

Validating an Online Adaptive System Using SVDD

Yan Liu Srikanth Gururajan Bojan Cukic Tim Menzies Marcello Napolitano
West Virginia University
Morgantown, WV 26505, U.S.A.

Abstract

One of the goals of verification and validation (V&V) activities for online adaptive control systems is providing assurance that they are able to detect novel system behaviors and provide adequate (safe) control actions. Novel (or abnormal) system behaviors cannot be enumerated or fully and explicitly described in requirements documentation. Therefore, they have to be observed and recognized during the operation. Novelty detection methods, therefore, provide an adequate approach for the V&V purposes.

We propose a novelty detection method based on Support Vector Data Description (SVDD) as a candidate approach for validating adaptive control systems. As a one-class classifier, the support vector data description is able to form a decision boundary around the learned data domain with very little or no knowledge of data points outside the boundary (outliers). We apply the SVDD techniques for novelty detection as part of the validation on an Intelligent Flight Control System (IFCS). Experimental results show that the SVDD can be adopted as an effective tool for finding indications of the safe region for the learned domain, whereby we are able to separate faulty behavior from normal events.

1 Introduction

Adaptive systems can be applied to domains where the requirement for autonomy is significant or environmental conditions are unpredictable. The aim of an adaptive system is to perform appropriately under both previously identified and unidentified circumstances through adaptation. If the adaptation occurs after the deployment of the system, the system is called an online adaptive system. In recent years, these systems have attracted increasing interest in application domains such as flight control, robotics and business critical applications.

Online adaptive systems are considered promising because of their plasticity. However, plasticity poses a significant problem in terms of Verification and Validation (V&V). These systems react distinctively with respect to specific data patterns and novel data might cause unstable

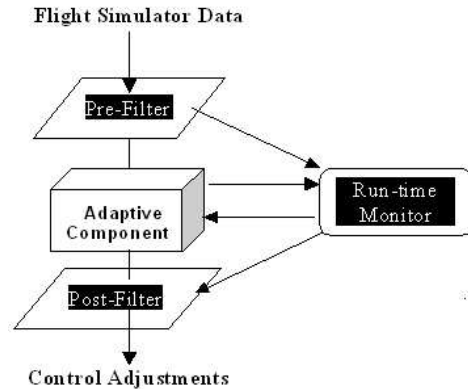


Figure 1: A validation framework for online adaptive systems.

states and potential failures. From safety and reliability assurance points of view, we propose an online validation framework consisting of:

1. Pre-filter, which should be able to block the unreliable or unreasonable training data from entering the online adaptive component and provide backward and forward recovery capabilities,
2. Run-time monitor, dealing with the real-time stability and convergence analysis of the learning device and indicating the current confidence measure, and
3. Post-filter, which should disallow the propagation of unsafe control adjustments into the controller, based on its knowledge of system safety requirements.

Figure 1 illustrates such an assurance architecture for online adaptive systems.

Our work on Step 2 in this validation framework has been focusing on the run-time performance checks using the Lyapunov theory of stability. Although not sufficient, Step 2 is a necessary part of this research ensuring learning (adaptation) stability and peak performance. In recent work [1], we proved that when an online neural network trains on a fixed feature manifold, the evolving state of the system due to the network's position adjustment is self-stabilizing in a globally asymptotically stable manner. This

stability measure, together with the properly applied statistical analysis performed by the pre-adaptation filter, should give us a reliable indication of the validity and safety of the resulting system outputs (control adjustments).

Steps 1 and 3 in our validation framework seek to define a novelty detection technique suitable for alerting to anomalous system adaptation and false prediction (control adjustment). In this paper, we propose a novel approach to building Pre and Post adaptation filters (see Figure 1) using support vector data description for detecting outliers based on the decision boundary obtained from learning on the target data only.

