Scenario Aggregation in Dynamic PRA Uncertainty Quantification

Diego Mandelli*, Tunc Aldemir, Alper Yilmaz

The Ohio State University
Nuclear Engineering Program
201 West 19th Avenue, Columbus, OH 43210
*mandelli.1@osu.edu

INTRODUCTION

A new generation of dynamic algorithms has started receiving attention for nuclear reactor probabilistic risk assessment (PRA). These new algorithms, such as the dynamic event tree (DET) methodology [1], are based on the event-tree and fault-tree logic but explicitly account for the time element in modeling the probabilistic system evolution and heavily incorporate system simulation tools to account for possible dependencies among failure events that may arise from the interactions between hardware, software, firmware, process and human.

The DET methodology is capable of quantifying both the effects of variability and model uncertainties on the risk metrics used in PRA. In that respect, it provides an important modality for the PRA modeling of passive safety systems. However, one challenge with the DET methodology is the large amount of data it produces which may be difficult to analyze without appropriate software tools.

The goal of this paper is to illustrate an approach to group and classify the set of scenarios generated by a DET methodology. The approach is based on the Mean-Shift algorithm presented for the first time in [2]. We will show how it is possible to adapt this methodology when we are dealing with a set of transient scenarios. A simple model of a steam generator level controller is used as an example system in order to illustrate its use for aggregation and the classification of transient scenarios.

METHODOLOGY PROPOSED

Mean-Shift Methodology (MSM) [2] is a non parametric iterative procedure that shifts each data point to the average of data points in its neighborhood. The idea behind the MSM is to determine the cluster centers and to assign each point to one cluster center only. By cluster center we mean a region with high observation density. Starting form a generic point (i.e., point s_A in Fig. 1), the algorithm associates a hyper-dimensional sphere which depends on the number of dimension of the state space, centered at that point. The radius of this ball is identified as the bandwidth *BW*.

The idea is to consider all the points that are inside this sphere and determine the center of mass of these points (point $m(s_A)$ in Fig. 1). From a mathematical viewpoint, given a point $s \in S$ in an n-dimensional Euclidean space, the center of mass of s_A is simply:

$$m(s_A) = \frac{\sum_{s\in S} K(s-s_A)s}{\sum_{s\in S} K(s-s_A)} \quad (1)$$

where $K(x)$ is often called the Kernel. The purpose of this function is its ability to assign different weights to different points during estimation of the center of mass. Several Kernels can be found in literature. In this work, we use the Epanechnikov kernel (see Fig. 2) [3]:

$$K(x) = \begin{cases} (1-\|x\|^2) & \text{if } \|x\| \leq BW \\ 0 & \text{if } \|x\| > BW \end{cases} \quad (2)$$

After estimating the center of mass, the MSM determines the calculated position $m(s_A)$ and repeats the calculation for the center of mass for the points included in the ball having identical value of radius (i.e., *BW)* but now centered on $m(s_A)$.

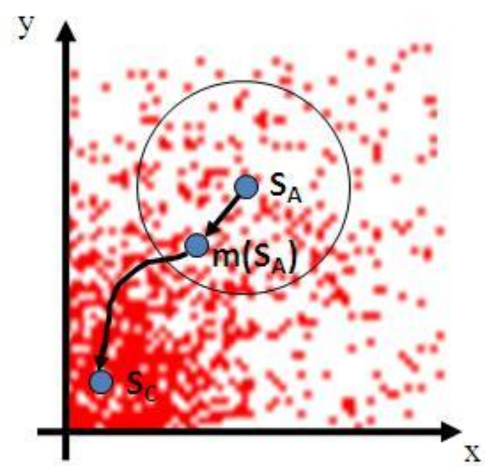

Fig. 1. Graphical representation of the Mean-Shift algorithm

This operation converges to the mode of the data distribution when the distance between the new center of mass and the old one is below a fixed threshold (i.e. s_C in Fig. 1 is reached). When this condition is reached:

- Point s_C is considered the center of a cluster, and
- The original point s_A is uniquely associated to the cluster centered by point s_C.

This procedure is repeated for all the points $s \in S$ to obtain:

- The center of all the clusters in S, and
- The cluster to which each point belongs (each point belongs to one cluster only).

