Accident Analysis and Prevention 41 (2009) 1180–1191

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Identification and classification of dynamic event tree scenarios via possibilistic clustering: Application to a steam generator tube rupture event

D. Mercurio[a,*], L. Podofillini[a], E. Zio[b], V.N. Dang[a]

[a] Paul Scherrer Institute, Villigen PSI, Switzerland
[b] Polytechnic of Milan, Italy

ARTICLE INFO

Article history:
Received 7 September 2007
Received in revised form 6 August 2008
Accepted 8 August 2008

Keywords:
Dynamic event tree
Probabilistic safety assessment
Dynamic safety assessment
Fuzzy
Possibilistic
Clustering

ABSTRACT

This paper illustrates a method to identify and classify scenarios generated in a dynamic event tree (DET) analysis. Identification and classification are carried out by means of an evolutionary possibilistic fuzzy C-means clustering algorithm which takes into account not only the final system states but also the timing of the events and the process evolution. An application is considered with regards to the scenarios generated following a steam generator tube rupture in a nuclear power plant. The scenarios are generated by the accident dynamic simulator (ADS), coupled to a RELAP code that simulates the thermo-hydraulic behavior of the plant and to an operators' crew model, which simulates their cognitive and procedures-guided responses.

A set of 60 scenarios has been generated by the ADS DET tool. The classification approach has grouped the 60 scenarios into 4 classes of dominant scenarios, one of which was not anticipated a priori but was "discovered" by the classifier. The proposed approach may be considered as a first effort towards the application of identification and classification approaches to scenarios post-processing for real-scale dynamic safety assessments.

1. Introduction

In the probabilistic safety assessment (PSA) of nuclear power plants (NPPs), accident scenarios, which are dynamic in nature, are analyzed with event trees and fault trees. This framework has proven useful for identifying unforeseen vulnerabilities leading to modifications of the plant, its operating procedures and the training of the personnel.

Yet, the current PSA framework has some limitations in handling the timing of automatic and personnel actions, whose variability may influence the successive evolution of the scenarios, and in modeling the interactions between the physical evolution of the process variables (temperatures, pressures, mass flows, etc., ...) and the behavior of the hardware components and the operating crew. Thus, differences in the sequential order of the same success and failure events and the timing of event occurrence along an accident scenario may affect its evolution and outcome; also, the evolution of the process variables (temperatures, pressures, mass flows, etc., ...) may affect the event occurrence probabilities and thus the developing scenario (Siu, 1994).

To overcome the above-mentioned limitations, dynamic methodologies have been investigated which attempt to capture the integrated response of the systems/components/operating crew during an accident scenario. Models of the process and human operator dynamics are embedded within stochastic simulation engines which generate the sequence of components failure and success transitions along the scenarios. A review of the methods developed can be found in Siu (1994); applications and theoretical advances in the subject can be found in Labeau et al. (2000), Sheng and Mosleh (1996), Cojazzi (1996), Dang (1996), Hofer et al. (2002), Kloos and Peschke (2007), Marseguerra and Zio (1996), Labeau (1996), Marseguerra and Zio (2002), Kopustinkas et al. (2005), Podofillini et al. (submitted for publication), Hu and Modarres (1999).

Abbreviations: ADS, accident dynamic simulator; DET, dynamic event tree; EFW, emergency feed water; EOP, emergency operating procedures; FCM, fuzzy C-means; HPI, high pressure injection; IDAC, information, decision, and action in crew context; IE, initiating event; MFW, main feed water; NPP, nuclear power plant; PSA, probabilistic safety assessment; PWR, pressurizer water reactor; PZR, pressurizer; RCP, reactor coolant pump; RCS, reactor coolant system; SCM, subcooling margin; SG, steam generator; SGTR, steam generator tube rupture; SL, steam line; TBV, turbine bypass valve; TBVA, turbine bypass valve of a loop; UOD, universe of discourse.

* Corresponding author.
E-mail addresses: davide.mercurio@psi.ch (D. Mercurio), luca.podofillini@psi.ch (L. Podofillini), enrico.zio@polimi.it (E. Zio), vinh.dang@psi.ch (V.N. Dang).

0001-4575/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2008.08.013

In particular, this paper focuses on dynamic event trees (DETs). The most evident difference between DETs and the event trees (ETs) is as follows. ETs, which are typically used in the industrial PSA, are constructed by an analyst, and their branches are based on success/failure criteria set by the analyst. These criteria are based on simulations of the plant dynamics. Instead, DETs are produced by a software that embeds the models that simulates the plant dynamics into stochastic models of components failure/success and of the crew response. More details on DETs can be found in Section 2.1. The differences between DETs and ETs are discussed in details in Siu (1994).

A challenge arising from the dynamic approach to PSA is that the number of scenarios to be analyzed is much larger than that of the classical fault/event tree approaches, so that the a posteriori information retrieval can become quite burdensome and difficult (Labeau et al., 2000). In the first attempt to overcome this difficulty, the authors of the present paper have proposed an approach to identifying and grouping the scenarios of a dynamic safety assessment, with the aim of finding the principal patterns of system evolutions towards failure (Podofillini et al., submitted for publication). The scenarios are grouped combining information from the end state, events sequence and process variables evolutions. A possibilistic evolutionary fuzzy C-means (FCM) clustering algorithm is at the basis of the grouping approach.

In the present paper, the approach is applied on the scenarios of the DET of a steam generator tube rupture (SGTR) event in a nuclear power plant (NPP). The DET are generated by the accident dynamic simulator (ADS) software (Sheng and Mosleh, 1996). The physical plant model and the operating crew model of the system of interest are externally linked to the ADS. In particular, the crew response is treated by the IDAC model, a crew model including cognitive, emotional, and physical activities during accident scenarios within a procedure-guided response framework (Chang and Mosleh, 1999, 2000; Chang and Mosleh, 2006); the simulation of the plant process is performed by a RELAP thermal-hydraulic transient model (Information Systems Laboratories, 2001).

The paper is organized as follows. Section 2 introduces the DET framework and discusses the problem of identifying and grouping scenarios derived from a dynamic risk assessment. The approach proposed to tackle the problem is presented in Section 3. The possibilistic and evolutionary FCM clustering algorithms are presented in Section 4. Sections 5 and 6 present the case study of application. Conclusions and directions of future research close the paper.

2. Problem description

Section 2.1 briefly presents the use of DET in dynamic safety assessments, within a framework that explicitly couples the stochastic process of system/component/operator states transitions with the deterministic evolution of the physical process variables (Sheng and Mosleh, 1996; Cojazzi, 1996; Dang, 1996). Section 2.2 presents the motivation underlying the classification approach proposed in this work for identifying and grouping similar scenarios.

2.1. Dynamic event trees

Discrete dynamic event trees are produced by software coupling stochastic failure/success events (e.g. hardware failures and operators crew actions) with the continuous-time behavior of the plant process variables, usually captured by a simulator that solves the differential equations governing the plant physical evolution (Dang, 1996). Examples of these software are: DYLAM (Cojazzi, 1996), MCDET (Kloos and Peschke, 2007), and ADS (Sheng and Mosleh, 1996). The latter has been used in this paper.

The basic idea of the DET is to identify branching points (also called nodes) in the system evolution (i.e. time-points along the simulation at which stochastic events occur), save the state of the system at each branching point and successively pursue the simulation of all branches. Branching points are generated whenever a system, a component or an operator action is called for. Each branch represents a possible outcome of the stochastic event. Simulating the scenario evolution from all branching points allows exploring all possible behaviors of the system parameters and process variables.

The evolution of a dynamic event tree is shown in Fig. 1; each path represents a different branch of the scenario. The top of the figure shows the tree after the simulation of sequence 1, whose end state (e.g. stable shutdown or plant damage) is determined by the calculated plant process variables and systems states. In this run, three branching events (represented by undeveloped nodes) arise and the complete human–machine system state (state of plant parameters, equipment states, and operator state) is stored for each node.

Next, the DET software returns to the last node in the sequence (marked ‘a’ in Fig. 1), reloads the overall state and simulates the sequence 2 to the end state of the undeveloped path, given that the event represented by node ‘a’ has failed. Once again, the end state is determined by the calculated plant conditions. In this example, no additional events arise. At this stage, the software returns to the next undeveloped node (marked ‘b’ in Fig. 1), reloads the overall state stored for ‘b’ and simulates the resulting sequence 3 up to the end state. Along this sequence, two new events are considered (marked ‘d’ and ‘c’ in Fig. 1). These are events that arise only after the occurrence of the failure represented in node ‘b’ and thus are not relevant for sequences 1 or 2. With sequence 3 completed, the next node is ‘c’ and the process repeats, resulting in sequence 4. Subsequently, node ‘d’ is developed, then ‘e’, and so on until all nodes have been developed.