In general, novelty detection techniques require beforehand knowledge of both nominal and off-nominal data domains. However, for a control system such as the Intelligent Flight Control System (IFCS), some of the fault situations (failure modes) are very difficult to anticipate and obtain representative data beforehand. Under such circumstances, the classification performance of regular models is relatively poor due to the restriction of their generalization capability and the low quality of the off-nominal data. As a one-class classification tool, the SVDD method is derived from the Support Vector learning theory by Tax et al [2, 3]. Different from general Support Vector classifiers, which decide the maximum margin hyperplane that separates two classes, the SVDD method tries to find the optimal decision boundary for one given data set. Thus, it provides the best representation of the target-class and offers inferences that can be used to detect the outliers from the nominal feature space, practically defined as the "safe region".

In this paper, we explore the SVDD method on a series of nominal flight data obtained from simulations of an intelligent flight control system. We conducted tests using five different known failure flight modes. Section 2 presents recent research work. In Section 3, we give a brief introduction on the support vector data description technique. The description of the flight control system simulation and the results obtained from the experiments with simulated flight data is provided in Section 4. We conclude the paper with a description of future work in Section 5.

2 Related Work

2.1 V & V of Online Adaptive Systems

Several approaches have been proposed for the verification and validation of online adaptive systems. In principle, analytical methods can provide assurance of system performance with respect to the predefined properties. These are static methods that provide reasoning about the system's functional behavior under certain assumptions. Approximation theory has been applied to analyze the approximation capabilities of adaptive paradigms. In the re-

lated literature, structures such as Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) networks have been proven to be universal approximators. In a recent research effort, Mili et.al. [4] proposed an abstract computational model for online adaptive systems. Their model attempts to capture the functional behavior of an online adaptive system by abstracting away random factors in the function of the system hence focusing exclusively on details that are relevant to the learning algorithm and the learning data. While this is a generic model that establishes functional properties of adaptive systems using refinement-based reasoning, it is difficult to apply it for real-time validation of complex systems.

Empirical methods are widely used in validating adaptive systems. Research work has focused on system evaluation through testing. Popular methods such as cross-validation during training, bias-variance trade-off, etc., are favorite approaches for balancing the memorization and generalization abilities. However, checking all possible inputs is impossible. In an attempt of validating the generalization performance of a RBF neural network by Leonard et. al. [5] the adaptive component is modified to provide support for testing based validation of results. Experimental success in research suggests its significant potential for future use.

2.2 Novelty Detection

In the past decades, several statistical methods for novelty detection have been developed. Popular parametric models are Parzen window, k -nearest model and Gaussian mixture model [6, 7]. These are well-known statistical approaches. Briefly, these methods utilize a certain number of parameters and kernels to build a model for the underlying data statistics. Novelty detection is achieved by evaluating the data based on the kernels and their combinations. Due to the requirement for extensive knowledge of a data domain when building a parametric model, these models are not as applicable and flexible as non-parametric models in real-world data domains.

Recently, learning paradigms using data mining techniques such as Support Vector Machines (SVM) have been investigated in the context of novelty detection. SVM is a classification algorithm that generates a maximum margin hyperplane, which provides "the greatest separation between the classes" [2]. Given a test instance, its distance from the hyperplane can be calculated and, following some threshold, we are able to determine whether the instance is novel. Sample applications in detecting novelties can be found in Scholkopf's paper [8].

Other machine learners can also provide models for novelty detection. Conventional neural networks such as MLP, RBF and Self Organizing Maps (SOM) are popular

mainly because they require no a priori knowledge about the domain [9]. However, such models usually require massive computational effort, thus making online monitoring infeasible. Therefore, we are more inclined to adopt computationally efficient learning techniques.

3 Support Vector Data Description

Support vector data description has been developed by Tax et. al. to solve the one-class classification problem based on Vapnik's Support Vector Machine learning theory [2]. The method of support vector data description originates from the idea of finding a sphere with the minimal volume to contain all data [10]. Given a data set S consisting of N examples $x_i, i = 1, \dots, N$, the SVDD's task is to minimize an error function containing the volume of this sphere. With the constraint that all data points must be within the sphere, defined by its radius R and its center a , the objective function can be translated into the following form by applying Lagrangian multipliers:

