Expert Systems with Applications 42 (2015) 7235–7244

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Clustering and visualization of failure modes using an evolving tree

Wui Lee Chang [a], Kai Meng Tay [a,*], Chee Peng Lim [b]

[a] Faculty of Engineering, Universiti Malaysia Sarawak, Sarawak, Malaysia
[b] Centre for Intelligent Systems Research, Deakin University, Geelong Waurn Ponds Campus, 75 Pigdons Road, Waurn Ponds, Victoria 3216, Australia

ARTICLE INFO

Article history:
Available online 7 May 2015

Keywords:
Clustering
Visualization
Failure Mode and Effect Analysis
Evolving tree
Edible bird nest industry
Euclidean similarity
Neural network

ABSTRACT

Despite the popularity of Failure Mode and Effect Analysis (FMEA) in a wide range of industries, two well-known shortcomings are the complexity of the FMEA worksheet and its intricacy of use. To the best of our knowledge, the use of computation techniques for solving the aforementioned shortcomings is limited. As such, the idea of clustering and visualization pertaining to the failure modes in FMEA is proposed in this paper. A neural network visualization model with an incremental learning feature, i.e., the evolving tree (ETree), is adopted to allow the failure modes in FMEA to be clustered and visualized as a tree structure. In addition, the ideas of risk interval and risk ordering for different groups of failure modes are proposed to allow the failure modes to be ordered, analyzed, and evaluated in groups. The main advantages of the proposed method lie in its ability to transform failure modes in a complex FMEA worksheet to a tree structure for better visualization, while maintaining the risk evaluation and ordering features. It can be applied to the conventional FMEA methodology without requiring additional information or data. A real world case study in the edible bird nest industry in Sarawak (Borneo Island) is used to evaluate the usefulness of the proposed method. The experiments show that the failure modes in FMEA can be effectively visualized through the tree structure. A discussion with FMEA users engaged in the case study indicates that such visualization is helpful in comprehending and analyzing the respective failure modes, as compared with those in an FMEA table. The resulting tree structure, together with risk interval and risk ordering, provides a quick and easily understandable framework to elucidate important information from complex FMEA forms; therefore facilitating the decision-making tasks by FMEA users. The significance of this study is twofold, *viz.*, the use of a computational visualization approach to tackling two well-known shortcomings of FMEA; and the use of ETree as an effective neural network learning paradigm to facilitate FMEA implementations. These findings aim to spearhead the potential adoption of FMEA as a useful and usable risk evaluation and management tool by the wider community.

1. Introduction

Clustering is a process of organizing a set of data attributed by multi-dimensional features into different groups based on a similarity measure (Rui & Donald, 2009). Usually, each group of data is represented by a unique weight vector, e.g. the centroid of the group (Rui & Donald, 2009). Clustering methods are useful in many applications, e.g. data mining (Lan, Frank, & Hall, 2005), data query (Lan et al., 2005), robotic arm movements (Kohonen, Simula, & Visa, 1996), noise reduction in telecommunication (Kohonen, 2001), and image segmentation (Chang, Luo, & Parker, 1998). Examples of popular clustering methods include the self-organizing map (SOM) (Vesanto & Alhoniemi, 2000), the evolving tree (ETree) (Pakkanen, Iivarinen, & Oja, 2006), fuzzy ART (Keskin & Özkan, 2009), as well as k-means (Chang et al., 1998) and fuzzy c-means (Rezaee, Leliveldt, & Reiber, 1998) clustering algorithms.

The SOM model is a neural network capable of mapping high dimensional data samples onto a lower dimensional space and representing them as nodes (Kohonen et al., 1996; Kohonen, 2001; Vesanto & Alhoniemi, 2000). It also provides a topological view of the underlying data structure (Kohonen et al., 1996; Kohonen, 2001; Vesanto & Alhoniemi, 2000). A number of enhanced SOM models have been proposed, e.g. growing SOM (GSOM) (Matharage, Alahakoon, Rajapakse, & Pin, 2011; Kuo, Wang, & Chen, 2012), growing hierarchical SOM (GHSOM) (Huang & Tsaih, 2012), and ETree (Pakkanen, Iivarinen, & Oja, 2004, 2006). These enhancements overcome two shortcomings of SOM, i.e., the requirement of a pre-defined map size before learning (Kohonen, 2001; Vesanto & Alhoniemi, 2000) and the long learning time when a large map size is initiated (Pakkanen et al., 2004, 2006). GSOM starts with a small map, and nodes are added during the

* Corresponding author.
E-mail addresses: wuileechang@gmail.com (W.L. Chang), kmtay@feng.unimas.my (K.M. Tay), chee.lim@deakin.edu.au (C.P. Lim).

http://dx.doi.org/10.1016/j.eswa.2015.04.036
0957-4174/

learning process; therefore realising a more efficient method with an incremental learning capability (Matharage et al., 2011). GHSOM has a hierarchical architecture, whereby it has a SOM-like adaptive architecture that builds various layers of hierarchy (Huang & Tsaih, 2012). For ETree, a tree structure is adopted whereby nodes are allowed to grow freely when new data samples are available (Pakkanen et al., 2004, 2006). The growing structure of ETree is important for analyzing complex data structures, e.g. in image clustering problems (Pakkanen et al., 2004, 2006). In our preliminary investigation, ETree has been shown useful for tackling textual document clustering problems (Chang, Tay, & Lim, 2013, 2014). The ETree-based approach increases the flexibility of clustering by allowing new clusters to be formed and updated, in response to new textual documents.

In this paper, the focus is on applying ETree to clustering and visualization of failure modes with the Failure Mode and Effect Analysis (FMEA) methodology. FMEA is an effective problem prevention and risk analysis methodology for defining, identifying, and eliminating failures of a system, design, process, or service (Stamatis, 2003). It has been used in a wide variety of application domains, e.g., aerospace (Bowles & Peláez, 1995), automotive (Stamatis, 2003), nuclear (Guimarães & Lapa, 2004), electronic (Zafiropoulos & Dialynas, 2005), manufacturing (Tay & Lim, 2006), chemical (Garrick, 1988), mechanical (Korayem & Iravani, 2008), healthcare and hospital (McNally, Page, & Sunderland, 1997), and agriculture (Jong, Tay, & Lim, 2013). FMEA identifies the failure modes of a system or process, understands the causes and effects of each failure mode, and determines suitable actions to eliminate or reduce the risk of the respective failure modes (Stamatis, 2003). Traditionally, the risk of a failure mode is determined by computing the Risk Priority Number (RPN) (Stamatis, 2003). The RPN model considers three factors as its inputs, i.e. Severity (S), Occurrence (O), and Detection (D), and produces an *RPN* score (i.e. multiplication of S, O, and D) as the output (Stamatis, 2003). S and O are seriousness and frequency of a failure, respectively, while D is the effectiveness of the existing measures in detecting a failure before the effect of the failure reaches the customer(s) (Stamatis, 2003).

While the effectiveness of FMEA has been demonstrated, the traditional RPN model is susceptible to a number of limitations (Bowles & Peláez, 1995; Tay & Lim, 2006; Liu, Liu, & Liu, 2013a). Many risk evaluation methods which can be used as an alternative to the traditional RPN model have been investigated. According to the review by Liu et al. (2013a), the existing risk evaluation methods can be grouped into five categories, i.e., multi-criteria decision making (MCDM) methods, mathematical programming methods, artificial intelligence methods, integrated methods, and other methods. Recently, a number of new methods for risk evaluation and/or ranking have also been developed. Bozdag, Asan, Soyer, and Serdarasan (2015) proposed an interval type-2 fuzzy set to capture both intra-personal and inter-personal uncertainties. Chang (2014) developed a soft set-based ranking technique for the prioritization of failure modes. Du, Mo, Deng, Sadiq, and Deng (2014) proposed a hybrid evidential reasoning (ER) and TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)-based method for group assessment in FMEA. ER was used to express FMEA users' assessments that contain imprecision and uncertainty. TOPSIS was then used to aggregate the risk factors. Liu, Liu, and Lin (2013b) proposed a fuzzy ER and belief rule-based method for risk evaluation and ranking. The Dempster rule of combination was used to aggregate all relevant rules.