During the scenario generation process, the probability is automatically “split” when an event occurs, i.e. a branching point is generated. In particular, the probability is updated by the conditional probability of a new branch. The process continues until all the sequences in the DET are stopped according to a predefined truncation rule, based on cut-off probabilities (or frequencies, depending on the implementation, when the cut-off includes the initiating event) in order to avoid the combinatorial explosion.

In the present paper the accident dynamic simulator (ADS) software was used to generate the dynamic event trees (Sheng and

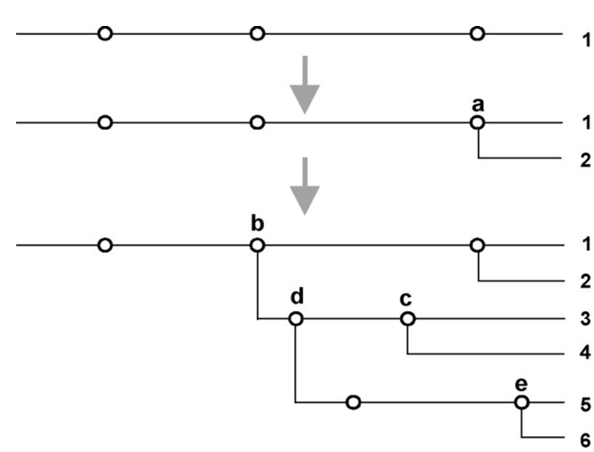

Fig. 1. Evolution of a dynamic event tree.

Mosleh, 1996). ADS is a computer code developed for accident dynamic simulation following an initiating event in nuclear power plants. ADS generates time-dependent scenarios by following the possibly dynamically changing states of various components and operators crew responses. The operator model implemented in the ADS software is the IDAC (information, decision, and action in crew context) operator model (Chang and Mosleh, 1999, 2000; Chang and Mosleh, 2006) combined with a model of the procedure-guided operators' response. In fact, in a real control room the preferred operator response is guided by the procedures; however, a response based on the operators' knowledge and training has been observed in some events and in simulated situations in which the procedures are perceived to be inadequate (De Carvalho and Paulo, 2006; Woods, 1984; Roth et al., 1994). In this context the IDAC models the cognitive response to a particular situation.

2.2. Motivation for the classification approach to scenario grouping

A potential difficulty of the dynamic approach to safety assessment is that the number of scenarios that arise in the analysis is much larger than that of the classical fault/event tree approaches and thus not only the computational burden of the simulation is increased but also the a posteriori information retrieval becomes difficult. On the other hand, the dynamic approach brings some clear advantages from the point of view of the completeness of the analysis and of the information content made available. First, there is potential for the identification of accident scenarios which may have been overlooked by the analyst when building the accident sequence models at the basis of the fault/event trees. Second, conservative simplifying assumptions made by the analyst, for example on the evolution of some process parameters, can be relaxed as the process evolution is simulated directly by the underlying dynamics model. Finally, additional informative content becomes available as a result of the dynamic analysis, in the form of time-dependent probability density functions of components states and process parameters values. In this respect, the amount of information retrievable from dynamic methodologies, in terms of number of scenarios and probability distributions, can be overwhelming and generally calls for a significant effort in the post-processing phase (Labeau et al., 2000).

Indeed, while the typical outcome of the dynamic safety analysis is the time evolution of the probability of the process variables exceeding predefined safety threshold, the focus is mainly on the system and process states at the end of the scenarios, with limited use of the information contained in the actual evolution towards these states. On the contrary, proper use of the information on the evolution of the scenarios can provide significant safety insights on the dominant scenarios with respect to the criticality and efficiency of the protections designed to counteract them.

The problem of identifying critical scenarios in dynamic reliability studies has been recently studied in Demmou et al. (2004), where the stochastic aspects of the system evolution are represented by means of Petri nets. The method proposed in Demmou et al. (2004) is based on the identification of the transitions through which a final state of interest can be reached. It is a qualitative method: the scenario probability does not enter in the search scheme. In addition, the problem of grouping similar sequences with respect to the process variables evolution is not addressed.

Within a DET simulation framework for dynamic safety analysis, the information on the evolution of the system is hidden in the simulated branches. Among these branches, there are sequences that reproduce qualitatively similar behaviors in terms of the evolution of the physical parameters and of the sequences of events, mainly differing for the actual timing at which these occur. Other sequences may instead differ in behavior, because characterized by different combinations of occurred events, and still reach the same final outcome state. Hence, the difficulty in identifying and grouping similar scenarios is in the fact that sequences composed of similar events can correspond to rather different process parameters evolutions and, possibly, end states, depending on the events timing or order of occurrence. Therefore, grouping the scenarios only on the basis of the occurred events and end states may be misleading and accounting of the physical behavior of the process variables ought to be included (Labeau et al., 2000).

3. The scenario classification approach

The underlying idea of the approach proposed in Kopustinkas et al. (2005) is to group the DET-generated scenarios in classes of "similarity", by combining information from both the event sequences and the patterns of evolution of the process variables.

In all generality, this leads to a task of pattern classification, i.e. the partitioning of objects into classes. Often in engineering, the complexity of the problems forces one to resort to empirical pattern classification techniques in which an algorithm is built through a process of learning based on a set of patterns labeled with the class they belong to. These kinds of techniques are termed "supervised" and the available pre-classified data are termed "training" data (Zio and Baraldi, 2005a).

In the case here considered, the objects to be classified are the DET scenarios and the basic steps for their classification are sketched in Fig. 2.

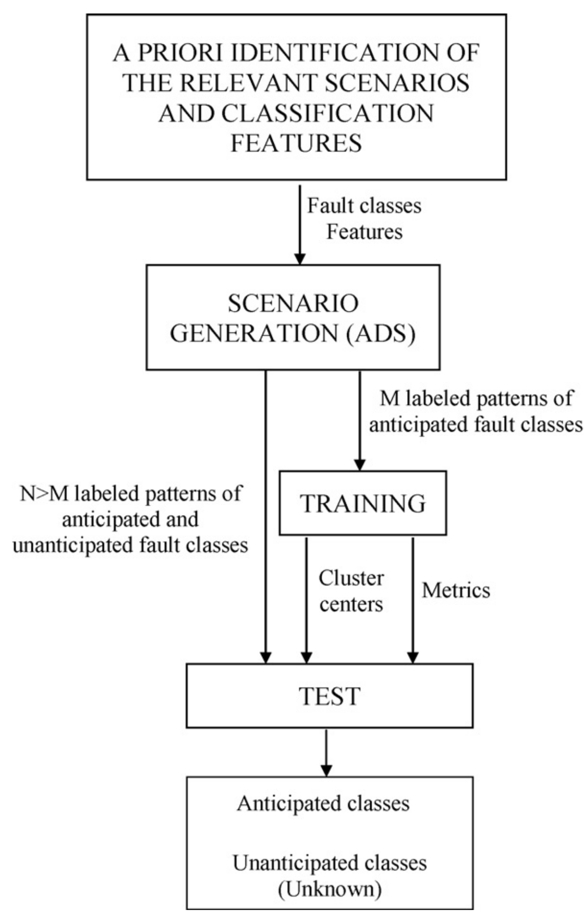

Fig. 2. The scenario classification approach.

The first step is the a priori identification of the anticipated scenario classes for the system under analysis and of the relevant classification features. The scenarios will eventually be classified as belonging to a particular class based on the affinity of their features to those characteristic of the class. Scenario classes should distinguish different reference scenarios that the system is expected to follow in its evolution. They must be defined a priori on the basis of available knowledge on the system operation. For example, in the system studied in Section 6, three classes of scenarios are expected: (1) the nominal operative scenarios; (2) scenarios involving the non-automatic startup of the high pressure injection (HPI) system; (3) scenarios involving both the non-automatic startup of HPI and the failure of a turbine bypass valve (TBV). The identification of the features relevant to the classification is necessary to condense the scenario description into an object vector $\underline{x}$, i.e. the pattern to be fed to the classification function. The features can be either binary or continuous variables. Binary variables characterize the scenarios based on the occurrence or not of certain events, for example the intervention or failure of a safety system; continuous variables characterize the scenario based on the evolution of the process variables.

The successive steps of the procedure are typical of a supervised classification scheme: training of the classifier on patterns of known classes and test of the classifier on new patterns. As for the classification technique, an evolutionary possibilistic FCM classifier paradigm is used in this work. An important asset of this technique is related to the use of the possibilistic classifier for recognizing unanticipated scenarios, referred to in the following as "unknown", i.e. patterns of evolution that were not foreseen as reference in the a priori analysis and thus do not fall in any scenario class. The evolutionary possibilistic FCM classifier filters them out avoiding that they be misclassified as known scenarios and, more importantly, revealing new dynamic failure patterns that were not identified in the a priori analysis. The identification of new, unforeseen evolutionary patterns completes the analyst knowledge on the system with information on unexpected failure scenarios and may aid to suggest additional and more effective safety-oriented improvements of the system.