$$L(R, a, \alpha_i) = R^2 - \sum_i \alpha_i \{R^2 - (x_i^2 - 2ax_i + a^2)\},$$

where $\alpha_i > 0$ is the Lagrange multiplier. L is to be minimized with respect to R and a and maximized with respect to α_i . By solving the partial derivatives of L , we also have:

$$\sum_i \alpha_i = 1; a = \sum_i \alpha_i x_i,$$

which gives the Lagrangian with respect to α_i :

$$L = \sum_i \alpha_i (x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j),$$

where $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$. By replacing some kernel functions $K(x, y)$ with the product of (x, y) in the above equations we have:

$$L = 1 - \sum_i \alpha_i^2 - \sum_{i \neq j} \alpha_i \alpha_j K(x_i, x_j).$$

According to the solution that maximizes L , a large portion of α_i 's become zero. Some α_i 's are greater than zero and their corresponding objects are those called support objects. Support objects lie on the boundary that forms a sphere that contains the data. Hence, object z is accepted by the description (within the boundary of the sphere) when:

$$\|z - a\|^2 = (z - \sum_i \alpha_i x_i)(z - \sum_i \alpha_i x_i) \leq R^2.$$

Similarly, by applying the kernel function, the formula for checking an object z now becomes:

$$1 - 2 \sum_i \alpha_i K(z, x_i) + \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \leq R^2.$$

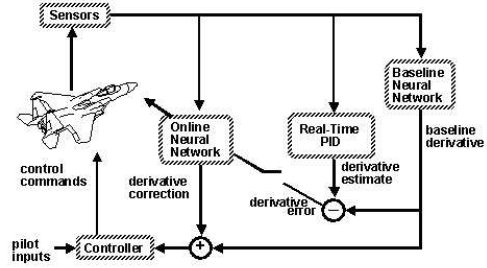


Figure 2: The Intelligent Flight Control System.

By applying kernel functions, we reach a more flexible and stable description of the boundary. However, since the SVDD is used as a one-class classifier, in practice, there are no actual outliers well defined other than those we randomly draw from the rest of the space outside the target class. Hence, by applying the SVDD, we only obtain a sound representation of the target class. To detect outliers, more precise criteria should be inferred from empirical testing or pre-defined thresholds.

4 Experiments

We applied SVDD method for novelty detection of a flight control system in a simulation environment and conducted a series of experiments based on data segments obtained from it. A data description inferred in this way is used for identifying outliers in datasets collected from five different failure mode simulation runs. The Intelligent Flight Control System (IFCS) was developed by the National Aeronautics and Space Administration (NASA) as "a revolutionary flight control system that can efficiently identify aircraft stability and control characteristics using neural networks and use this information to optimize aircraft performance in both normal and simulated failure conditions [11]." Figure 2 shows the architecture of the IFCS employing the Dynamic Cell Structure neural network, referred to as the Online Neural Network. Briefly, the online neural network learns from the discrepancies between the baseline neural network and the real-time parameter identification (PID) according to the sensor data sequences generated by the PID. It also produces parameters to be used by the controller to optimize the flight response of the aircraft under a variety of maneuvering conditions. Furthermore, when abrupt changes to the flight control system occur, the online neural network helps to simulate failure or damage to the aircraft control surfaces. In either scenario, the online learning of the network should provide reliable performance for further mapping of those key control parameters, which motivates us to seek a practical

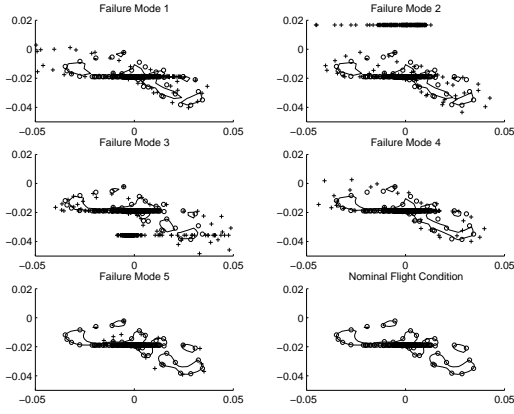


Figure 3: 2D Plot of SVDD of IFCS simulation data testing on five failure modes and the nominal flight condition simulations.

methodology to detect such abrupt changes for validation purposes.