A number of new fuzzy sets related approaches have also been reported. These include an intuitionistic fuzzy set method with the weighted Euclidean distance (Liu, Liu, & Li, 2014a), a fuzzy weighted average with fuzzy decision-making method (Liu, You, Lin, & Li, 2014c), a fuzzy set theory and a multi-MOORA (multi-objective optimization by ratio analysis) method (Liu, Fan, Li, & Chen, 2014e), a hybrid intuitionistic fuzzy-TOPSIS method (Liu, You, Shan, & Shao, 2014f), and a hybrid fuzzy digraph and matrix method (Liu, Chen, You, & Li, 2014h). Besides that, Liu, You, Fan, and Lin (2014b) also proposed an D numbers and grey relational projection method, in which the assessment results were expressed in D numbers. Liu, Li, You, and Chen (2014d) presented an FMEA method comprising interval 2-tuple linguistic variables with gray relational analysis to capture FMEA users' diverse opinions. Liu, You, and You (2014g) proposed an interval 2-tuple hybrid weighted distance measure –based method for risk evaluation. In addition, a fuzzy SIRM (Single-Input-Rule-Module)-based RPN model (Jong, Tay, & Lim, 2014) and a two-stage Sugeno fuzzy-based RPN model with similarity reasoning (Jee, Tay, & Lim, 2015) were also proposed.

In our previous research, fuzzy inference systems (Tay & Lim, 2006; Jong et al., 2013, 2014; Jee et al., 2015) and an adaptive clustering method, i.e., fuzzy adaptive reasoning theory (Tay, Jong, & Lim, 2015), have been applied to FMEA. Based on the findings, we have proven the importance of maintaining the monotonicity relationship between inputs (S, O, D) and the output (*RPN* score) (Tay & Lim, 2008a, 2008b; Jee et al., 2015). Our works (Tay & Lim, 2006; Jong et al., 2013, 2014; Jee et al., 2015) focus on the development of fuzzy inference system–based frameworks for risk evaluation, with the aim of reducing the number of fuzzy rules while maintaining the monotonicity property. Fuzzy ART was used by Keskin and Özkan (2009) as a clustering method to tackle the problem whereby different combinations of S, O, and D could produce the same *RPN* scores. In addition to this reason, we further justify the advantages of using clustering methods in FMEA, as follows: (1) clustering deals with the original S, O, and D scores directly; (2) clustering allows the failure modes to be compared and visualized in the input space as groups of information; (3) the use of the original S, O, and D scores (instead of the mapped S, O, and D scores into a common domain) avoids loss of information or modification of important information for decision making purposes.

In the literature, studies on combining clustering (sometime together with visualization) and FMEA (or risk management) are not new. Arunajadai, Uder, Stone, and Tumer (2004) proposed a statistical clustering procedure to identify potential failures of FMEA. A function-failure matrix that can be used as a knowledge base to identify and analyze potential failures for new designs and redesign as well as to allow grouping of the failure modes was developed. The underlying principle was based on similarities between different failure modes pertaining to the product/component functionality. The resulting failure modes were then arranged (or visualized) in a table according to their types and clusters. The importance of grouping failure modes was again suggested by Mandal and Maiti (2014) recently. They suggested the use of a similarity to group failure modes that have similar risk levels. Besides that, Romuald Iwańkowicz and Rosochacki (2014) used a clustering method to process and analyze a database of accidents and predict the process risk. Recently, Li, Chen, and Xiang (2015) proposed a grey clustering-based indicator system to avoid the arbitrary selection of indicators in risk management of an airport safety evaluation program in China.

Studies on combining visualization and FMEA (or risk analysis and management) are also available in the literature. Inoue and Yamada (2010) visualized complicated processes in pharmaceutical research using FMEA. Grøndahl, Lund, and Stølen (2011) investigated how graphical effects (e.g., size, color, shape) and text labels introduced in the CORAS risk modeling language affected the understanding of the subject of interest. Wintle and Nicholson (2014) used Bayesian networks as graphical modeling tools for exploring what-if scenarios, visualizing systems and problems, and communicating with stakeholders during decision

making. In risk analysis, the failure risks are at times visualized with respect to the likelihood and impact (a.k.a. S) dimensions, i.e., a two-dimensional space (Feather, Cornford, Kiper, & Menzies, 2006). Such visualization is useful for understanding the distribution of risks. However, the method is not applicable to FMEA, which involves a three-dimensional input (i.e., S, O, and D) space.

Building on the findings of our previous investigations (Tay et al., 2015), the main aim of this paper is to examine the use of the ETree for analyzing (i.e., both clustering and visualization) of failure modes in FMEA, instead of just performing clustering only as in the use of fuzzy ART (Tay et al., 2015). To the best of our knowledge, even though studies on synthesizing clustering and/or visualization with FMEA is not new, little attention is focused on clustering and visualization of the failure modes in FMEA, as conducted in this paper. In addition, the method presented in this study is useful for tackling two other important and practical issues related to FMEA implementation (Signor, 2002; Montgomery, Pugh, Leedham, & Twitchett, 1996), i.e., (1) complexity of the FMEA worksheet; (2) intricacy of its use. Entries in a FMEA worksheet are voluminous (Signor, 2002). An example of an FMEA table (e.g. see Stamatis, 2003, pages 231-242) could take up to 11 pages. The FMEA worksheet is hard to produce (Signor, 2002; Montgomery et al., 1996), hard to understand and read, as well as hard to maintain (Signor, 2002; Montgomery et al., 1996). From our literature research, limited investigations on using computing techniques to solve the aforementioned issues have been reported so far. As a result, we incorporate visualization techniques into FMEA, in an attempt to offer an effective solution for both aforementioned issues. Visualization serves as a communication means between FMEA and its users. A good visualization (or effective communication) of the failure modes or corrective actions is important because (1) it presents the failure modes or actions as a structure that is easy to understand, as compared with the original failure modes or actions in complex FMEA worksheets; (2) it allows users to access or analyze FMEA with a large number of failure modes and corrective actions quickly, which may not be otherwise possible; (3) it provides users an overview of all failure modes and corrective actions, which mitigates the problem in having a good understanding and insight into the overall FMEA exercise in situations where no visualization is available; (4) it leads to more efficient processes for making decisions and taking actions.

Compared with other existing methods, our proposed method does not focus on risk evaluation issues. Instead, we tackle issues related to complexity of the FMEA worksheet and difficulties in understanding and maintaining the FMEA table. The main advantages of the proposed method lie in its ability to transform the failure modes in a complicated FMEA worksheet to a tree structure for better visualization, while maintaining the risk evaluation and ordering features. The proposed method is applicable to the conventional FMEA methodology without requiring additional information or data. It can also be used with other existing methods, e.g., the fuzzy inference system–based RPN model (Bowles & Peláez, 1995; Jee et al., 2015), for obtaining more accurate risk evaluation and ordering outcomes. Nevertheless, one limitation of the proposed method is that it requires consensus from FMEA users with different expertise levels, in order to provide a meaningful visualization covering all the S, O, and D scores.

In this study, ETree is chosen for its adaptive feature, flexibility to grow or expand with new data samples, and ease of representation in the form of a tree for visualization and clustering purposes. ETree is a type of neural network (Pakkanen et al., 2004; Pakkanen et al., 2006), which can create its own organization of the information it receives during learning. Such visualization is useful as it provides a quick and easily understandable representation of the FMEA worksheet, which is usually complex and lengthy, to facilitate decision making tasks. In this paper, we depict the characteristics and relationship of the failure modes or actions in a tree structure. Two concepts, i.e., risk interval and risk ordering of clusters, are introduced to allow failure modes to be analyzed in a group. To evaluate the proposed method, a real-world case study from the edible bird nest (EBN) industry is used. The experimental results show that the complicated and lengthy FMEA worksheets pertaining to the EBN processes can be visualized as comprehensible tree structures. From the discussion with FMEA users engaged in the case study, such visualization is helpful in understanding and analyzing the failure modes, instead of studying information in the complex FMEA worksheets.

This paper is organized as follows. In Section 2, the background of FMEA and the RPN model are explained. In Section 3, the Euclidean similarity–based metric is formulated, and the use of ETree in FMEA is described. In Section 4, the experimental results are presented and discussed. Finally, concluding remarks are provided in Section 5.

2. Preliminaries

To make this paper self-contained, the traditional RPN and FIS-based RPN models are explained. The distance-based similarity metric (Tay et al., 2015) is also reviewed.