The procedure for constructing the evolutionary possibilistic clustering algorithm is composed by a possibilitic clustering algorithm that finds the geometric clustering in the feature space based on a Mahalanobis metrics. Then, if the obtained clustering partitions are not close to the physical classes, the Mahalanobis matrix is updated using an evolutionary fuzzy C-means algorithm (Zio and Baraldi, 2005a). The optimal metrics thereby obtained are then used for the classification problem (Krishnapuram and Keller, 1993).

Once the scenarios are classified properly, the probability of each class can be estimated and the dominant evolutionary patterns identified, in terms of both failure event sequences and process variable evolutions.

4. The supervised evolutionary possibilistic clustering algorithm for classification

This section focuses on the classification technique adopted in the work. The combination of the fuzzy clustering classification with a possibilistic clustering for recognizing unknown patterns has been recently introduced by some of the authors in connection with the classification of nuclear plants transients, in an effort to aid the plant operators in diagnosing the causes of the transients (Zio and Baraldi, 2005a,b; Zio et al., 2005). The approach was developed to avoid misclassification of scenarios that were possibly overlooked in the a priori identification step (Section 3). As a result of the introduced approach, the unknown transients are labeled as unknown by the evolutionary possibilistic FCM clustering algorithm.

4.1. Fuzzy and possibilistic clustering

Fuzzy clustering algorithms have been widely studied and applied in various domains such as taxonomy, medicine, geology, business, engineering, image processing and others. The interested reader is referred to Yang (1993) for an extensive literature review.

Fuzzy clustering classification is based on the fuzzy partition of each pattern $\vec{x}_k$, $k = 1, 2, \ldots, N$, into c available classes:

$$0 \le \mu_{ik} \le 1 \quad i = 1, 2, \ldots, c, \quad k = 1, 2, \ldots, N \tag{1}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1 \quad k = 1, 2, \ldots, N \tag{2}$$

In particular, the 'probabilistic' constraint (2), that the memberships μ_{ik} of a given pattern must sum up to 1, is a generalization of the condition which ensures that in a 'hard' (crisp) partition a pattern is a member of one class only and avoids the trivial solution of all memberships equal to 0. As a result of this constraint, the membership of a pattern to a cluster depends on the memberships to all other clusters, i.e. geometrically speaking, it depends on where the pattern is located with respect to not only that cluster but also the others. Hence, in the framework of fuzzy clustering the membership functions take the meaning of degrees of sharing, i.e. they measure how much a pattern belongs to a cluster relatively to the others.

The values of the found memberships can serve as a confidence measure in the classification (Keller et al., 1985): for example, if a pattern is assigned 0.9 membership in one class and 0.05 membership in two other classes we can be reasonably sure that the class of 0.9 membership is the class to which it belongs. On the other hand, if a pattern is assigned 0.55 membership in class A, 0.44 membership in class B, and 0.01 membership in class C, then we should be hesitant to assign it to a specific class based on these results.

Under these conditions, two major drawbacks arise (Dubois and Prade, 1988):

1. The constrained memberships cannot distinguish between 'equal evidence' and 'ignorance' or, in other words, between 'equally likely' and 'unknown' membership to a cluster.
2. Since most distance functions used in fuzzy clustering are geometric in nature, 'noisy' patterns, i.e. lying far from the clusters, can drastically influence the estimates of the clusters prototypes and, hence, the final partition and the resulting classification.

In this situation, an 'unknown', atypical pattern not belonging to any cluster would still belong more to one cluster than to the others, relatively speaking, even if it lies far from all clusters in the feature space and thus it may receive high membership values to some clusters. On the contrary, in our application it is required that unknown, atypical patterns be recognized as such, i.e. bear low membership to all clusters. In this respect, thus, the 'conservation of total membership' constraint (2) is too restrictive since it gives rise to relative membership values, dependent on the number of clusters.

To overcome the above limitations, the clustering problem can be recast into the framework of possibility theory (Dubois and Prade, 1988; Klir and Folger, 1988). In this view, the memberships of representative (typical) patterns are high, while unrepresentative (atypical) points bear low memberships to all clusters. In this interpretation, the membership function μ_{ik} represents the degree of compatibility of the pattern $\vec{x}_k$ with the prototypical member $\vec{v}_i$, i.e. the center, of cluster i. If the classes represented by the clusters are thought of as a set of fuzzy sets defined over the universe of discourse (UOD), then there should be no constraint on the sum of the memberships. The only constraint is that the membership values do represent degrees of compatibility, or possibility, i.e. they must lie in [0,1]. This is achieved by substituting the fuzzy clustering constraints (1) and (2) with the following (Krishnapuram and Keller, 1993):

$$0 \le \mu_{ik} \le 1 \quad i = 1, 2, \ldots, c, \quad k = 1, 2, \ldots, N \tag{3}$$

$$\max_i \mu_{ik} > 0 \quad k = 1, 2, \ldots, N \tag{4}$$

where constraint (4) simply ensures that the set of fuzzy clusters covers the entire UOD.

A possibilistic partition derived under these constraints defines a set of distinct, uncoupled possibilistic distributions (and the corresponding fuzzy subsets) over the UOD (Krishnapuram and Keller, 1993).

4.2. The evolutionary possibilistic FCM clustering

The traditional, unsupervised possibilistic algorithm based on a Euclidean metric to measure compatibility leads to spherical clusters that rarely are adequate to represent the data partition in practice. A significant improvement in classification performance is achieved by considering a different Mahalanobis metric for each cluster, thus obtaining different ellipsoidal shapes and orientations of the clusters that more adequately fit the a priori known data partition (Zio and Baraldi, 2005b; Yuan and Klir, 1997).

The information on the membership of the available patterns $\vec{x}_k$, $k = 1, \ldots, N$, to the c a priori known classes, can be used to supervise the algorithm for finding the optimal Mahalanobis metrics such as to achieve geometric clusters as close as possible to the a priori known physical classes. Correspondingly, the possibilistic clustering algorithm is said to be constructed through an iterative procedure of 'training' based on a set of available patterns, pre-labeled with their possibilistic memberships to the a priori classes. The training procedure for the optimization of the metrics is carried out via an evolutionary procedure, presented in the literature within a supervised fuzzy clustering scheme (Yuan et al., 1995) and further extended to diagnostic applications (Zio and Baraldi, 2005a). Here, the procedure is employed within the possibilistic clustering scheme.

To this purpose, the distance $D(\Gamma_i^t, \Gamma_i^*)$ between the set Γ_i^t (t = true) of memberships of the N available patterns to the a priori known class i and the corresponding set Γ_i^* of the possibilistic memberships to cluster $i = 1, 2, \ldots, c$, is computed by:

$$D(\Gamma_i^t, \Gamma_i) = \sum_{k=1}^{N} \frac{|\mu_{ik}^t - \mu_{ik}^*|}{N} \tag{5}$$

where $0 \le \mu_{ik}^t \le 1$ is the a priori known (possibilistic) membership of the kth pattern to the ith physical class and $0 \le \mu_{ik}^* \le 1$ is the possibilistic membership to the corresponding geometric cluster in the feature space.

The target of the supervised optimization is the minimization of the distance $D(\Gamma^t, \Gamma^*)$ between the a priori known physical class partition $\Gamma^t \equiv \Gamma_1^t, \Gamma_2^t, \ldots, \Gamma_c^t$ and the obtained geometric cluster partition $\Gamma^* \equiv \Gamma_1^*, \Gamma_2^*, \ldots, \Gamma_c^*$:

$$D(\Gamma^t, \Gamma^*) = \sum_{i=1}^{c} \frac{D(\Gamma_i^t, \Gamma_i^*)}{c} = \sum_{i=1}^{c} \sum_{k=1}^{N} \frac{|\mu_{ik}^t - \mu_{ik}^*|}{N \cdot c} \tag{6}$$

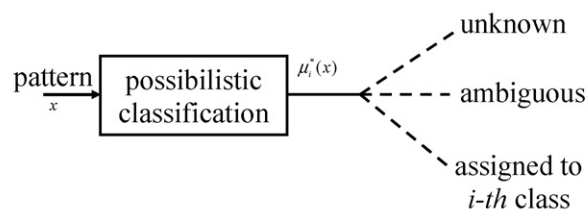

Fig. 3. Classification of pattern $\vec{x}$.

