynamic alliance. Decision Support Systems 2007;42:2043–59.
[29] Wang YM, Yang JB, Xu DL. Environmental impact assessment using the evidential reasoning approach. European Journal of Operational Research 2006;174(3):1885–913.
[30] Wang YM, Yang JB, Xu DL, Chin KS. The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. European Journal of Operational Research 2006;175:35–66.
[31] Xu DL, Yang JB, Wang YM. The evidential reasoning approach for multi-attribute decision analysis under interval uncertainty. European Journal of Operational Research 2006;174(3):1914–43.
[32] Yang JB, Wang YM, Xu DL, Chin KS. The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. European Journal of Operational Research 2006;171(1):309–43.
[33] Yang JB, Xu DL. On the evidential reasoning algorithm for multiattribute decision analysis under uncertainty. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans 2002;32(3):289–304.
[34] Yang JB, Xu DL. Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty. IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans 2002;32(3):376–93.
[35] Shafer GA. Mathematical theory of evidence. Princeton, NJ: Princeton University Press; 1976.
[36] Doyle JR, Green RH, Bottomley PA. Judging relative importance: direct rating and point allocation are not equivalent. Organizational Behavior and Human Decision Processes 1997;70:55–72.
[37] Roberts R, Goodwin P. Weight approximations in multi-attribute decision models. Journal of Multi-Criteria Decision Analysis 2002;11:291–303.
[38] Takeda E, Cogger KO, Yu PL. Estimating criterion weights using eigenvectors: A comparative study. European Journal of Operational Research 1987;29:360–9.
[39] Srinivasan V, Shocker AD. Linear programming techniques for multidimensional analysis of preferences. Psychometrika 1973;38:337–69.
[40] Hwang CL, Yoon K. Multiple attribute decision-making: methods and applications. Berlin: Springer; 1981.
[41] Wang YM, Greatbanks R, Yang JB. Interval efficiency assessment using data envelopment analysis. Fuzzy Sets and Systems 2005;153:347–70.