2.1. Severity, occurrence, detection and risk priority number

Traditionally, FMEA adopts the RPN model, which considers three risk factors, i.e., S, O, and D, for analyzing the failure modes and prioritizing the corresponding actions. These three risk factors form the input space, as follows.

Definition 1. An input space, i.e., $S \times O \times D$, is considered. Variables s, o, and d are the elements of S, O, and D, respectively, i.e., $s \in$ S, $o \in$ O, and $d \in$ D. The lower and upper boundaries of S is represented by $\underline{s}$ and $\bar{s}$, respectively. Similarly, the lower or upper boundaries of O and D are represented by $\underline{o}$ and $\bar{o}$, as well as $\underline{d}$ and $\bar{d}$, respectively.

A set of data samples in the S, O and D space, as defined in Definition 1, i.e., $[s, o, d]$, is considered. Traditionally, the risk of $[s, o, d]$ is compared with those from other data samples in the *RPN* space (i.e. the output space) according the following definition.

Definition 2. The *RPN* space is the output space containing all possible *RPN* scores. The lower and upper boundaries of the RPN space is represented by $\underline{RPN}$ and $\overline{RPN}$, respectively. The *RPN* space follows a monotonic, ordered sequence, i.e., the higher the *RPN* score, the higher the risk.

2.2. The traditional RPN model

Traditionally, the *RPN* (output) score of an input $[s, o, d]$ is obtained using Eq. (1):

$$RPN = s \times o \times d \tag{1}$$

Eq. (1) can be viewed as a mapping of $[s, o, d]$ to the *RPN* space.

2.3. The Euclidean distance-based similarity metric

A set of failure modes as defined in Definition 3 is considered. Note that each failure mode is represented by a data sample in the input (S, O, and D) space.

Definition 3. A set of data samples with m failure modes is considered. Each failure mode is denoted as $\overline{x_k} = [s_k, o_k, d_k]$ in the S, O, and D space (Definition 1), $k = 1, 2, 3, \ldots, m$.

An Euclidean distance-based similarity metric for two failure modes, i.e., $\overline{x_k} = [s_k, o_k, d_k]$ and $\overline{x_j} = [s_j, o_j, d_j]$, $j, k = 1, 2, 3, \ldots, m$, is computed using Eq. (2).

$$Euclidean\ Similarity(\overline{x_k}, \overline{x_j}) = 1 - \sqrt[2]{\frac{(s_k - s_j)^2 + (o_k - o_j)^2 + (d_k - d_j)^2}{(\bar{s} - \underline{s})^2 + (\bar{o} - \underline{o})^2 + (\bar{d} - \underline{d})^2}} \quad (2)$$

3. The proposed methodology

In this section, the ETree model for FMEA application is described in detail, as follows.

3.1. The evolving tree

Fig. 1 depicts the structure of ETree. The tree consists of a total of n_{node} nodes. Each node is denoted as $N_{l,p}$, where $l = 1, 2, 3, \ldots, n_{node}$ is the identity of a node, and $p = 0, 1, 2, \ldots$ is its parent node. As an example, the parent node of $N_{2,1}$ is $N_{1,0}$, or $N_{2,1}$ is a child node of $N_{1,0}$. There are three types of nodes, i.e., a root node, trunk nodes, and leaf nodes, $N_{leaf}, leaf \in l$. The root node is the first node created in ETree (e.g., $N_{1,0}$ in Fig. 1). It is located at the top layer of the tree, and has no parent node (i.e., $j = 0$). A trunk node is a parent node (labelled as white in Fig. 1). A leaf node, which is located at the bottom layer of a tree, has no child node (labelled as black in Fig. 1). Each node is also attributed with a weight vector (i.e., $W_l = [w_{l,S}, w_{l,O}, w_{l,D}]$) and a Best Matched Unit (BMU) hit counter (i.e., b_l).

The root node and trunk nodes provide a series of paths connecting a leaf node to other leaf nodes. The tree distance of two leaf nodes, i.e., N_{l_1} and $N_{l_2}, l_1, l_2 \in l$, is denoted as $d_T(N_{l_1}, N_{l_2})$. Note that is determined by the number of trunk nodes (including the root node) that connect and N_{l_2} at the shortest path. As an example, $d_T(N_4, N_7) = 3$, i.e., the number of trunk nodes connecting $N_{4,2}$ and $N_{7,3}$ through the shortest path is 3 (2 trunk nodes and the root node). The tree size is measured by the number of nodes, i.e., n_{node}. Based on the example in Fig. 1, the tree size is 9. The depth of a tree is the number of layers, e.g. the depth of the tree in Fig. 1 is 4.

The ETree learning algorithm for clustering and visualization of failure modes is depicted in Fig. 2. Each failure mode, i.e., $\overline{x_k}$, is provided to ETree for learning. There are three parameters that need to be predefined before learning, i.e., $epoch, \theta_{splitting}$, and $\theta_{childnode}$. Parameter $epoch$ is the number of iterations, i.e., the number of times $\overline{x_k}, k = 1, 2, 3, \ldots, m$, should be learned. Parameter $\theta_{splitting}$ is the splitting threshold, which is also known as the tree growing rate. Parameter $\theta_{childnode}$ is the number of child nodes generated when a node is expanded.

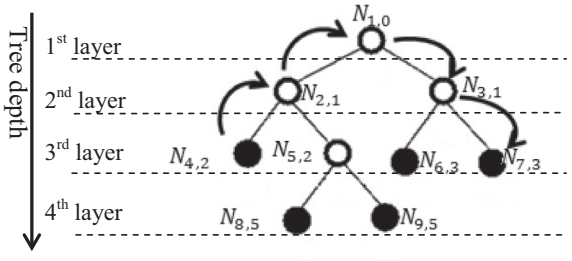

Fig. 1. The structure of ETree.

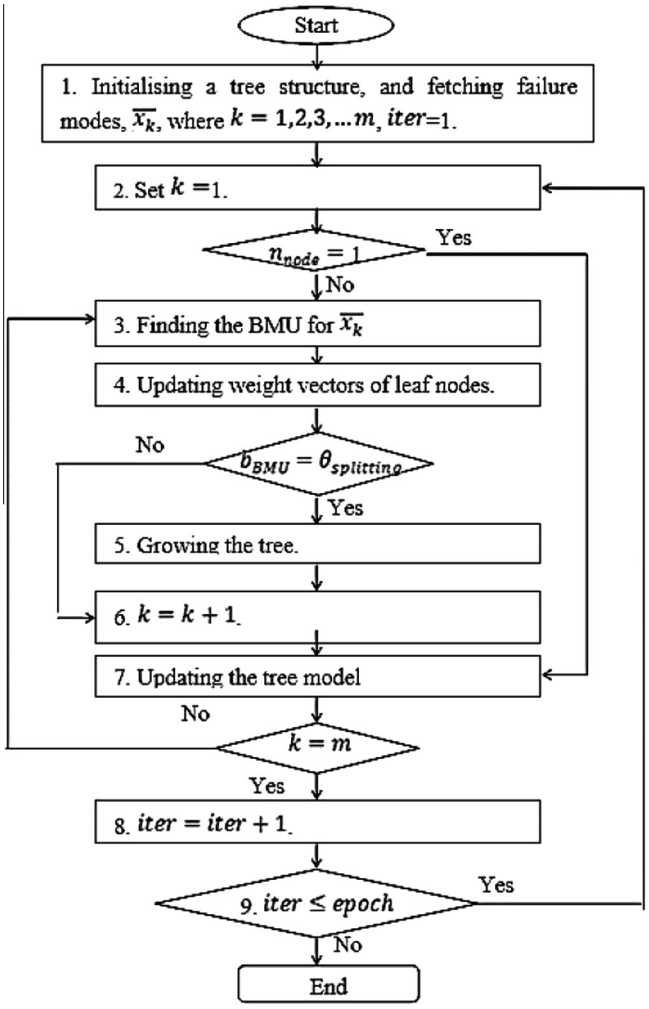

Fig. 2. The ETree learning algorithm for clustering and visualization of failure modes in FMEA.

The important steps of ETree, as depicted in Fig. 2, are explained, as follows.

Step1: Initialise the tree structure and provide a set of failure modes. A root node is created with a weight vector of $W_1 = [111]$, and $l = n_{node} = 1$. Then, the failure modes, $\overline{x_k}, = 1, 2, 3, \ldots, m$, are processed.