The optimal membership functions μ_{ik}^* $i = 1, 2, \ldots, c$, $k = 1, 2, \ldots, N$, result from the minimization of the objective function:

$$J_m(\Gamma, \vec{v}) = \sum_{i=1}^{c} \sum_{k=1}^{N} (\mu_{ik})^{r_m} s_{ik} + \sum_{i=1}^{c} \eta_i \sum_{k=1}^{N} (1 - \mu_{ik})^{r_m} \tag{7}$$

where η_i are suitable positive numbers and r_m is an index that determines the fuzziness of the final possibilistic partition and the shape of the possibility distribution ($r_m \to 1$, the membership functions are hard, $r_m \to \infty$ they are maximally possibilistic).

The distance $s_{ik} \equiv s_i(\vec{x}_k, \vec{v}_i^*)$ in Eq. (7) between the pattern $\vec{x}_k$ and the optimal cluster center $\vec{v}_i^*$ is computed by:

$$s_i(\vec{x}_k, \vec{v}_i^*) = (\vec{x}_k - \vec{v}_i^*)^{\mathrm{T}} \underline{\underline{M_i}} (\vec{x}_k - \vec{v}_i^*) \tag{8}$$

$\underline{\underline{M_i}}$ being the metric for the cluster i proposed by the evolutionary supervised procedure and T denoting the transpose operator.

4.3. The classification algorithm

When fed with a new pattern $\vec{x}$ the classification algorithm provides the values of the membership functions $\mu_i^*(\vec{x})$, $i = 1, 2, \ldots, c$, to the possibilistic clusters

$$\mu_{ik}^* = \frac{1}{1 + (s_{ik}/\eta_i)^{1/r_m - 1}} \tag{9}$$

These values give the degree of compatibility or "typicality" of $\vec{x}$ to the c clusters. In practice, three situations may arise (Fig. 3):

1) $\vec{x}$ does not belong to any cluster with enough membership, i.e. all the membership values $\mu_i^*(\vec{x})$ are below a given threshold ε_f (degree of ignorance): this means that $\vec{x}$ is an unanticipated (unknown) pattern with respect to the training patterns.
2) at least two membership values are above the threshold ε_c (degree of confidence): $\vec{x}$ is thus ambiguous. In this case, the ambiguity must be regarded as "equal evidence", i.e. the pattern is typical of more than one class and thus cannot be assigned to a class with enough confidence. This situation occurs if $\vec{x}$ is at the boundary between two classes.
3) $\vec{x}$ belongs only to a cluster with a membership value greater than the threshold ε_c: in this case, it is assigned to the corresponding class.

5. Description of the case study: SGTR scenario in a NPP

In this work we consider a nuclear power plant which behaves like a typical commercial pressurized light water reactor (Fig. 4). The main parts of the plant are: the reactor core which creates heat, the pressurized-water in the primary coolant loop which carries the heat to the steam generator, and the steam generator which vaporizes the water in a secondary loop to drive the turbine which produces electricity.

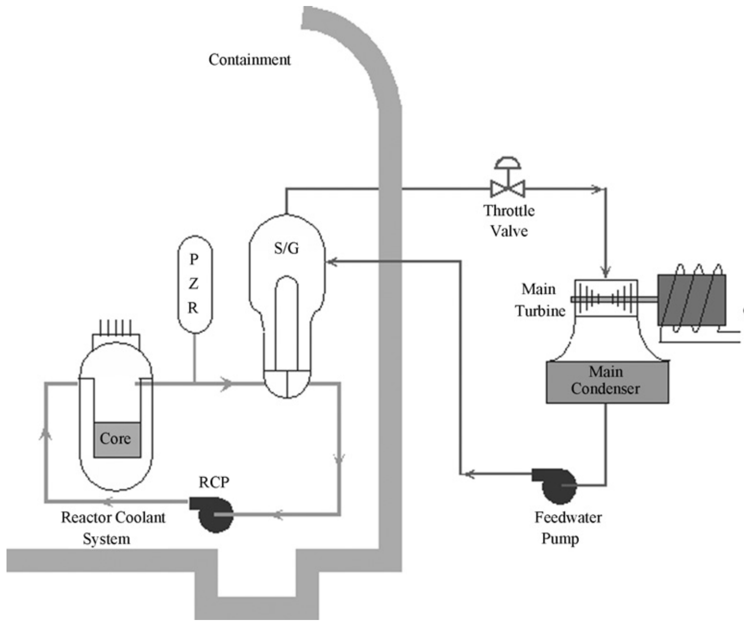

Fig. 4. Diagram of a typical pressurized light water reactor (from http://www.nrc.gov/reactors/pwrs.html).

5.1. Steam generator tube rupture (SGTR) scenario

The steam generator (SG) is an important component of a nuclear power plant (NPP) not only because it transfers heat from the reactor coolant system (RCS) to the secondary steam system, but also because it is a barrier against the release of radioactive material to the secondary side, which in turn presents potential leakage paths outside the containment. In a SGTR accident this barrier is broken due to the rupture of one or more SG tubes.

In this study we consider a SGTR event in loop A of a two-loop pressurizer water reactor (PWR) NPP. The SGTR event is an accident initiating event (IE) which induces a number of abnormal conditions (e.g. low RCS pressure, high radiation in the secondary steam line, etc., . . .), which automatically command: (1) the actuation of the high pressure injection (HPI) system (after about 80 s) to inject cool water to cooldown the RCS and provide RCS inventory make-up, and (2) trip of the reactor and of the Main Feed Water (MFW) system (after about 115 s). Upon the reactor trip, the turbine trips as well, the core power rapidly decreases, and the TBVs open to control the secondary side pressure.

The response of the control room operators is guided by the emergency operating procedures (EOPs) and conditioned by their training. Upon the IE, the operators enter into EOP *immediate and subsequent manual actions*. The first steps entail checking the behavior of the key plant parameters and the status of the key plant safety systems. Then, the operators must go through a number of decision points that would direct them to initiator-specific EOPs: the *loss of subcooling margin* EOP, the *excessive heat transfer* EOP; the *steam generator tube leak* EOP. In particular, the entry criteria for the *steam generator tube leak* EOP are: high radiation levels in the secondary side, incontrollable level increase in one SG, and mismatch in the FWs to the SGs.

Once the operators have entered the *steam generator tube leak* procedures, their main goals are to:

- maintain the pressurizer (PZR) level,
- maintain the subcooling margin,
- depressurize and cooldown the RCS (so as to minimize the leak from the RCS to the secondary side),
- isolate the ruptured SG,
- control and maintain key plant parameters.

The crew model implemented in the ADS includes three types of actions, in an effort to reproduce the operators behavior as realistically as possible: actions guided by the EOPs, actions guided by the so-called mental procedures, and cognitive actions.

The actions of the first type are directly taken from the EOPs and are intended to model the crew's following the procedures. Actions belonging to *immediate and subsequent manual actions* and *steam generator tube leak* EOPs have been implemented in this work (Table 1). The second type of actions is carried out following the so-called mental procedures (Dang, 1996). These are plant operating procedures memorized by the operator and are based on formal procedures. Mental procedures are task-oriented, i.e. each mental procedure corresponds to a specific task. Table 2 lists the modeled actions of this type.

Rule-based actions comprise a third type (De Carvalho and Paulo, 2006). These are intended to model the fact that, although during an accident (as well as during normal operations) the typical control room operators' response is mainly guided by procedures, in many situations it has been observed that operators do not follow the procedure by rote. In other words, they do not apply these "mechanically" or without judgment. Even when guided by procedures, situation assessment and response planning based on the knowledge and training of the operators continue to be important for successful operator performance (Roth et al., 1994). In this work, rule-based actions model this assessment or planning. These rules are instructions acquired by the operator through training, e.g. provided by supervisors. These are listed in Table 3.

5.2. Branching generation

As said earlier, the ADS is used to generate the DET that develop from the SGTR initiating event. The frequency of the IE is 0.005/*y*. The simulation covers the first 2000 s (about 33 min) after the IE. The branching points and the corresponding branching probabili-

Table 1
Operators' crew actions directed by the emergency operating procedures (EOPs) modeled in the SGTR scenario

EOP	Modeled actions
Immediate and subsequent manual actions	Manually trip the reactor Verify reactor shut down If emergency system is not required, maintain pressurizer level above 100 in. If any subcooling margin = 0 °F, go to loss of subcooling procedures If the heat transfer is or has been excessive, go to excessive heat transfer procedures If any of the radiation alarms start, go to steam generator tube leak procedures
Steam generator tube leak	Identify the SG with the leak Maintain the pressurizer level >80 in. Adjust the turbine bypass valves to maintain the steam pressure below 950 psig Depressurize the reactor coolant system to minimize the core subcooling margin Depressurize and cooldown the RCS When the RCS is below 532 °F and if more than one Reactor Coolant Pump (RCP) loop is operating then secure RCPs for one RCP/loop operation Check any subcooling margins If pressurizer level is above 375 in. reduce RCP pressure below 2000 psig and open the pressurizer relief block Isolate the SG with the larger tube leak Control the secondary side contamination

Table 2
Operators' crew actions directed by the mental procedures

Mental procedure	Action
Pressurizer level <100 in.	Maintain PZR level to 110 in. by manual control
Subcooling margin (SCM) less than 5 °F	Maintain SCM to 50 °F by manual control

ties are listed in Table 4. Branching points correspond to hardware failure events as well as to human failure events.

