Annals of Nuclear Energy 121 (2018) 89–100

Contents lists available at ScienceDirect

Annals of Nuclear Energy

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

SVR optimization with soft computing algorithms for incipient SGTR diagnosis

Abiodun Ayodeji [a,b], Yong-kuo Liu [a,*]

[a] Fundamental Science on Nuclear Safety and Simulation Technology Laboratory, Harbin Engineering University, Harbin, Heilongjiang 150001, China
[b] Nuclear Power Plant Development Directorate, Nigeria Atomic Energy Commission, Abuja, Nigeria

ARTICLE INFO

Article history:
Received 15 March 2018
Received in revised form 3 July 2018
Accepted 5 July 2018

Keywords:
Fault diagnosis
Support vector regression
Feature selection algorithms
Steam generator tube rupture

ABSTRACT

Fault severity awareness and fault identification are some of the key steps to a successful diagnosis in nuclear power plants. Currently, faults such as leak detection are being done using the N-16 method. However, traditional leak monitors are not sensitive to small leak rate changes, hence cannot be used for low-level leak rate detection under incipient fault conditions and are limited to post-accident analysis of significant releases. In this work, we present a diverse and implementable data-driven Support Vector Regression (SVR) model whose capability compensates for the weaknesses in the already established N-16 methods in the nuclear plant. The method can be integrated with the conventional N-16 method to form a robust hybrid diagnostic system, effective for detecting both incipient and large leakage in the steam generator. The purpose of the SVR model is to estimate uncertain parameters that are sensitive to certain faults, and the parameter estimation efficiency is evaluated using the mean squared error values (MSE). To obtain efficient predictive model capable of supporting decision-making process and to further optimize the model, minimize false alarm rate and reduce computation cost, we also utilized Particle Swarm Optimization algorithm, Sequential Feature Selection algorithm, and Genetic Algorithm for feature selection purposes. To demonstrate the method and evaluate the predictive model, we simulated steam generator tube rupture (SGTR) faults with varying severity in the reactor coolant system of CNP300 NPP, with RELAP5/SCDAP Mod4.0 code. The SVR's relative error (MSE) with and without feature selection algorithms were compared using different solver algorithms. The feature selection performance of the algorithms and the resulting SVR model fault diagnosis performance evaluation are discussed in this paper.

1. Introduction

In a safety critical and complex systems like nuclear power plant (NPP), incipient fault detection and isolation (FDI) is important for a number of reasons: For safe shutdown and timely repair or replacement of failed components or subsystem, on-line fault by-pass to avoid unscheduled downtime and increase availability, and to quantify and evaluate possible consequences of the failure and the likely protective action. NPP online condition monitoring has developed from routine noise analysis technique for plant parameter transmitters and sensing line blockages to beyond sensor equipment condition monitoring such as reactor internal vibration, loose parts monitoring, leak detection, as well as performance evaluation of rotating parts and valves. Moreover, the presence of huge database, derived from parallel measurement of hundreds of plant parameters with diverse sensors, provide better insights into the current state of the plant. Recently, the database has been indispensable in the utilization of the artificial intelligence system and soft computing approaches for early and accurate diagnosis and prognosis of plant transients, and in supporting operator decisions. A few applications of majorly signal-based FDI techniques for instrument calibration monitoring, instrument dynamic performance monitoring, equipment condition monitoring, reactor core monitoring, loose part monitoring, and some transients are reviewed in Ma (2015). However, there is an obvious disconnect between many of these proposed diagnostic techniques and the actual implementation in operating NPPs. Apart from a few utilities and research centers that have developed and/or apply some of these techniques, a significant gap still exists between theory and practice (See Table 1).

Some challenges affecting the practical FDI implementation in operating plants are the complexity of process dynamics, limited ranges of validity of the models, incomplete uncertain data, and

* Corresponding author.
E-mail address: liuyongkuo@hrbeu.edu.cn (Y.-k. Liu).

https://doi.org/10.1016/j.anucene.2018.07.011
0306-4549/

Table 1
Generalized structure of Genetic Algorithm:

```
GA():
InitRandP();
EvaluateFitnessP(V);
While (termination Criteria) do:
Parent selection;
Crossover with probability Pc to form a new offspring;
Mutation with probability Pm;
Fitness Calculation;
Survivor selection
If best solution found then:
Exit
Return the best solution
```