Step2: Set $k = 1$.

Step3: Find the BMU for $\overline{x_k}$. $\overline{x_k}$ is matched with the child nodes of the root note using the Euclidean distance-based metric (Eq. (2)). $Euclidean\ Similarity(\overline{x_k}, \overline{x_j})$ is obtained. The node with the highest score of $Euclidean\ Similarity(\overline{x_k}, \overline{x_j})$ is the BMU. If the BMU $\notin$ leaf nodes, $\overline{x_k}$ is matched with the child nodes of the BMU using the same manner, and the BMU is determined among the child nodes again. If there are more than one nodes with the same score of $Euclidean\ Similarity(\overline{x_k}, \overline{x_j})$, the node with the largest index is chosen. If the BMU $\in$ leaf nodes, it is N_{BMU}.

Step4: Update the weight vectors of the leaf nodes. If $n_{node} = 1$, the root node is assumed to be the leaf node. The Kohonen learning rule (Kohonen, 2001) is applied to update the leaf nodes, i.e.,

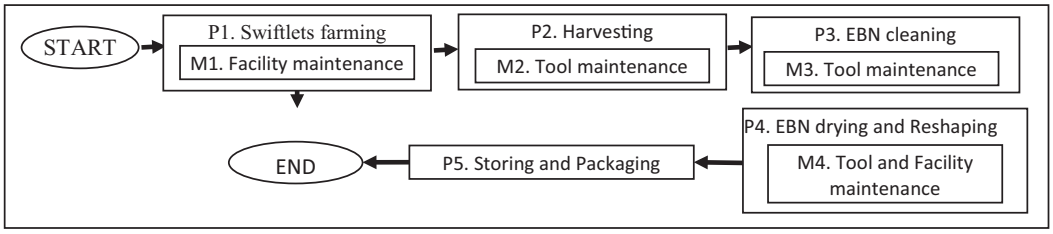

Fig. 3. The process of EBN production (Jong et al., 2013).

$$W_1(new) = W_1(old) + \left[\propto(iter) \times exp\left(\frac{-d_T(N_{\mathrm{BMU}}, N_{leaf})^2}{2\sigma^2(iter)}\right)\right] \times (\overline{x_k} - W_1(old)) \quad (3)$$

where $\propto(iter)$ and $\sigma(iter)$ decrease monotonically as *iter* increases. Update the b_{BMU} counter, i.e., $b_{BMU} = b_{BMU} + 1$.

Step5: Grow the tree, i.e., N_{BMU} is split to $\theta_{childnode}$ nodes. The weight vectors of $\theta_{childnode}$ nodes are cloned from N_{BMU}.

Step6: Set $k = k + 1$.

Step7: Update the tree structure.

Step8: Set $iter = iter + 1$.

Step9: If $iter \leqslant epoch$, go to Step 2. Otherwise, end.

3.2. Risk interval of a cluster

In this paper, the risk associated with the failure modes in each cluster is represented as a risk interval, i.e., $RPN_{leaf} = \left[\underline{RPN_{leaf}}, \overline{RPN_{leaf}}\right]$. Eq. (1) is used to obtain the *RPN* score of each $\overline{x_k}$ pertaining to a cluster (i.e., $k \in N_{leaf}$), and is denoted as $RPN(\overline{x_k})$. The risk interval of N_{leaf} that contains a group of failure modes is obtained using Eqs. (4) and (5).

$$\underline{RPN_{leaf}} = \min\left(RPN(\overline{x_k}), \text{ where } k \in N_{leaf}\right) \quad (4)$$

$$\overline{RPN_{leaf}} = \max\left(RPN(\overline{x_k}), \text{ where } k \in N_{leaf}\right) \quad (5)$$

4. Experimental study

In this section, a series of experiments with real data/information from the EBN industry in Sarawak, Borneo Island, is presented. EBN (or known as "the *Caviar of the East*") is the nest of swiftlets, which is consumed as a type of popular (healthy) food (Hobbs, 2004). With a high demand of EBN from China, swiftlets farming and EBN processing have emerged as a major urban industry among Southeast Asia countries, including Malaysia (Jordan, 2009; Lim, 2002). Note that Sarawak and Sabah (two states of Malaysia in the Borneo Island) are the second resource area (after Indonesia) of the world for EBN production (Hobbs, 2004; Lim, 2002). In general, the EBN production cycle consists of five processes, i.e., swiftlets farming (P1), harvesting (P2), EBN cleaning (P3), EBN drying and reshaping (P4), and storing and packaging (P5) (Jong et al., 2013), as summarized in Fig. 3. The failure modes associated with each sub-process need to be analyzed. In addition, the failure modes for a number of tools and/or facilities used for maintenance in the first four sub-processes (i.e., M1, M2, M3 and M4 in Fig. 3) also need to be examined.

There are a total of 46 failure modes, denoted as $\overline{x_k} = [s_k, o_k, d_k]$, where $k = 1, 2, \ldots, 46$ (i.e., $m = 46$), for the entire EBN production cycle, as summarised in Table 1. Column "Processes" indicates the EBN production processes, i.e., from P1 to P5 and M1 to M4. Column "Failure modes" indicates the label of each failure mode (i.e., k) and its description. Columns s_k, o_k, and d_k indicate the three risk factors, i.e., S, O, and D.

The scale tables of S, O, and D are presented in Tables 2–4, respectively. In each scale table, column "Ranking" states the score intervals. These intervals are tagged with a linguistic term, as in column "Linguistic Term". A detailed description of each interval is summarized in column "Description". As an example, an S score from 1 to 2 is assigned with the linguistic term of "Very Low". This interval is used to indicate a failure with a minor effect, which can be ignored. Besides that, even if the failure occurs, the yield and the product quality are still excellent.

In this paper, the failure modes are analyzed using the proposed method in Section 3. To illustrate the usefulness of the Euclidean similarity metric in measuring the similarity level of two failure modes, the failure modes from two sub-processes, i.e., $\overline{x_{14}}$ to $\overline{x_{26}}$ in P3 and P4 (highlighted in Table 1), is exemplified in Section 4.1. Besides that, all failure modes in Table 1 are clustered and visualized with ETree as presented in Section 4.2. The parameters are empirically set as follows: $\theta_{childnode} = 2, epoch = 5$, and $\theta_{splitting} = 10, 20$, and 30. The risk of each cluster of failure modes is obtained based on its risk interval (Eqs. (4) and (5)) using the traditional RPN model. The clustering outcomes are then ordered and ranked, as discussed in Section 4.3.

4.1. Similarity measures of two failure modes

Consider the comparison between $\overline{x_3}$ and $\overline{x_{39}}$, where $\overline{x_3}$ is [10 1 1] and $\overline{x_{39}}$ is [1 10 1], as shaded in Table 1. Using Eq. (1), the *RPN* scores of both $\overline{x_3}$ and $\overline{x_{39}}$ are 10; however, they have different combinations of S, O, and D. Both $\overline{x_3}$ and $\overline{x_{39}}$ share the same D score, i.e., 1 (excellent detection). However, $\overline{x_3}$ is associated with S= 10 (i.e., potentially lead to product safety and quality) and O = 1 (i.e., happen at least once ever). $\overline{x_{39}}$ is associated with S = 1 (i.e., effects of the potential failure mode is not obvious and can be ignored) and O=10 (i.e., happen many times within an hour). With the Euclidean distance-based similarity metric (i.e., Eq. (2)), their similarity measure is 0.1835. This implies that even though both failure modes are associated with the same *RPN* score, they can be differentiated by the similarity metric. Such metric is useful for clustering and visualization of failure modes. As such, it is useful for solving a key issue of risk ranking in FMEA, i.e., different combinations of S, O, and D produce the same *RPN* scores.

4.2. Visualization with ETree

In this section, visualization with ETree as a result of the Euclidean distance-based metric is presented. To simplify the diagram, the nodes, i.e., $N_{l,p}$, are indexed as *l*. Fig. 4 depicts the tree structure with $\theta_{splitting} = 30$, while Figs. 5 and 6 depict the tree structures with $\theta_{splitting} = 20$, and 10, respectively. As shown in Fig. 4, the failure modes are clustered into thirteen nodes. The nodes located at the lower layers describe the detailed information in FMEA. As an example, the weight vector of $N_{12,4}$([2.0209 5.8461 1.0002]) is associated with three unique failure modes, i.e., [1 10 1], [2 6 1], and [2 5 1]. $N_{13,4}$, which is located next to $N_{12,4}$ in the tree, has the weight vector of

Table 1
Failure modes of the EBN production process.