Concerning the hardware failure events, the failure on demand of the HPI automatic startup and the failure on demand of the TBV of loop A (TBVA) are modeled. Concerning human action events, the following types of interactions are modeled: cognitive actions (i.e. failure to perceive that the PZR level is below the lower threshold and failure to manually startup HPI); delay in following a mental procedure (turn HPI on when the PZR level is below the lower threshold) as well as a written procedure (SGTR procedure step 7.4);

Table 3
Rule-base cognitive actions

Perceived information	Action
Rate of power >1 (or <−1)	Trip the reactor
Rate of the loop 1 (or 2) hot leg temperature is <−1 (or 1)	Trip the reactor
Pressurizer level >320 in. and the rate of pressurizer level >1	Trip the reactor
Pressurizer level <200 in. and the rate of pressurizer level <1	Trip the reactor
RCS pressure >2350 psig and the RCS rate >0.25	Trip the reactor
RCS pressure <2000 psig and the RCS rate <−0.25	Trip the reactor
All the RCPs are not working	Trip the reactor
The SG operating range levels >70 and alarm flow imbalance on	Trip the reactor
The SG operating range levels <40, alarm flow imbalance on and MFWs has been tripped	Trip the reactor
Alarm flow imbalance on	Trip the reactor
The SG operating range levels >80	Trip the reactor
The SG operating range levels <30	Trip the reactor
The SG pressure <900 psig	Trip the reactor
Pressurizer level <200 in.	Turn on HPI
SCM less than 5 °F	Trip all RCPs

Table 4
Branching points considered in the SGTR event model

	Branching point	Probability[a]
Hardware failure events	Non-automatic startup of HPI	0.01
	TBVA failure	0.01
Human failure events	Failure to perceive that PZR level is below the lower threshold	0.2
	Failure to manually startup HPI	0.4
	Delay (40 s) to turn on HPI when the PZR level <200 in.	0.3
	Delay (20 s) to carry out SGTR procedure step 7.4 (maintain PZR level >80 in.)	0.3
	Failure to check if HPI has automatically started and in the negative case manually start it	0.2

[a] Although the probability values have not been validated, in practice they are regarded as credible values which can be used to apply the proposed classification methodology.

checks on the status of a system (i.e. check if the HPI is automatically started and in the negative case manually re-start it).

Except for the human actions listed in Table 4, the outcome of all the other human actions mentioned in Section 5.1 is assumed to guarantee success. Similarly, all other components and systems involved in the SGTR event except for the HPI and the TBV are assumed to work as designed.

As described in Section 2.1, during the simulation of each branch, the branches frequencies are updated based on the branching point probabilities. The branches are truncated (i.e. the simulation of the branch is interrupted and the ADS moves on to simulate the next branch) when their frequency is below a threshold value of $10^{-9}/y$.

The truncation limit is set to reduce the number of events that constitute each sequence and, therefore, the number of scenarios produced by the ADS. This was done to test the classification approach on a limited number of scenarios. Further research efforts will be devoted to apply the approach on more scenarios, with lower thresholds and longer simulation time windows.

At the end of the simulation, the ADS generated a total of 60 branches (i.e. scenarios to be classified).

6. Application of the approach

6.1. A priori identification of the anticipated scenarios and relevant classification features

Four classes of scenarios are identified a priori:

Class 1: Nominal scenarios (described in Section 5.1), with all components and operators' actions successful.
Class 2: Scenarios with HPI failure to start on demand and success of the TBV of loop A.
Class 3: Scenarios with HPI failure to start and stuck open TBV of loop A.

Figs. 5–7 show the behaviors of three relevant process variables (the RCS pressure, the steam line pressure, and the steam generator level) representative of the three scenario classes. The HPI activation signal is generated on low RCS pressure. At around 115 s, the reactor trips and the RCS pressure immediately decreases. In case of successful HPI start (class 1 in Fig. 5), the RCS pressure recovers until the operators reach the depressurization step in the corresponding EOP and perform it accordingly (around 900 s), using the PZR spray system. The level in SG A increases (Fig. 7) due to the RCS-to-secondary side leak (note that the uncontrollable rise in the level in one SG is one of the cues for diagnosing the SGTR event).

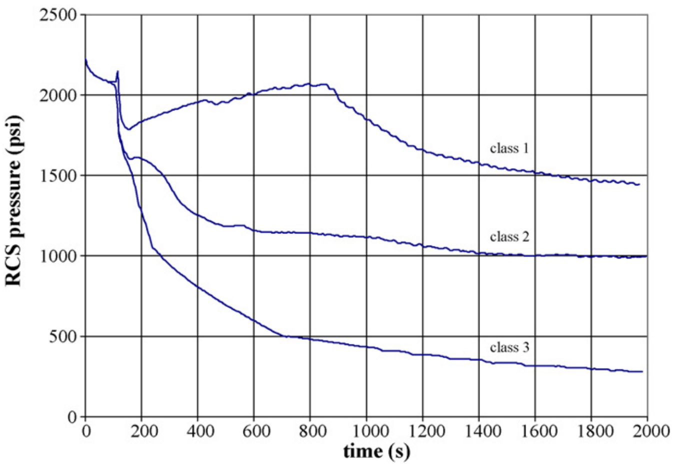

Fig. 5. Behavior of the RCS pressure for classes 1–3 scenarios.

In case the HPI fails to start (classes 2 and 3), the operators are directed by the EOPs to start it manually. In case the operators succeed to manually start the HPI, the scenario evolves as class 1; in case of failure, the RCS pressure continues to decrease (Fig. 5, classes 2 and 3). The pressure decrement in the RCS is faster in case of the stuck open failure of the TBV (class 3), which also causes the drop in the secondary side pressure (Fig. 6).

In case of HPI failure, the level in the ruptured SG remains lower than in case of HPI success (classes 2 and 3 compared to class 1 in

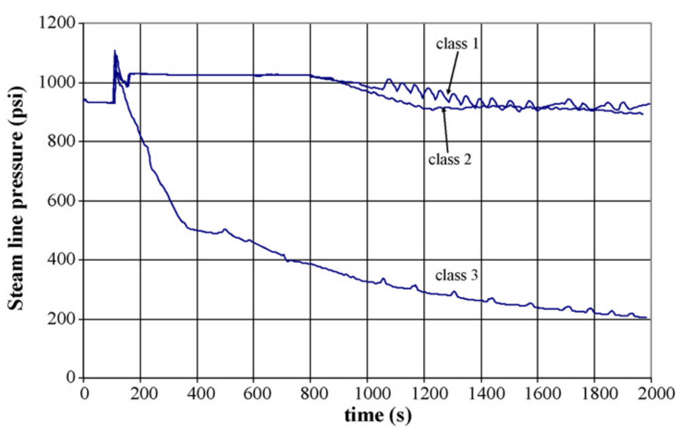

Fig. 6. Behavior of the steam line pressure for classes 1–3 scenarios.

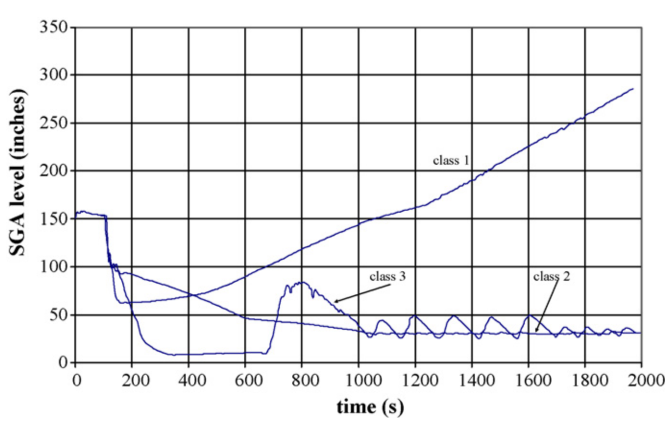

Fig. 7. Behavior of the steam generator A level for classes 1–3 scenarios.