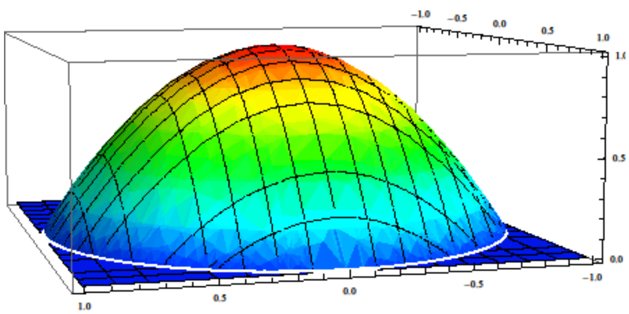

Fig. 2. Epanechnikov kernel with BW=1 in a 2-D space

SYSTEM CONSIDERED

The first test for this methodology is on the analysis of the transients generated by the DET applied to the level controller described in [4] (see Fig. 3). The liquid level is actively controlled through the actuation of three components: two inlet pumps and one outlet valve, hereafter called Unit 1, 2 and 3, respectively.

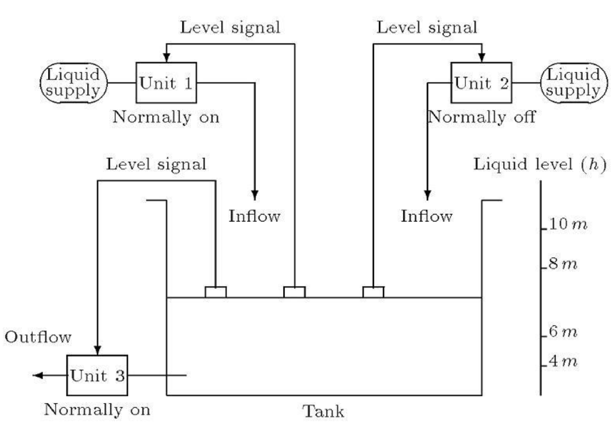

Fig. 3. The example level controller

Each unit is a multi-state component operating either correctly, stuck ON or stuck OFF. At *t=0*, the system is assumed to be in its nominal state (ON,OFF,ON), with equilibrium values of 30.93 ºC of the liquid temperature (*T*) and 7 m of the level (*L*). The temperature of the liquid is assumed to directly affect the failure rates of the components. A thermal power source heats up the fluid to keep it around the nominal temperature *T*, in spite of the level fluctuations. Two possible Top Events are considered: dry-out ($L < 4$ m) and over-flow ($L > 10$ m).

In our DET analysis, the branching is dictated by the failure of the three active components. As shown in [4], the probabilistic behavior of the units can be modeled through Markov approach. Starting at time *t=0* when the system is in nominal condition, the DET algorithm generates a series of different scenarios every hour of simulation depending on the state of the components.

RESULTS

We performed the analysis of the scenarios generated by the DET code with the MSM. The mission time for the DET generation is 4 hours and branching is occurring every hour. Since the system state space consists of 3 variables (time, *L* and *T*), we represented each scenario s_i as a vector in a 10-dimensional space as:

$$s_i = [T(0), L(0), T(1), L(1), \dots, T(4), L(4)] \qquad (3)$$

where *T(i)* and *L(i)* represents the values of temperature and level at time *i* (unit is hour), respectively.

Figures 4 and 5 show the cluster centers for the data generated by the DET for 2 different values of *BW* (i.e., *BW* = 5, 6) and for the 2 Top Events separately. As mentioned earlier, a cluster center can be viewed as the representative scenario for a subset of scenarios (i.e., a cluster of scenarios) where the size of the cluster itself depends on the chosen value of *BW*.

With a broader value of *BW*, the algorithm identifies only 3 different clusters while a narrower value of BW is able to identify a larger numbers clusters (and hence improving the accuracy of the clustering process).

CONCLUSIONS

This paper illustrates a practical methodology to analyze the set of transients generated by the DET approach. Using an example level controller, we have illustrated how it is possible to group the scenarios in clusters and to assign uniquely each scenario to one cluster.

Future work will include the analysis of more complex transients such as sets of data generated by transient analysis codes such as (e.g., RELAP5 [5]).

ACKNOWLEDGMENTS

This work is a product of the project “Risk-Informed Balancing of Safety, Non-Proliferation, and Economics for the Sodium-Cooled Fast Reactor (SFR)” supported by the US Department of Energy under a NERI contract (DE-FG07-07ID14888). The views presented here are those of the authors and do not necessarily represent the views of the US Department of Energy.

REFERENCES

1. B. RUTT, U. CATALYUREK, A. HAKOBYAN, K. METZROTH, T.ALDEMIR, R. DENNING, S. DUNAGAN, D. KUNSMAN, “Distributed Dynamic Event Tree Generation for Reliability and Risk Assessment”, 1-4244-0420-7/06, *IEEE* (2006).

2. K. FUKUNAGA, L. HOSTETLER, “The estimation of the gradient of a density function, with applications in pattern recognition”, *IEEE Transactions on Information Theory*, **21**, 32-40 (1975).

3. Y. CHENG, “Mean Shift, Mode Seeking, and Clustering”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **17**, no. 8, (1995).