Processes	Failure modes		s_k	o_k	d_k
	k	Description			
Swiftlets farming (P1)	1	Farm temperature goes beyond the acceptable range	5	1	2
	2	Air humidity goes beyond the acceptable range	4	1	2
	3	The density for nitrite gas and nitride gas is too high	10	1	1
	4	The farm is too bright	1	1	1
	5	Thieves steal the bird's nest from the farm.	9	1	1
	6	Existence of owls in the farm	5	1	2
	7	Existence of Asian glossy starling in the farm	6	1	6
	8	Existence of bats in the farm	5	1	1
	9	Existence of home lizards in the farm	2	1	6
	10	Existence of rats in the farm.	3	1	2
	11	Existence of cockroaches in the farm.	2	1	7
	12	Existence of ants in the farm.	3	1	1
Harvesting (P2)	13	Tearing of raw EBN	3	1	1
EBN cleaning (P3)	14	Tearing of raw EBN	4	9	1
	15	Dissolution of EBN	4	6	2
	16	Dirty EBN	3	10	1
	17	Tearing of raw EBN	4	10	1
EBN drying and reshaping (P4)	18	EBN is not dry enough for the reshaping process	3	7	4
	19	Spraying is uneven	4	7	4
	20	Cracking of the EBN	4	8	4
	21	Cracking of the EBN Too much gaps	3	8	4
	22	Failure in molding	3	8	4
	23	EBN is too dry	3	7	4
	24	Spraying is uneven	4	7	4
	25	Cracking of the EBN	4	7	4
	26	EBN is too dry and may crack	3	7	4
Storing and packaging (P5)	27	Cracking of the EBN	3	1	1
	28	Cracking of the EBN	3	7	1
Swiftlets farming (M1)	29	Malfunction of the alarm system	9	1	2
	30	Malfunction of the lighting system	5	1	4
	31	Failure of power supply	9	1	10
	32	Malfunction of sound system	5	1	4
	33	Malfunction of the humidifier	6	1	4
Harvesting (M2)	34	Abrasion of the taping knife	2	1	1
	35	Cracking of the swiftlets corner mirror	1	1	1
EBN cleaning (M3)	36	Pincers is blunt	4	6	1
	37	The workstation platform is dirty	2	6	3
	38	Unstable water pressure	2	2	3
	39	Too much contaminant in the water	1	10	1
	40	Dirty sifter	2	6	1
	41	The magnifier is dirty	1	2	1
EBN drying and reshaping (M4)	42	Dirty casing	2	5	1
	43	Malfunctioning of the bulb	1	5	3
	44	Dirty fan	2	4	3
	45	Dirty net	2	6	1
	46	Dirty mold	2	6	1

Table 2
Scale table for severity (from Jong et al. (2013)).

Ranking	Linguistic term	Description
1–2	Very low	• Effect of the potential failure mode is not obvious and can be ignored • Excellent yield and product quality
3–4	Low	• Very minor impact to the production yield • Failures cause a minor impact to EBN food production process control. The consequence will cause a minor effect to the products' cosmetic appearance and packaging
5–7	Medium	• Failures lead to the issue of minor security breaches of the farm, habitat of the swiftlets is affected by some of the pests and enemies of the swiftlets. The consequence will cause a reduction in the population of the swiftlets and the yield of the farm • Failures cause a minor impact to the production yield
8–9	High	• Failures lead to the issue of serious security breaches of the farm. Safety of the swiftlets will be threatened by its enemies, such as thieves and predators • Failures cause a major impact to the production yield
10	Very High	• Failures lead to impacts to product safety and quality • Compliance to law • Major impact to the reputation of the company and the products • Lead to failure to yield management

[3.7196 6.5797 1.0203], and is associated with six unique failure modes. i.e., [4 9 1], [4 6 2], [3 10 1], [4 10 1], [3 7 1], and [4 6 1]. Referring to Tables 2–4, node $N_{12,4}$ compresses the failure modes pertaining to the S score from 1 to 2 (i.e., effects of the potential failure mode is not obvious and can be ignored, with excellent yield and product quality), the O score from 5 to 10 (i.e., occurs more than at least once within 1 to 6 months.), and the D score of 1 (i.e., excellent detection). Node $N_{13,4}$ compresses the failure modes with the S score from 3 to 4 (i.e., have very minor impact to the production yield and cause minor effect to the products' cosmetic appearance and packaging), the O score from 6 to 10 (i.e., occurs more than at least once within 1 to 6 months.), and the D score of 1 (i.e., excellent detection). $N_{12,4}$ and $N_{13,4}$ are the child nodes of $N_{4,2}$, while the weight vector of $N_{4,2}$ is the same as that of $N_{13,4}$. As such, it is clear that the parent nodes represent more general information pertaining to the failure modes while their child nodes represent more specific or detailed information about the failure modes in FMEA.

The results can also be displayed in a three-dimensional (S, O, and D) space in accordance with Definition 1, as shown in Fig. 7. To allow a clear visualization, the failure modes are labelled with l. In Fig. 7, the failure modes mapped onto $N_{12,4}$ (labelled as 12) and $N_{13,4}$ (labelled as 13) comprise $\overline{x_k}, k = 39, 40, 42, 45, 46$, and $\overline{x_k}, k = 14, 15, 16, 17, 28, 36$, respectively. These failure modes have the closest distance measurement in the Cartesian space. Such method allows the failure modes with different combinations of S, O and D but the same *RPN* score to be clustered and visualized.

Referring to Figs. 4 and 7, $\overline{x_3}$ (i.e., [10 1 1]) and $\overline{x_{39}}$ (i.e., [1 10 1]) have the same *RPN* score (i.e., 10), and they belong to $N_{9,6}$ and $N_{12,4}$, respectively. Using the traditional RPN model (Eq. (1)), both $\overline{x_3}$ and $\overline{x_{39}}$ have the same risk level; therefore they should have the same level of importance or priority. However, it should be noted that $\overline{x_3}$ is associated with S=10 (i.e., potentially lead to product safety and quality) and O = 1 (i.e., happen at least once ever) while $\overline{x_{39}}$ is associated with S = 1 (i.e., effects of the potential failure mode is not obvious and can be ignored) and O = 10 (i.e., happen many times within an hour). As such, both failure modes should have different prioritization actions in practice. By using ETree, these two failure modes (although having the same *RPN* score) are grouped into two different clusters. This outcome indicates that ETree is able to differentiate both failure modes so that FMEA users can prioritize different actions to overcome each failure mode

Table 3
Scale table for occurrence (from Jong et al. (2013)).

Ranking	Linguistic term	Description
1	Extremely Low	• Failures happen at least once ever
2–3	Very Low	• Failures happen at least once within 6 to 12 months
4–5	Low	• Failures happen at least once within 1 to 6 months
6–7	Medium	• Failures happen at least once within 1 to 30 days
8–9	High	• Failures happen at least once within 1 to 8 working hours (1 day)
10	Very High	• Failures happen many times within an hour

Table 4
Scale table for detect (from Jong et al. (2013)).

Ranking	Linguistic term	Description
1–3	Very High	Detection is excellent • Control actions can almost detect the failure on the spot and appropriate actions are taken to solve the failure and the weakness • Prevent the excursion from occurring
4–6	High	Detection is good • Control actions can almost detect the failure on the spot within the same process module or steps • In farm management, control actions can detect the failure within one day • Appropriate actions are available to solve the failure and the weakness
7–8	Medium	Detection is acceptable • Control actions can detect the failure within one to two process modules or steps • In farm management, control actions can detect the failure within one to three days • Appropriate actions are available. However the failure can be tricky and hard to solve
9	Low	Hard to detect • Control actions may not detect the failure • Appropriate actions may not be available and the failure cannot be solved
10	Very Low	Detection is almost impossible • No control action is available • No solution is available for solving the failure

separately. The difference between $\overline{x_3}$ and $\overline{x_{39}}$ is also indicated by their Euclidean similarity metric (0.1835), whereby they are located far apart in the input (S, O, and D) space, as shown in Fig. 7. Besides that, $\overline{x_3}$ belongs to the same group of $\overline{x_1}, \overline{x_2}, \overline{x_5}, \overline{x_6 x_7}, \overline{x_8}$, $\overline{x_9}, \overline{x_{11}}, \overline{x_{29}}, \overline{x_{30}}, \overline{x_{32}}$, and $\overline{x_{33}}$, with relatively high similarity measures, i.e., ranging from 0.3585 to 0.9358. On the other hand, $\overline{x_{39}}$ belongs to the same group of $\overline{x_{40}}, \overline{x_{42}}, \overline{x_{45}}$, and $\overline{x_{46}}$, with high similarity measures too, i.e., ranging from 0.6729 to 0.7355. This observation highlights the usefulness of ETree coupled with the Euclidean similarity metric to differentiate failure modes into distinctive clusters that can be visualized easily by FMEA users for decision making, e.g. allowing different prioritization actions to be taken.