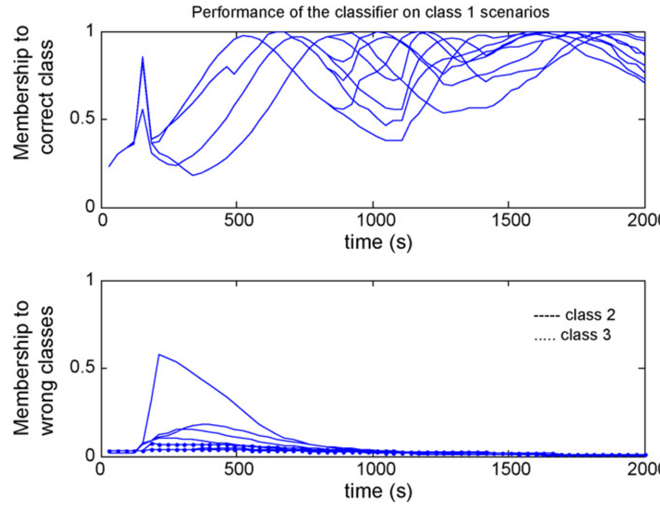

Fig. 8. Performance of the classifier on class 1 scenarios. Upper plot: membership to the correct class 1; lower plot: membership to the wrong classes 2 and 3.

Fig. 7). It is also interesting to notice the saw-tooth behavior of the SG level in Fig. 7, which results from the intermittent functioning of the emergency FW system for SG level control.

Scenarios are classified based on the instantaneous values of the triplet of process variables:

1. RCS pressure;
2. loop A steam line (SL) pressure;
3. loop A SG level.

As previously mentioned, the DET has generated 60 scenarios. The patterns to be classified are triplets of values of the mentioned variables, taken every 5 s along the evolution. Therefore, every scenario is converted into 400 patterns (if the simulation is not interrupted earlier than 2000 s because the truncation frequency threshold is reached). This entails that the classification of the scenarios, which is made on the instantaneous pattern, is time-dependent. It can be foreseen that at early stages in the transient, the classification will be poor since the scenarios belonging to the three classes have similar characteristics (note that up to the HPI branching point all scenarios are perfectly overlapped). On the

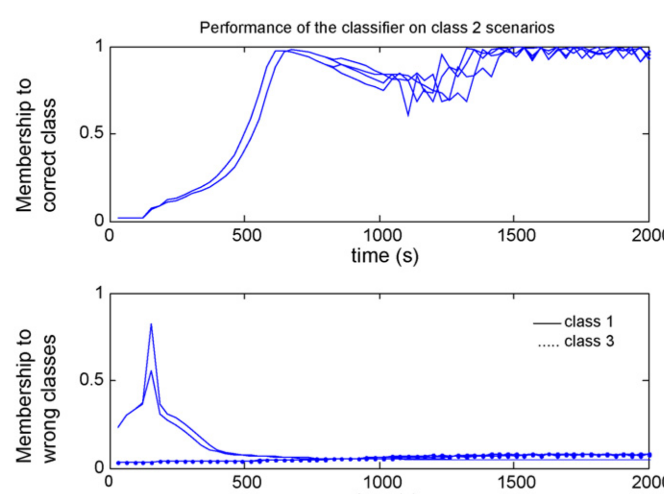

Fig. 9. Performance of the classifier on class 2 scenarios. Upper plot: membership to the correct class 2; lower plot: membership to the wrong classes 1 and 3.

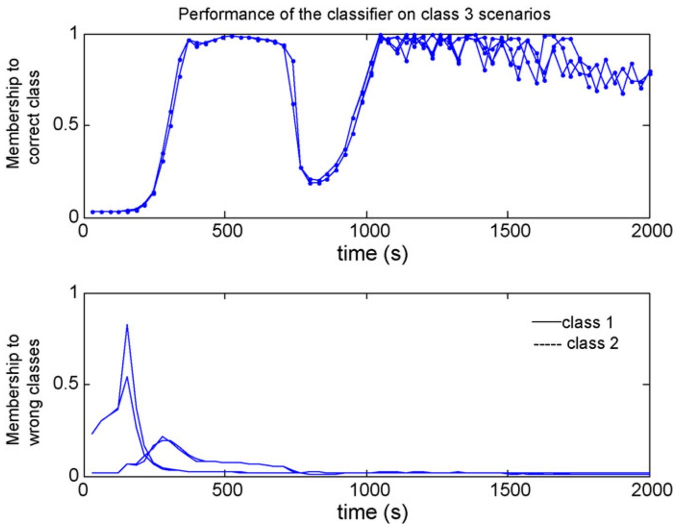

Fig. 10. Performance of the classifier on class 3 scenarios. Upper plot: membership to the correct class 3; lower plot: membership to the wrong classes 1 and 2.

other hand, it is expected that the scenario classification will be more solid as the transient progresses and the scenarios features become distinct.

6.2. Scenario classification and scenario analysis

The main outcome of this step is the grouping of the sampled scenarios in the three a priori defined classes and the identification of unanticipated scenario classes that were possibly left out in the a priori identification of relevant scenarios. These latter unanticipated scenario classes will be referred to as "unknown".

This procedure can be subdivided into two steps:

1. Training of the possibilistic evolutionary FCM classifier on the basis of a set of the nine labeled scenarios (three evolutions for each one of the three process variables) representative of the three classes (Section 6.1). This amounts to a set of 9×400 labeled patterns, each one constituted by a triplet of values of RCS pressure, SL pressure and SG level at a given time, and the integer label of the corresponding scenario class.
2. Use of the possibilistic and fuzzy classifiers to classify the 60 DET-generated scenarios.

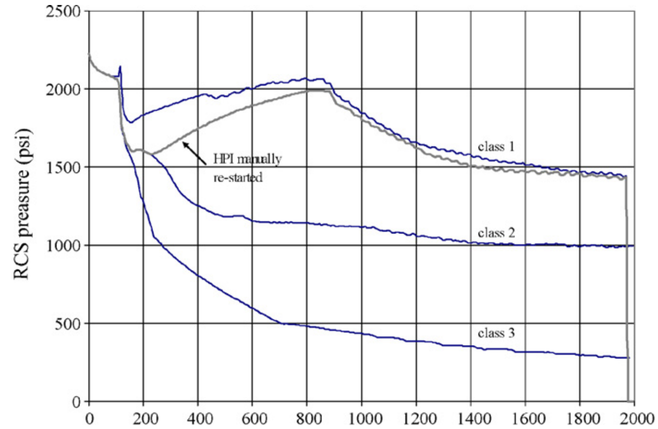

Fig. 12. Evolution of the RCS pressure: effect of manual HPI start on class 1 scenarios.

The search of the optimal membership functions has been done setting the value of the degree of fuzziness r_m to 2.0, which means a low degree of fuzziness in the resulting partition. The possibilistic evolutionary FCM procedures are fed with the above labeled patterns to identify the centers and metrics of the clusters in the three-dimensional space.

The second step of the procedure entails that the 60 DET-generated scenarios be fed into the possibilistic evolutionary FCM classifier for the scenario classification. Results are shown in Figs. 8–10. The upper part of the figures shows the membership values assigned to the scenarios of a given class (e.g. class 1 for Fig. 8) to that specific class (i.e. the correct class 1), as a function of time. The lower part of the figures shows the membership values

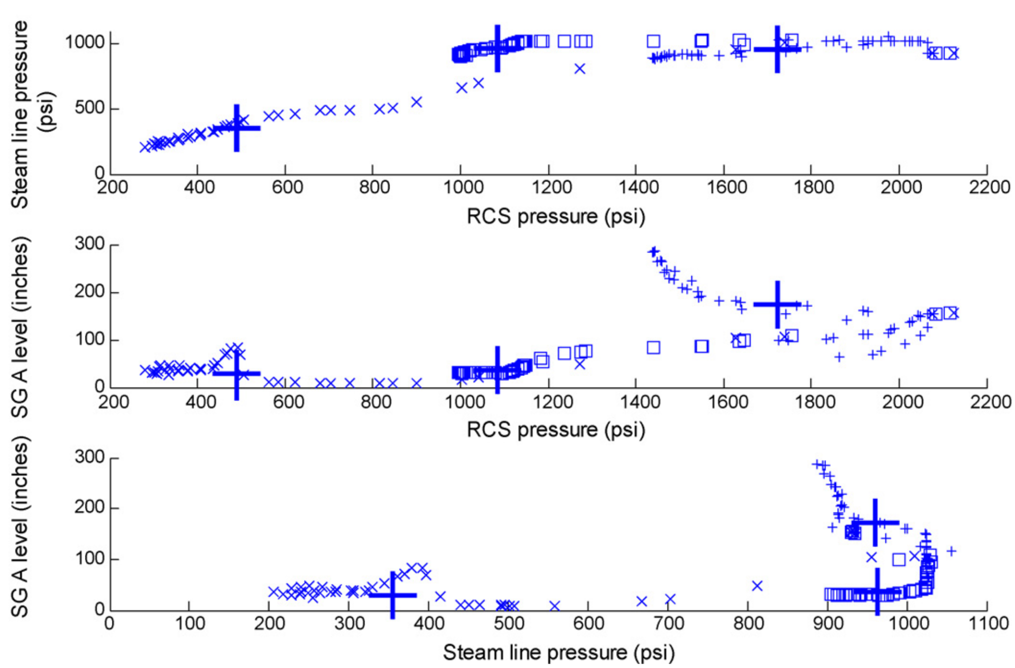

Fig. 11. Position of the patterns belonging to class 1 (+), class 2 (□), class 3 (×) and position of the corresponding cluster centers (+).