The experiment is conducted by training and testing the SVDD with data from the flight simulator. For the purpose of this experiment, we choose to fly the aircraft in the open loop mode, with and without induced actuator failures. The simulator also allows the user to set the time of failure of the actuator as well as the option to choose the actuator to fail. SVDD training is done with data obtained from the simulator, under nominal conditions. Here, nominal conditions imply that there is no actuator failure induced on the aircraft, during the simulation. Once the SVM is trained and the decision boundary is formed, the test phase is carried out with data from the simulator, under different actuator failure conditions. Each data point can be considered to have multiple dimensions, representing different flight parameters including the angular rates, control surface deflections and Euler angles, among others. For the purposes of initial experiments, we choose the pairing of the longitudinal parameters q , α and θ with respect to the average of the stabilator deflections and monitor the decision boundaries by applying the Gaussian kernel function for the SVDD method.

By applying the support vector data description method, the nominal domain is well learned and a decision boundary is formed. We then collect data sequences under five different actuator failure modes listed below. A failure mode is selected at the beginning of each simulation run.

Mode 1 Actuator Failure - stuck left stabilator at current position (0 degree)

Mode 2 Actuator Failure - stuck left stabilator at predefined deflection (+3 degree)

Mode 3 Actuator Failure - stuck left stabilator at predefined deflection (-3 degree)

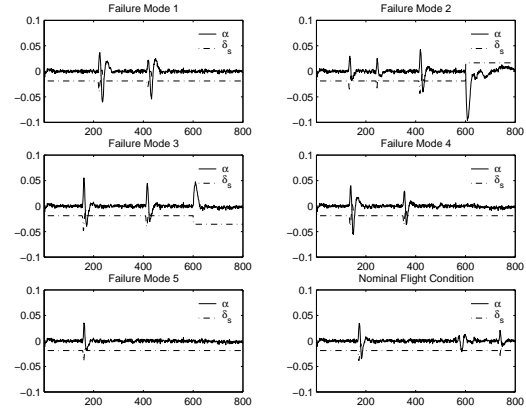


Figure 4: Data sequences of IFCS simulation under five failure modes and the regular mode.

Mode 4 Actuator Failure - stuck right aileron at predefined deflection (+3 degree)

Mode 5 Actuator Failure - stuck right aileron at predefined deflection (-3 degree).

For each mode, we select the following three pairs of highly correlated parameters to test our approach.

Pair 1 (q the pitch rate, δ_s the average of the Left Stabilator and the Right Stabilator)

Pair 2 (α the angle of attack, δ_s the average of the Left Stabilator and the Right Stabilator)

Pair 3 (θ the pitch attitude, δ_s the average of the Left Stabilator and the Right Stabilator)

Figure 3 shows the test results using parameter pair 3 for each mode. The crosses represent the testing data points while the circles represent the learned data points. The closed line surrounding the data points is the decision boundary obtained through the SVDD learning on the nominal flight data. The time history of the parameter pair under test, in each of the five test modes as well as the nominal case is plotted in Figure 4. Figure 5 shows the detected outliers for each failure mode as well as the nominal case corresponding to the time history. For those data points whose value equals one, they are considered outliers detected by the SVDD method, which can be referred to as the ones falling outside the boundary in Figure 3. Similarly, the data points with value zero in Figure 5 fall within the boundaries described in Figure 3. Comparing Figure 4 and Figure 5, we are able to see that for each spike that reflects abrupt changes in the data sequence, our method is sensitive to such changes and has the capability of detecting suspicious novelties. Please note that x axis represents

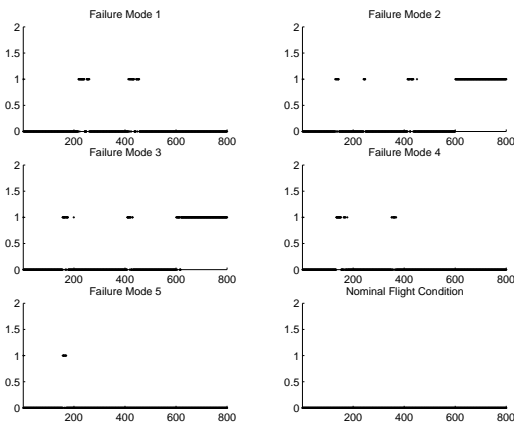


Figure 5: Detected novelties (Value 1) of IFCS simulation under five failure modes and the nominal mode.

actual time (20 frames per second) in all graphs of Figures 4 and 5.