As expected, different $\theta_{splitting}$ settings lead to different tree structures, as depicted in Figs. 4–6. A smaller $\theta_{splitting}$ results in a larger and more detailed tree structure. Table 5 shows the overall comparison among different tree structures with different $\theta_{splitting}$ settings, i.e., in terms of the tree size, tree depth, and number of leaf nodes. A discussion with FMEA users engaged in this case study indicated that $\theta_{splitting} = 30$ provided the best clustering and visualization outcome, whereby all forty-six failure modes could be visualized easily as seven clusters for decision making purposes.

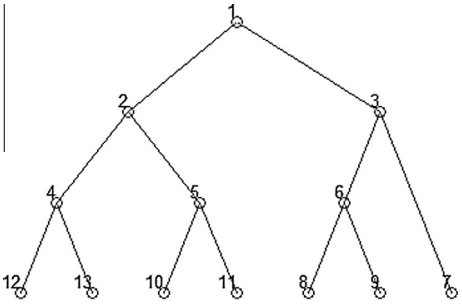

Fig. 4. The tree model at $\theta_{splitting} = 30$.

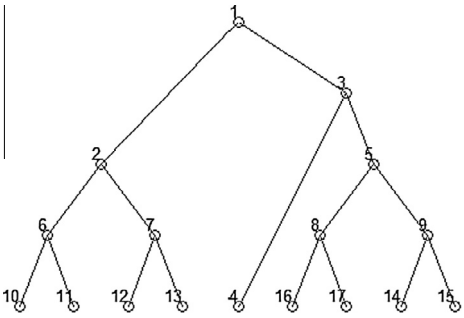

Fig. 5. The tree model with $\theta_{splitting} = 20$.

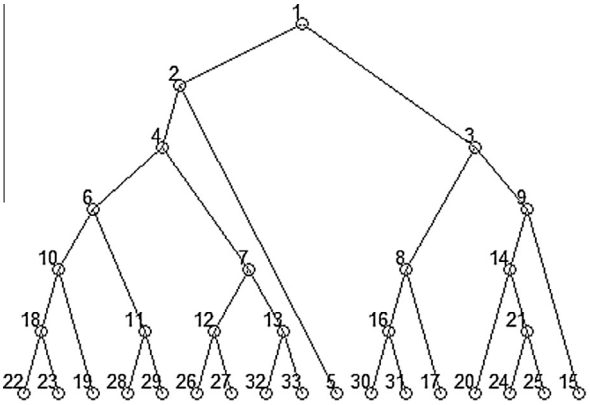

Fig. 6. The tree model with $\theta_{splitting} = 10$.

4.3. Risk interval and ordering analysis

The ETree model enables the risks of failure modes to be ordered and analyzed. Specifically, the clusters of failure modes are ordered, and the risk intervals are analyzed. Table 6 summarizes the risk intervals of failure modes with $\theta_{splitting} = 30$. Column l indicates the cluster label. Columns $w_{l,S}, w_{l,O}$, and $w_{l,D}$ indicate the elements of the weight vector, i.e., $[w_{l,S} w_{l,O} w_{l,D}]$. The risk interval of each cluster is summarized in columns $\underline{RPN_l}$ and $\overline{RPN_l}$. As an example, in Table 6, $N_{7,3}$ with the weight vector of [8.778405 1.000083 7.368034] is associated with an risk interval of [90 90].

From Table 6, notice that $\left[\underline{RPN_{12}}, \overline{RPN_{12}}\right] = [10, 12]$ and $\left[\underline{RPN_{13}}, \overline{RPN_{13}}\right] = [21, 48]$. As such, the risk associated with $N_{12,4}$ is within the risk interval of $N_{13,4}$. Besides that, $\left[\underline{RPN_7}, \overline{RPN_7}\right] = [90, 90]$, and $\left[\underline{RPN_8}, \overline{RPN_8}\right] = [1, 24]$. This means that the risk

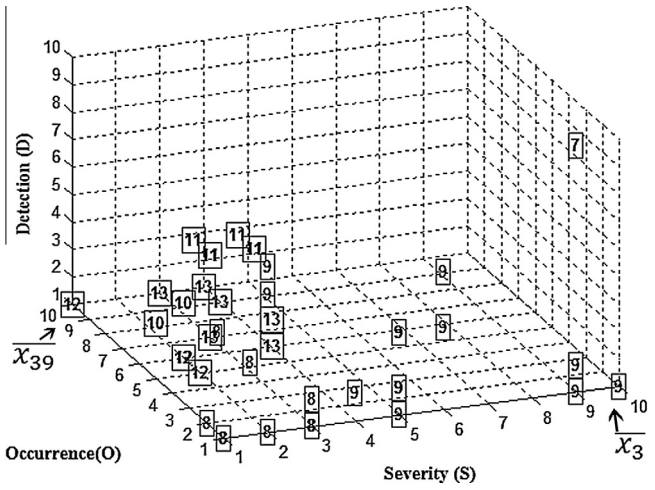

Fig. 7. Clustering outcomes of identical failure modes in the S, O, and D space with $\theta_{splitting} = 30$.

Table 5
Comparison of the tree size, tree depth and number of clusters with different $\theta_{splitting}$ settings.

$\theta_{splitting}$	Tree size	Tree depth	Number of clusters (leaf nodes)
30	13	4	7
20	17	5	9
10	33	7	17

associated with $N_{7,3}/N_{11,5}$ is higher than that of $N_{8,6}$, i.e., $N_{8,6} < N_{7,3}/N_{11,5}$. It can also be observed that $N_{8,6} < N_{9,6} < N_{13,4} < N_{7,3}/N_{11,5}$. The results in Table 6 suggest that the failure modes in $N_{7,3}/N_{11,5}$ have higher risks than those in $N_{13,4}$, and then follow by those in $N_{9,6}$ and $N_{8,6}$.

It is evident that the ETree model allows failure modes to be ordered and interpreted as a group, instead of interpreting each failure mode individually. This is useful for FMEA users to have a relatively simple ordering outcome, which can be visualized quickly. Instead of ordering all forty-six failure modes, only seven clusters of failure modes need to be ordered after clustering. The results show that the priority should be given to $N_{11,5}$ and $N_{7,3}$, which comprise failure modes of $\overline{x_{18}}, \overline{x_{19}}, \overline{x_{20}}, \overline{x_{21}}, \overline{x_{22}}, \overline{x_{23}}, \overline{x_{24}}, \overline{x_{25}}, \overline{x_{26}}$, and $\overline{x_{31}}$, with the risk intervals of $[84, 128]$ and $[90, 90]$, respectively. From Fig. 7, failure modes in $N_{11,5}$ and $N_{7,3}$ fall in different regions of the input space, and can hardly be grouped together owing to different implications. However, it is important to identify both groups of failure modes because they have the highest risk intervals. In short, the ETree model can identify the failure modes with high risks with different implications, and categorise them into different groups for decision making.

4.4. Computational times

In this paper, a laptop equipped with Intel (R) Core (TM) i7–3612QM @3.10 GHz and 4.0 GB of RAM was used for the experimental study. The computational times required for $\theta_{splitting} = 10, 20$, and 30 were 0.9933, 0.2744 and 0.1531 s, respectively, i.e., less than one second for each $\theta_{splitting}$ setting. In short, the proposed ETree model could generate its tree structure with a low computational requirement.