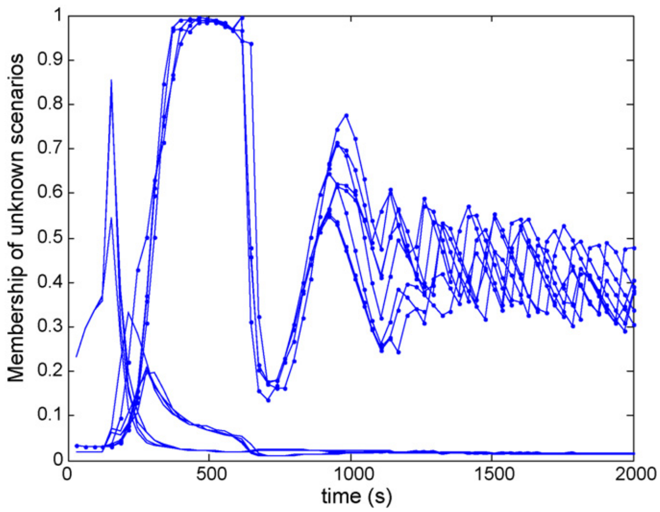

Fig. 13. Performance of the classifier on "unknown" scenarios: (—) membership to class 1, (- -) membership to class 2, (- · -) membership to class 3.

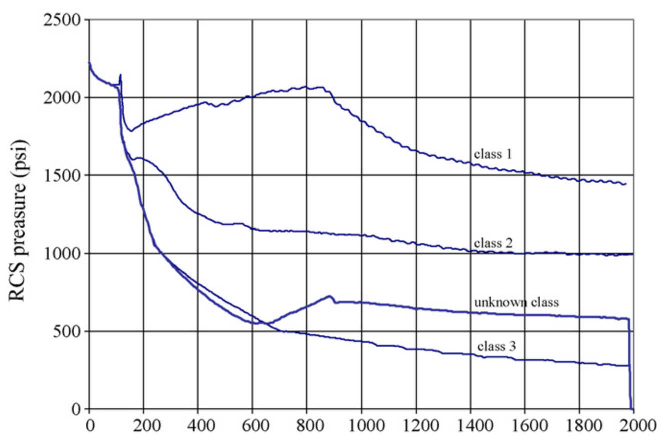

Fig. 14. Evolution of the RCS pressure for classes 1–3 and "unknown" scenarios.

assigned to the other two classes (i.e. the incorrect classes 2 and 3). The figures confirm the good performance of the classifier, which produces high values of membership to the correct class and low values of membership to the incorrect classes.

Figs. 8–10 show that in all the three cases the possibilistic evolutionary FCM classifier is able to recognize the correct class in the mid-late part of the scenario. In fact, as anticipated in Section 6.2, at the beginning of the scenario (until about 100 s) the process behavior is in practice the same for the three classes (Figs. 5–7). The early stage of any scenario is usually assigned to class 1(upper part of Fig. 8, and lower part of Figs. 9 and 10) because this class covers better the features of the early scenarios, i.e. "high" values of RCS pressure, SL pressure and SG level (Fig. 11). After about 100 s the correct scenario classes start to be identified, as their characteristic features become distinguishable. Note that the early time classification as class 1 is reasonable since, at early times, all scenarios actually belong to the nominal class 1 (since failures of the HPI and/or of the loop A TBV have not came in yet).

It is interesting to note that a number of class 1 scenarios have low membership value at the beginning of the scenario, around 500 s (Fig. 8, upper part). These are scenarios in which the HPI failed to startup on demand and the operators were successful to start it manually, as they are instructed to do by procedures. The behavior of the RCS temperature for this type of scenarios is shown in Fig. 12. The earlier stage of the scenario is similar to class 2 scenarios (HPI failed), thus explaining the rather high values of membership to class 2 (lower part of Fig. 8). After the operators have recovered the HPI, the scenarios proceed as those of class 1 (Fig. 12). Accordingly, they are recognized as such by the classifier (the membership to class 1 returns high and that to class 2 drops to zero).

Fig. 10 shows that the membership values of class 3 scenarios drop between about 800 and 1000 s, then restoring back to high values. This can be understood by looking at the behavior of the SG level for class 3 scenarios in Fig. 7. In the mentioned 800–1000 s time window there is a rise in the SG level before the emergency feed water (EFW) control brings the level down to below 50 in.. The level increment carries the patterns away from the cluster centers, thus making the classification more problematic. These patterns

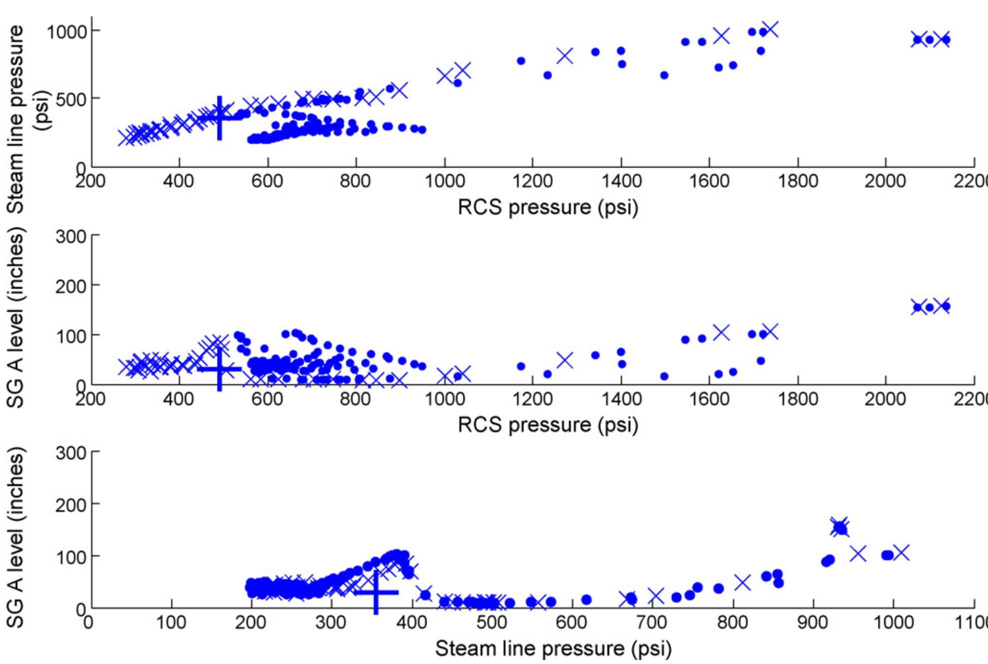

Fig. 15. Position of the patterns belonging to class 3 (×) and unknown (●) and position of the cluster centers (+).

can be identified in the far right side of Fig. 15, which shows the projections of the clusters onto the features planes.

6.2.1. Analysis of the "unknown" scenarios

The possibilistic evolutionary FCM classifier identified 18 scenarios as unknown. Their membership values are shown in Fig. 13. It can be seen that these values are not large for any of the available classes: membership to class 3 has the highest values, ranging within 0.2 and 0.5, while those to the other classes are in practice zero.

The events that lead to these scenarios have been analyzed a posteriori and it was found that they are characterized by failure of the loop A TBV, with correct functioning of the HPI. Fig. 14 shows the behavior of the RCS pressure for these scenarios. The evolution is rather similar to that of class 3 scenarios (failure of both HPI and loop A TBV), and in practice coincident up to about 600 s after the IE. After 600 s, the effect of HPI makeup is to increase the RCS pressure, compared to the class 3 behavior with HPI failed (see Fig. 14). This explains the large values of membership to class 3 up to about 600 s and why these scenarios still maintain some degree of membership to class 3 at later times. Fig. 15 shows the position of the unknown patterns as compared to class 3 patterns, further confirming that the features of class 3 and unknown patterns are rather similar.