model complexity (Dash and Venkatasubramanian, 2000). To address these issues, a number of researchers (Costache and Minzul, 2012; Ye et al., 2014) have proposed the development and utilization of distributed Support Vector Machine (SVM) and Support Vector Regression (SVR) algorithms with various capabilities for FDI. Ye et al. (2014) present a wavelet and support vector regression (SVR) based method for locating grounded faults in radial distribution systems. The method utilizes traveling wave data recorded at a substation and used the maxima of modal components in each scale as the candidate features for training an SVR. A comprehensive method for integrating the predictive capability of two different intelligent systems to a knowledge-based operator support system for nuclear plant fault diagnosis is also presented in (Ayodeji et al., 2018); although the architectures of the support vector regression can be more easily determined than that of neural networks. Ding and Fang (2017) utilized particle filter and nonlinear regression to predict faults in a nonlinear stochastic system with incipient faults. The effectiveness of the proposed method is verified by the simulations of the three-tank system. In (Ayodeji and Liu, 2018), a fault diagnosis method that utilizes the flexibility of data-driven Support Vector Machine (SVM) for component-level fault diagnosis was proposed. They evaluated the method on a PWR coolant system and compare the performances of different multiclass implementation algorithms, using different coding matrices, as well as different kernel functions on the representative data from a nuclear plant. Liu et al. (2017) proposed the hybrid of Elman Neural Network (ENN) and Signed Directed Graph (SDG) for fault recognition. In (Keskes et al., 2013), wavelet-SVM model was used to detect broken rotor faults in induction machine, and PCA was also introduced for feature extraction.

A practical limitation of SVR is in the rate of false alarm generation, as a result of redundant features, signal noises and model uncertainties. NPP operations are characterized by disturbances, noisy measurements, state variations and a number of false alarms are generated as a result of these transients. Also, uncertainties such as model abstractions, as well as disturbances and high background noise as found in NPPs can obscure fault detection by raising false alarms. Moreover, regression models that involve a large number of redundant features may result in reduced learning speed, performance degradation, and increased probability of over-fitting (Moradi and Gholampour, 2016). Over-parameterization often leads to models that require the estimation of too many parameters for their practical application to be feasible. To compensate for the model uncertainties as a result of plant fluctuations and instability, and to avoid over-parameterization of the SVR model, there is a need for effective feature selection and extraction. A good feature subset selection contributes towards boosting the performance of a regression model - especially when dealing with a high dimensional feature space – by identifying the most discriminative features and removing redundant ones. The identified significant feature subsets constitute the optimized representations of the original problem. Empirical studies indicate that a proper feature selection method not only enhances the decision-making processes in various domains but also accords the learned models with an efficient computational cost and enhance their performance for decision support (Xue et al., 2016).

A number of feature selection algorithm and non-evolutionary statistical metrics such as Principal Component Analysis, Greedy Search, Sequential Feature Selection, and Bayesian Optimization have been explored to assess features' discriminative power (Liu et al., 2017; Xue et al., 2016; Jothi and Inbarani, 2016; Yan et al., 2017). However, some of these methods suffer from local optima, high computational cost, and generalization problems. Comparatively, swarm intelligence algorithm and evolutionary computation such as Genetic Algorithms (GA), Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Differential Evolution (DE), Artificial Bee Colony (ABC), Glowworm Swarm Optimization (GSO), and Cuckoo Search Algorithm (CSA) among others have been shown to possess great capabilities in finding global optima. Specifically, GA and PSO have been shown to demonstrate consistent superiority over other methods in undertaking feature selection problems by employing an evolutionary and swarm-based strategy to yield multiple solutions for complex and non-linear problems (Zhang et al., 2017; Ephzibah, 2011). Xue et al. (2014) proposed a PSO-based feature selection method to select a smaller number of features and achieve better classification performance than using all features. A hybrid method has been proposed in Lin et al. (2008) using a PSO-based method to find optimal feature subsets by combining the PSO with support vector machine (SVM). In addition, Ma et al. (Ma and Xia, 2017) utilized GA for feature selection for a pattern recognition problem. Discussions on these algorithms and their variants can be found in Wahab et al. (2015).

Currently, faults, such as leak detection is being done using the N-16 method. Traditional leak detection system in nuclear plant involves the use of sensors, and the type of sensor largely depends on the type of reactor and the location being monitored. Early indications of a primary-to-secondary leak can be obtained from sensor measurement of certain parameters. However, this method's peculiar drawback is in detection and evaluation of the severity of small leaks. Traditional leak detectors are not sensitive to small leak rate changes and cannot estimate small faults. These monitors would only respond if there was a significant source term in the reactor cooling secondary system, hence, cannot be used for low-level leak rate detection and are limited to post-accident of significant releases (Sohn et al., 2006). In low leak rate scenario, evaluating the size of the leak is equally important, as failure to recognize the potential problems associated with the methods used to calculate primary-to-secondary leak rates can lead to un-conservative estimates of the leak rate. Moreover, many of the models presented as a diagnostic fault size estimation system in the literature lack the capability to evaluate incipient faults, and retrofitting the systems into operating plant is problematic.

Consequently, this work presents a practical method to combine the existing N-16 FDI in a nuclear plant with a data-driven technique, implementable on the operator support system in the NPP. We propose the hybrid of the N-16 method and optimized support vector regression. In this work, high-level fault detection is done by the N-16 method and the support vector regression model diagnose incipient, low-level fault and specifies the severity of the fault. We acknowledge that every feature selection method may give different results in solving different problems. Hence, to select appropriate features to train the SVR model, we experimented with three different feature selection algorithms: Particle Swarm Optimization algorithm, Genetic Algorithm and Sequential Feature Selection Algorithm. To test the method presented in this study, incipient steam generator tube rupture (SGTR) event in the Chinese CNP300 PWR is simulated using a thermal-hydraulic system code, RELAP5/Mod4.0.