4.5. Remarks

This empirical study demonstrates that the failure modes in FMEA can be easily visualized with ETree, which leads to the following advantages. Firstly, the failure modes captured in a complex FMEA worksheet can be represented in a tree structure, whereby the failure modes within the same node share a higher similarity measure than those from other nodes. From the feedback of the FMEA users engaged in this case study, the tree structures are useful for visualizing and analyzing the associated failure modes, which result in more effective and efficient prioritization actions and decision making. Secondly, based on the information and knowledge elucidated from the tree structures, understanding the entire FMEA problem with a large number of failure modes and a complex FMEA worksheet can be made feasible within a short period of time, as acknowledged by the FMEA users of this case study. These advantages directly mitigate the challenging issues related to practical implementation of FMEA, i.e., complex FMEA worksheets and intricacy of use as highlighted in Montgomery et al. (1996), Signor (2002). From the discussion with the FMEA users engaged in the case study, the failure modes captured in a tree structure can be understood quickly and effectively, as compared with the information in a complex FMEA worksheet.

In addition to the above two advantages, the tree structure provides a hierarchy of layers that represents the failure modes with increasing details when the FMEA users transverse down the layers in the tree. This property allows the FMEA users to choose an alternative solution pertaining to the nodes at a higher layer, instead of only at the leaf nodes. Choosing clustering outcomes at a higher layer provides a set of more general prioritization actions as compared with that from a lower layer. Conversely, choosing clustering outcomes at a lower layer provides a set of detailed prioritization actions. Such property is useful as it provides options for the FMEA users to plan their prioritization actions subject to the available resources.

5. Summary

In this paper, an ETree model has been proposed to solve two well-known shortcomings in FMEA, i.e., complicated FMEA worksheet and intricacy of use. The proposed model has been evaluated comprehensively using a case study with real EBN information. The ETree has shown its usefulness for clustering the S, O, and D scores, and analyzing the failure modes in groups. The characteristics and relationship among the failure modes or prioritization actions can be visualized in a tree structure. The failure modes in groups can also be easily analyzed and ordered by using the proposed risk interval and risk ordering equations.

Table 6
Risk interval analysis of failure modes with $\theta_{splitting} = 30$.

l	$w_{l,S}$	$w_{l,O}$	$w_{l,D}$	$\underline{RPN_l}$	$\overline{RPN_l}$	Failure modes
7	8.778405	1.000083	7.368034	90	90	$\overline{x_{31}}$
8	1.826295	3.549084	2.637303	1	24	$\overline{x_4}, \overline{x_{10}}, \overline{x_{12}}, \overline{x_{13}}, \overline{x_{27}}, \overline{x_{34}}, \overline{x_{35}}, \overline{x_{38}}, \overline{x_{41}}$, and $\overline{x_{44}}$
9	5.505737	1.034637	3.918611	5	36	$\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_5}, \overline{x_6}, \overline{x_7}, \overline{x_8}, \overline{x_9}, \overline{x_{11}}, \overline{x_{29}}, \overline{x_{30}}, \overline{x_{32}}$, and $\overline{x_{33}}$
10	1.305596	5.318108	3.071625	15	36	$\overline{x_{37}}$, and $\overline{x_{43}}$
11	3.817895	7.01255	3.981091	84	128	$\overline{x_{18}}, \overline{x_{19}}, \overline{x_{20}}, \overline{x_{21}}, \overline{x_{22}}, \overline{x_{23}}, \overline{x_{24}}, \overline{x_{25}}$, and $\overline{x_{26}}$
12	2.020925	5.84615	1.000247	10	12	$\overline{x_{39}}, \overline{x_{40}}, \overline{x_{42}}, \overline{x_{45}}$, and $\overline{x_{46}}$
13	3.719644	6.579667	1.020294	21	48	$\overline{x_{14}}, \overline{x_{15}}, \overline{x_{16}}, \overline{x_{17}}, \overline{x_{28}}$, and $\overline{x_{36}}$

There are two main contributions of this study. From the perspective of intelligent systems, this study contributes toward a new application of the ETree model, i.e., FMEA. To the best of our knowledge, this is the first attempt to use the ETree model in FMEA applications with real information in the literature. From the perspective of FMEA, this study contributes toward a new computing technique for tackling two existing shortcomings of FMEA. To complement the ETree model in FMEA, the ideas of risk interval and risk ordering for analyzing the failure modes in groups have also been introduced. These notions are useful for clustering and visualization problems, as they allow failure modes to be grouped and ordered. Specifically, we have demonstrated a practical example on the use of risk interval and risk ordering, coupled with ETree, to allow both visualization and ordering of failure modes in groups in the EBN case study.

The proposed ETree model suffers from a number of limitations too. Firstly, it is difficult to have an accurate visualization of the failure modes without the consensus from the FMEA users with different expertise levels. Besides that, the ETree model requires a few new parameters to be pre-determined. As such, the trial-and-error method is required to obtain the preferred tree structure.

For further research, other visualization modalities, e.g., graphical effects (size, color, shape), text labels, and hierarchical maps, can be examined for visualizing the failure modes in FMEA effectively. The use of other clustering models, e.g., the SOM network (Kohonen, 2001) and self-organizing tree (Campos & Carpenter, 2001), for visualization of the failure modes or prioritization actions can be investigated too. Potential applications of other types of neural network learning paradigms, e.g., supervised, reinforcement, and/or semi-supervised learning, for FMEA applications can be studied. Visualization of the failure modes with fuzzy S, O, and D (also known as linguistic S, O, and D) measurements constitutes another area for further research.

Acknowledgements

The financial support of the FRGS grant (FRGS/ICT02(01)/997/2013(38)) and the RACE grants (RACE/F2/TK/UNIMAS/5, and RACE/(c)1/1252/2015(8)), are gratefully acknowledged.

References

Arunajadai, S. G., Uder, S. J., Stone, R. B., & Tumer, I. Y. (2004). Failure mode identification through clustering analysis. *Quality and Reliability Engineering International, 20*(5), 511–526.

Bowles, J. B., & Peláez, C. E. (1995). Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. *Reliability Engineering and System Safety, 50*, 203–213.

Bozdag, E., Asan, U., Soyer, A., & Serdarasan, S. (2015). Risk prioritization in failure mode and effects analysis using interval type-2 fuzzy sets. *Expert Systems with Applications, 42*(8), 4000–4015.

Campos, M. M., & Carpenter, G. A. (2001). S-TREE: self-organizing trees for data clustering and online vector quantization. *Neural Networks, 14*, 505–525.

Chang, K. H. (2014). A more general risk assessment methodology using a soft set-based ranking technique. *Soft Computing, 18*(1), 169–183.

Chang, W. C., Luo, J., & Parker, K. J. (1998). Image segmentation via adaptive K-Mean clustering and knowledge-based morphological operations with biomedical applications. *IEEE Transaction on Image Processing, 7*(12), 1673–1683.

Chang, W. L., Tay, K. M., & Lim, C. P. (2013). Enhancing an evolving tree-based text document visualization model with Fuzzy c-means clustering. *IEEE International Joint Conference on Fuzzy System*, 1–6.

Chang, W. L., Tay, K. M., & Lim, C. P. (2014). An evolving tree for text document clustering and visualization. *Soft computing in industrial applications*, 141–151.

Du, Y., Mo, H., Deng, X., Sadiq, R., & Deng, Y. (2014). A new method in failure mode and effects analysis based on evidential reasoning. *International Journal of System Assurance Engineering and Management, 5*(1), 1–10.

Feather, M. S., Cornford, S. L., Kiper, J. D., & Menzies, T. (2006). Experiences using visualization technique to present requirements, risk to them, and options for risk mitigation. In *REV '06, first international workshop on requirement engineering visualization* (p. 10).

Garrick, B. J. (1988). The approach to risk analysis in three industries: nuclear power, space systems, and chemical process. *Reliability Engineering & System Safety, 23*, 195–205.

Grøndahl, I. H., Lund, M. S., & Stølen, K. (2011). Reducing the effort to comprehend risk models: Text labels are often preferred over graphical means. *Risk Analysis, 31*(11), 1813–1831.

Guimarães, A. C. F., & Lapa, C. M. F. (2004). Fuzzy FMEA applied to PWR chemical and volume control system. *Progress in Nuclear Energy, 44*, 191–213.

Hobbs, J. J. (2004). Problems in the harvest of edible birds' nests in Sarawak and Sabah, Malaysian Borneo. *Biodiversity & Conservation, 13*(12), 2209–2226.