7. Conclusions

This paper has addressed the problem of identifying and grouping scenarios resulting from dynamic safety and reliability assessments, for which the number of scenarios that are simulated is much larger than that of the classical fault/event tree approaches and the post-simulation information retrieval becomes a difficult challenge.

Information on the dominant system scenarios can significantly aid the analyst in identifying safety bottlenecks and verifying the efficiencies of the devised protections. For this information to be of practical use, the scenarios should be grouped together on the basis not only of the occurred events and their end states as done in the classical fault/event tree approaches, but also of the physical evolution of the process variables, which may depend on the order and timing of occurrence of the events.

The case study presented in the paper considers scenarios generated in a dynamic event tree (DET) analysis of a steam generator tube rupture event, using the accident dynamic simulator (ADS). The ADS is coupled to a RELAP code that simulates the thermo-hydraulic evolution of the plant process and to an operators' crew model, which simulates their cognitive and procedures-guided responses. Hardware as well as human failure events are modeled. A total of 60 scenarios are generated by the ADS, resulting from the combination of the possible realizations of the success/failure events.

The approach used to tackle the problem is based on evolutionary possibilistic fuzzy C-means clustering classification. Reference scenarios that the plant is anticipated to follow in its evolution are defined a priori on the basis of the expected intervention or failure of safety systems and operators. Then, DET simulation is used to generate scenarios which are classified based on their affinity with the reference scenarios. An important asset of the technique proposed is related to the use of the possibilistic framework: if during the process, unanticipated scenarios are generated, i.e. patterns of evolution that were not foreseen as reference in the a priori analysis, the possibilistic paradigm enables to classify them as unknown scenarios and, more importantly, reveal new dynamic failure patterns that were not identified a priori.

In the case study the classification approach allowed grouping the 60 DET-generated scenarios into 4 classes of dominant scenarios, one of which was not anticipated a priori but was "discovered" by the classifier.

This work represents the first step towards the application of identification and classification approaches to scenarios post-processing for real-scale dynamic safety assessments. The scenarios result from a sophisticated DET code that couples realistic models of the plant and crew behaviors. For demonstration purposes, the number of scenarios generated has been kept low by including only a limited number of failure events in the sequences (which do not reach the core damage state). Future steps will be devoted to the operalization of the approach to cases with many scenarios which can realistically be in the order of several thousands.

Acknowledgements

Part of this work is funded by the Swiss Nuclear Safety Inspectorate (HSK), under DIS-Vertrag No. 82610. The views expressed in this article are solely those of the authors.

References

Chang, Y.H.J., Mosleh, A., 1999. Cognitive Modeling and Dynamic Probabilistic Simulation of Operating Crew Response to Complex System Accidents (Ads-Idacrew). Center for Technology Risk Studies, University of Maryland, College Park, Maryland.

Chang, Y.H., Mosleh, A., 2000. ADS–IDACrew: dynamic probabilistic simulation of operating crew response to complex system accidents. In: PSAM 5—Probabilistic Safety Assessment and Management, November 27–December 1, 2000. Osaka, Japan, Universal Academy Press Inc.

Chang, Y.H.J., Mosleh, A., 2006. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents, part 1–part 5. Reliability Engineering and System Safety 82 (8), 997–1101.

Cojazzi, G., 1996. The DYLAM approach for the dynamic reliability analysis of the systems. Reliability Engineering and System Safety 52, 279–296.

Dang V.N., 1996. Modeling cognition for accident sequence analysis: development of a dynamic operator-plant simulation. Ph.D. thesis. Massachusetts Institute of Technology, May 1996.

De Carvalho, Paulo, V.R., 2006. Ergonomic field studies in a nuclear power plant control room. Progress in Nuclear Energy 48, 51–69.

Demmou, H., Khalfaoui, S., Guilhem, E., Valette, R., 2004. Critical scenarios derivation methodology for mechatronic systems. Reliability Engineering and System Safety 84, 33–44.

Dubois, D., Prade, H., 1988. Possibility Theory: An approach to Computerized Processing of Uncertainty. Plenum Press, New York.

Hofer, E., Kloos, M., Krzykacz-Hausmann, B., Peschke, J., Sonnenkalb, M., 2002. Dynamic event trees for probabilistic safety analysis. EUROSAFE-Berlin, vol. 2. Gesellschaft fur Anlagen - und Reaktorwsicherheit, Cologne, Germany, p. 14.

Hu, Y.-S., Modarres, M., 1999. Evaluating system behavior through Dynamic Master Logic Diagram (DMLD) modeling. Reliability Engineering and System Safety 64 (2), 241–269.

RELAP5/MOD3.3 CODE MANUAL, December 2001. Prepared for the Division of Systems Research Office of Nuclear Regulatory Research U.S. Nuclear Regulatory Commission Washington, DC 20555. Information Systems Laboratories, Inc., Rockville, Maryland Idaho Falls, Idaho.

Keller, J., Gray, M., Givens, J., 1985. A fuzzy k-nearest neighbor algorithm. IEEE Transactions on Systems, Man, and Cybernetics SMC-15 (4), 580–585.

Klir, G., Folger, T., 1988. Fuzzy Sets, Uncertainty and Information. Prentice Hall, Englewood Cliffs, New Jersey.

Kloos, M., Peschke, J., 2007. Consideration of human actions in combination with the probabilistic dynamics method MCDET. In: Aven, Vinnem (Eds.), Proceeding of ESREL 2007, Risk, Reliability, and Societal Safety. Taylor and Francis Group, London.

Kopustinkas, V., Augutis, J., Rimkevicius, S., 2005. Dynamic reliability and risk assessmenent of the accident localization system of the Ignalina NPP RBMK-1500 reactor. Reliability Engineering and System Safety 87, 77–87.

Krishnapuram, R., Keller, J.M., 1993. A possibilistic approach to clustering. IEEE Transactions on Fuzzy systems 1 (2), 98–110.

Labeau, P.E., 1996. Probabilistic dynamics: estimation of generalized unreliability thought efficient Monte Carlo simulation. Annals of Nuclear Energy 17, 1355–1369.

Labeau, P.E., Smidts, C., Swaminathan, S., 2000. Dynamic reliability: towards an integrated platform for probabilistic risk assessment. Reliability Engineering and Systems Safety 68, 219–254.

Marseguerra, M., Zio, E., 1996. Monte Carlo approach to PSA for dynamic process systems. Reliability Engineering and Systems Safety 52, 227–241.

Marseguerra, M., Zio, E., 2002. Basics of the Monte Carlo Method with Application to System Reliability. LiLoLe-Verlag GmbH (Publ. Co. Ltd.).

Podofillini L., Zio E., Mercurio D., Dang V.N., 2007. Dynamic safety assessment: scenario identification via a fuzzy clustering approach. Reliability Engineering and System Safety, submitted for publication.

Roth E.M., Mumaw R.J., Lewis P.M., 1994. An empirical investigation of operator performance in cognitively demanding simulated emergencies. NUREG/CR-6208.

Sheng, K.S., Mosleh, A., 1996. The development and application of the accident dynamic simulator for dynamic probabilistic risk assessment of nuclear power plats. Reliability Engineering and System Safety 52, 297–314.

Siu, N., 1994. Risk assessment for dynamic systems: an overview. Reliability Engineering and Systems Safety 43, 43–73.

Woods, D.D., 1984. Some Results on Operator Performance In Emergency Events, vol. 90. Institute of Chemical Engineers Symposium Series, p. 21–31.

Yang, M., 1993. A survey of fuzzy clustering. Mathematical and Computer Modelling 18 (11), 1–16.

Yuan, B., Klir, G., 1997. Data driven identification of key variables. In: Ruan, D. (Ed.), Intelligent Hybrid Systems Fuzzy Logic, Neural Network, and Genetic Algorithms. Kluver Academic Publishers, pp. 161–187 (Chapter 7).

Yuan, B., Klir, G., Swan-Stone, J., 1995. Evolutionary fuzzy c-means clustering algorithm. In: Proc. Fourth IEEE International Conference on Fuzzy Systems, pp. 2221–2226.

Zio, E., Baraldi, P., 2005a. Identification of nuclear transients via optimized fuzzy clustering. Annals of Nuclear Energy 32, 1068–1080.

Zio, E., Baraldi, P., 2005b. Evolutionary fuzzy clustering for the classification of transients in nuclear components. Progress in Nuclear Energy 46 (3–4).

Zio, E., Baraldi, P., Mercurio, D., 2005. Fuzzy clustering classification of nuclear transients with a possibilistic filter for unknown conditions. In: Enlarged Halden Programme Group Meeting, Radisson SAS Lillehammer Hotel, Norway, 16–21 October.