From Figure 3 we can see that although some of the data points still fall inside the decision boundary, a large portion of data points that fall outside are considered as outliers. Our SVDD detectors show very effective outlier (failure mode) detection capabilities. Similar results are obtained from experiments on the other two pairs of data.

5 Conclusion

This paper proposes a novel approach to validating on-line adaptive systems using the support vector data description method. We offer a validation framework comprising of two filters that check the validity of inputs and classifications. Empirical results are obtained from running the SVDD algorithm on a sequence of simulation data for an actual flight control system. We observe that this method provides us a very good technique for separating nominal flight conditions from those representing possible failures. Its computational efficiency gives us the confidence to continue pursuing SVDD for online monitoring of the adaptive flight control system.

In the future, we expect to be able to test our method in an online (real-time) fashion. Meanwhile, statistical methods for determining the cross correlations among different dimensions will be investigated. We believe with the assistance of such methods, more accurate and sensitive detections can be made. However, more experiments need to be conducted to verify the domain specific performance of the novelty detectors implemented using the SVDD algorithm. Thus, we will focus on finding appropriate thresholds in further experiments as more data becomes available.

Acknowledgements

The authors acknowledge the partial support for this research received from NASA through cooperative agreement NCC 2-979. The opinions, findings, conclusions and recommendations expressed herein are those of the authors and do not necessarily reflect the views of the sponsor. We thank Dr. Tax and Dr. Duin for making the SVDD tools (PRTools and dd_tools) available.

References

- [1] S. Yerramalla, B. Cukic and E. Fuller. Lyapunov Stability Analysis of Quantization Error for DCS Neural Networks. *Int'l Joint Conference on Neural Networks*, July 2003.
- [2] V. N. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
- [3] D.M.J. Tax and R.P.W. Duin, Outliers and data descriptions, *Proc. ASCI 2001, 7th Annual Conf. of the Advanced School for Computing and Imaging*, ASCI, Delft, 2001, 234-241.
- [4] A. Mili, B. Cukic, Y. Liu, and R. Ben Ayed. Towards the Verification and Validation of On-Line Learning Adaptive Systems, In *Computational Methods in Software Engineering*. Kluwer Scientific Publishing, 2003.
- [5] J.A. Leonard, M.A. Kramer, and L.H. Ungar. Using radial basis functions to approximate a function and its error bounds. *IEEE Transactions on Neural Networks*, July 1992.
- [6] Roberts S.J. Extreme Value Statistics for Novelty Detection in Biomedical Signal Processing, *Proceedings of IEE: Science, Technology & Measurement*, Vol 147, No 6, pp 363-367.
- [7] C. Bishop. Novelty detection and neural network validation. *IEE Proceedings: Vision, Image and Signal Processing*, 141(4):217-222, 1994.
- [8] B. Scholkopf, R.C. Williamson, A.J. Smola, J. Shawe-Taylor, and J. Platt. Support vector method for novelty detection. In *Neural Information Processing Systems*, 2000.
- [9] S.J. Roberts, W.D. Penny, D. Pillot. Novelty, Confidence & Errors in Connectionist Systems. *Proceedings of IEE Colloquium on Intelligent Sensors and Fault Detection*, September 1996.
- [10] D.M.J. Tax, A. Ypma, and R.P.W. Duin, Support vector data description applied to machine vibration analysis, *Proc. 5th Annual Conference of the Advanced School for Computing and Imaging*, June 1999, 398-405.
- [11] Perhinschi M.G., Campa G., Napolitano M.R., A Simulation Tool for On-Line Real Time Parameter Identification, *Proceedings of the 2002 AIAA Modeling and Simulation Conference*, Monterey, August 2002.