Huang, S. Y., & Tsaih, R. H. (2012). The prediction approach with growing hierarchical self-organizing map. In *International joint conference on neural network*, (pp. 1–7).

Inoue, H., & Yamada, S. (2010). Failure mode and effects analysis in pharmaceutical research. *International Journal of Quality and Service Sciences., 2*(3), 369–382.

Jee, T. L., Tay, K. M., & Lim, C. P. (2015). A new two-stage fuzzy inference system-based approach to prioritize failures in failure mode and effect analysis. *IEEE Transaction on Reliability*. http://dx.doi.org/10.1109/TR.2015.2420300.

Jong, C. H., Tay, K. M., & Lim, C. P. (2013). Application of the fuzzy failure mode and effect analysis methodology to edible bird nest processing. *Computer and Electronics in Agriculture, 96*, 90–108.

Jong, C. H., Tay, K. M., & Lim, C. P. (2014). A single input rule modules connected fuzzy FMEA methodology for edible bird nest processing. In *Soft computing in industrial applications* (pp. 165–176). Springer International Publishing.

Jordan, D. (2009). Globalization and bird's nest soup. *International Development Planning Review., 26*, 97–110.

Keskin, G. A., & Özkan, C. (2009). An alternative evaluation of FMEA: Fuzzy art algorithm. *Quality and Reliability Engineering International, 25*, 647–661.

Kohonen, T. (2001). *Self-organizing maps* (Third ed.). Berlin: Springer.

Kohonen, T., Simula, O., & Visa, A. (1996). Engineering applications of the self-organizing map. *Proceeding of the IEEE, 84*(10), 1358–1384.

Korayem, M. H., & Iravani, A. (2008). Improvement of 3P and 6R mechanical robots reliability and quality applying FMEA and QFD approaches. *Robotic and Computer-Integrated Manufacturing, 24*, 472–487.

Kuo, R. J., Wang, C. F., & Chen, Z. Y. (2012). Integration of growing self-organizing map and continuous genetic algorithm for grading lithium-ion battery cells. *Applied Soft Computing, 12*, 2012–2022.

Lan, H. W., Frank, E., & Hall, M. A. (2005). *Data mining: Practical machine learning tools and techniques* (3rd ed.). Morgan Kaufmann.

Li, C., Chen, K., & Xiang, X. (2015). An integrated framework for effective safety management evaluation: Application of an improved grey clustering measurement. *Expert Systems with Applications, 42*(13), 5541–5553.

Lim, C. K. (2002). *Earl of cranbrook. Swiflets of borneo: Builder of edible nest*. Borneo: Natural History Publications.

Liu, H. C., Chen, Y. Z., You, J. X., & Li, H. (2014h). Risk evaluation in failure mode and effects analysis using fuzzy digraph and matrix approach. *Journal of Intelligent Manufacturing*. http://dx.doi.org/10.1007/s10845-014-0915-6.

Liu, H. C., Fan, X. J., Li, P., & Chen, Y. Z. (2014e). Evaluating the risk of failure modes with extended MULTIMOORA method under fuzzy environment. *Engineering Applications of Artificial Intelligence, 34*, 168–177.

Liu, H. C., Liu, L., & Li, P. (2014a). Failure mode and effects analysis using intuitionistic fuzzy hybrid weighted Euclidean distance operator. *International Journal of Systems Science, 45*(10), 2012–2030.

Liu, H. C., Liu, L., & Lin, Q. L. (2013b). Fuzzy failure mode and effects analysis using fuzzy evidential reasoning and belief rule-based methodology. *IEEE Transactions on Reliability, 62*(1), 23–36.

Liu, H. C., Liu, L., & Liu, N. (2013a). Risk evaluation approaches in failure mode and effects analysis: A literature review. *Expert Systems with Applications, 40*, 828–838.

Liu, H. C., Li, P., You, J. X., & Chen, Y. Z. (2014d). A novel approach for FMEA: Combination of interval 2-tuple linguistic variables and gray relational analysis. *Quality and Reliability Engineering International*. http://dx.doi.org/10.1002/qre.1633.

Liu, H. C., You, J. X., Fan, X. J., & Lin, Q. L. (2014b). Failure mode and effects analysis using D numbers and grey relational projection method. *Expert Systems with Applications, 41*(10), 4670–4679.

Liu, H. C., You, J. X., Lin, Q. L., & Li, H. (2014c). Risk assessment in system FMEA combining fuzzy weighted average with fuzzy decision-making trial and evaluation laboratory. *International Journal of Computer Integrated Manufacturing*. http://dx.doi.org/10.1080/0951192X.2014.900865.

Liu, H. C., You, J. X., Shan, M. M., & Shao, L. N. (2014f). Failure mode and effects analysis using intuitionistic fuzzy hybrid TOPSIS approach. *Soft Computing*, 1–14.

Liu, H. C., You, J. X., & You, X. Y. (2014g). Evaluating the risk of healthcare failure modes using interval 2-tuple hybrid weighted distance measure. *Computers & Industrial Engineering, 78*, 249–258.

Mandal, S., & Maiti, J. (2014). Risk analysis using FMEA: Fuzzy similarity value and possibility theory based approach. *Expert Systems with Applications, 41*(7), 3527–3537.

Matharage, S., Alahakoon, D., Rajapakse, J., & Pin, H. (2011). Fast growing self-organizing map for text clustering. *Neural Information Processing, 7063*, 406–415.

McNally, K. M., Page, M. A., & Sunderland, V. B. (1997). Failure-mode and effects analysis in improving a drug distribution system. *American Journal of Health-System Pharmacy, 54*, 171–177.

Montgomery, T. A., Pugh, D. R., Leedham, S. T., & Twitchett, S. R. (1996). FMEA automation for the complete design process. In *Proceedings annual reliability & maintainability symposium*, (pp. 30–36).

Pakkanen, J., Iivarinen, J., & Oja, E. (2004). The evolving tree-a novel self-organizing network for data analysis. *Neural Processing Letter, 20*, 199–211.

Pakkanen, J., Iivarinen, J., & Oja, E. (2006). The evolving tree – analysis and applications. *IEEE Transaction on Neural Network, 17*(3), 591–603.

Rezaee, M. R., Leliveldt, B. P. F., & Reiber, J. H. C. (1998). A new cluster validity index for the fuzzy c-mean. *Pattern Recognition Letters, 19*, 237–246.

Romuald Iwańkowicz, R., & Rosochacki, W. (2014). Clustering risk assessment method for shipbuilding industry. *Industrial Management & Data Systems, 114*(9), 1499–1518.

Rui, X., & Donald, C. W. (2009). *Clustering. IEEE series on computational intelligence*. Hoboken: Wiley.

Signor, M. C. (2002). The failure-analysis matrix: A Kinder, gentler alternative to FMEA for information systems. In *Proceeding of reliability and maintainability symposium 2002*, (pp. 173–177).

Stamatis, D. H. (2003). *Failure mode and effect analysis: FMEA from theory to execution* (Second ed.). ASQ Press.

Tay, K. M., Jong, C. H., & Lim, C. P. (2015). A Clustering-based failure mode and effect analysis with an application to the edible bird nest industry. *Neural Computing and Applications, 26*(3), 551–560.

Tay, K. M., & Lim, C. P. (2006). Fuzzy FMEA with a guided rules reduction system for prioritization of failures. *International Journal of Quality & Reliability Management, 23*, 1047–1066.

Tay, K. M., & Lim, C. P. (2008a). On the use of fuzzy inference techniques in assessment models: Part I - theoretical properties. *Fuzzy Optimization and Decision Making, 7*, 269–281.

Tay, K. M., & Lim, C. P. (2008b). On the use of fuzzy inference techniques in assessment models: Part II: Industrial applications. *Fuzzy Optimization and Decision Making, 7*, 283–302.

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transaction on Neural Network, 11*(3), 586–600.

Wintle, B. C., & Nicholson, A. (2014). Exploring risk judgments in a trade dispute using bayesian networks. *Risk Analysis, 34*(6), 1095–1111.

Zafiropoulos, E. P., & Dialynas, E. N. (2005). Reliability prediction and failure mode effects and criticality analysis (FMECA) of electronic devices using fuzzy logic. *International Journal of Quality & Reliability Management, 22*, 183–200.