4. T. ALDEMIR, “Utilization of the cell-to-cell mapping technique to construct Markov failure models for process control systems”, *Proc. of Probabilistic Safety Assessment and Management: PSAM1*, Elsevier, New York, 1991, pp. 1431-1436.

5. The RELAP5-3D Code Development Team, RELAP5-3D Code Manual, Volume I: Code Structure, System Models, and Solution Methods, INEEL-EXT-98-00834, June 2005.

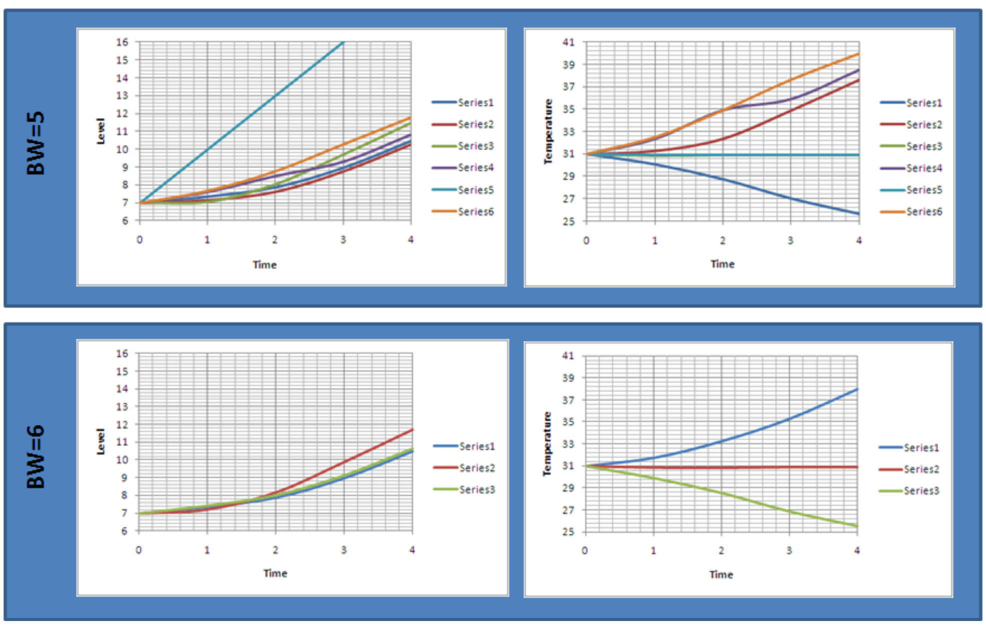

Fig. 4. Representation of the cluster centers for Over-flow for 2 different values of bandwidth

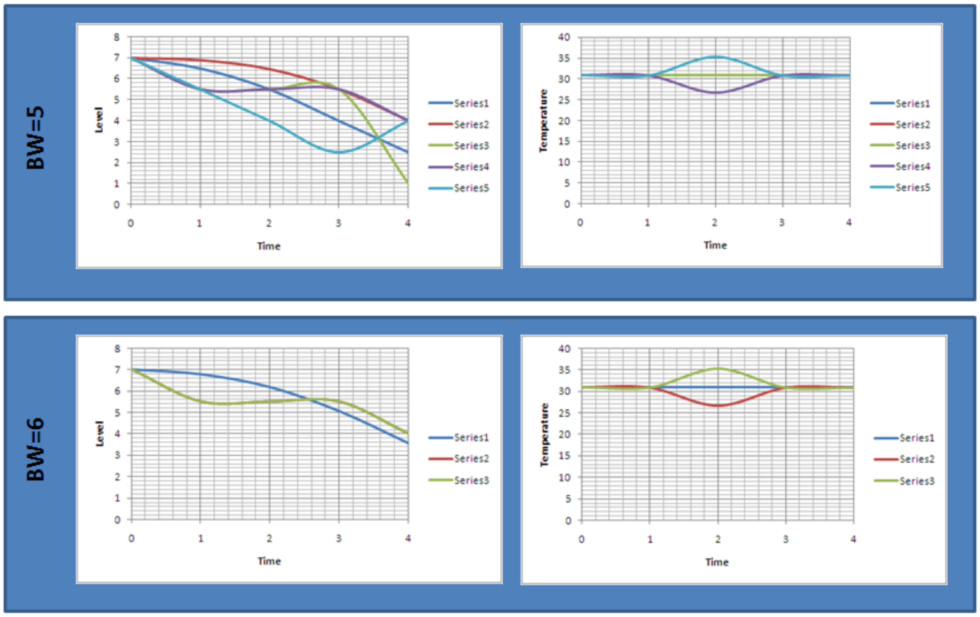

Fig. 5. Representation of the cluster centers for Dry-out for 2 different values of bandwidth