Management Science and Financial Engineering
Vol 18, No 1, May 2012, pp.49-53
ISSN 2287-2043 | EISSN 2287-2361 |

Estimating Basin of Attraction for Multi-Basin Processes Using Support Vector Machine

Daewon Lee*
School of Industrial Engineering, University of Ulsan
Jaewook Lee
Department of Industrial Engineering, Seoul National University

(Received: April 2, 2012 / Accepted: May 7, 2012)

ABSTRACT

A novel method of transient stability analysis is presented in this paper. The proposed method extracts data points near the basin-of-attraction boundary and then builds a support vector machine (SVM) model learned from the generated data. The constructed SVM classifier has been shown to reduce dramatically the conservativeness of the estimated basin of attraction.

Keywords: Basin-of-Attraction Boundary, Support Vector Machines, Closest Unstable Equilibrium Points, Nonlinear Dynamical System, Power Systems, Transient Stability Analysis

* Corresponding Author, E-mail: dwlee0402@ulsan.ac.kr

1. INTRODUCTION

Multi-basin process is one of the widely used dynamical systems that have plenty of application areas, including machine learning, electrical engineering, industrial engineering, applied mathematics, and so on (Chiang *et al*., 1995; Jardim *et al*., 2000; Jung *et al*., 2010; Jung *et al*., 2011; Kim and Lee, 2007; Lee, 2003; Lee and Chiang, 2004; Lee and Lee, 2005; Lee, 2005; Lee and Lee, 2006; Lee and Lee, 2006; Lee *et al*., 2009; Lee and Lee, 2010). In particular, let us consider a nonlinear dynamic system described by

$$dz(t)/dt = F(z(t)) \tag{1}$$

where F: $R^n \rightarrow R^n$ are continuously differentiable for the existence of solutions. We will call the process in Eq. (1) as a multi-basin process if it decomposes the whole phase space into multiple basins of attractions, where each basin is represented by its representative stable equilibrium point, i.e.,

$$\mathrm{R}^n = \bigcup_i \; cl(A(s_i)) \tag{2}$$

Here, s_i is the stable equilibrium point, and $A(s_i)$, called the basin of attraction of s_i, is the set of all points that converges to s_i when Eq. (1) is applied. Quite a large class of Eq. (1) under some mild conditions can be shown to be multi-basin processes (Chiang *et al*., 1995; Jung *et al*., 2010; Lee and Chiang, 2004; Lee and Lee, 2005; Lee and Lee, 2006).

One of the key issues in the applications of multi-basin processes is the estimation of the basin of attraction for a given stable equilibrium points. The closest unstable equilibrium point (C-UEP) method can determine the optimal critical level value of the Lyapunov energy function to find a larger region within the basin of attraction and is one of the most popular methods adopted for transient stability analysis (TSA),an important tool in planning and operating power and control systems to predict the system response to various disturbances (Chiang *et al*., 1995; Jung *et al*., 2010; Jung *et al*., 2011). However, it has some practical limitations. First, the estimated region is very conservative, which can lead to serious TSA consequences because conservative estimates of the basin of attraction could result in unnecessary interruptions in the operation of power systems and in expensive overdesign of control systems. Second, by only comparing the current energy values without numerically integrating the point, the C-UEP method cannot guarantee that the current operating point is within the basin of attraction of a desired stable equilibrium point. These handicaps constitute the main obstacles of the C-UEP method to perform reliable and efficient actions, especially for TSA, in due time.

Recently, the pattern recognition approach has come up as an alternative approach to overcome these drawbacks. In this paper, we employ support vector (SV) machines (SVMs) to estimate a basin of attraction. Considering that SVM is originally designed for binary classifications and the verification of whether a point is inside a basin of attraction of interest or not can be considered as a binary classification problem, SVM is a good choice as an assessment model for TSA. The basic idea is as follows: first, the original space is transformed into a reduced space of lower dimension, hopefully preserving the original phase space information. Then, a set of points lying on the basin-of-attraction boundary is generated using the C-UEP method, and a binary class dataset from the set is extracted after some perturbations, where the dataset is composed only of potential SVs. Finally, an SVM classifier is trained by the extracted dataset. Considering that the dataset is extracted near the exact basin-of-attraction boundary, the conservativeness of the estimated basin of attraction can be reduced dramatically, and a fast reliable assessment for TSA is enabled or the stability margin can be predicted.

2. DESCRIPTION OF SVMS

In this section, the SVMs to be employed as assessment model for an online TSA in the proposed method are briefly reviewed. SVMs are primarily designed to minimize the upper bound of the expected error by optimizing the trade-off between the empirical risk (or training error) and model complexity captured by the VC dimension (Vapnik, 1999). Let us consider a training dataset $=\{x_i, y_i\}_{i=1}^N$, where $x_i \in \{-1, +1\}$ and $y_i \in \{-1, +1\}$, and consider a nonlinear transformation to a higher dimensional feature space $\varphi(\cdot): R^d \to R^H$. To minimize the expected error, SVMs can be formulated into the following quadratic optimization problem:

$$\begin{aligned} \min_{w,b,\varepsilon} \quad & \tfrac{1}{2}\|w\|^2 + C\sum_{i=1}^N \varepsilon_i \\ \text{s.t.} \quad & y_i\left(\left\langle w^T, \varphi(x_i)\right\rangle + b\right) \ge 1-\varepsilon_i, \ \forall i \\ & \varepsilon_i > 0, \ \forall i \end{aligned} \tag{3}$$

where w and b define the linear classifier in the feature space and the regularization parameter $C > 0$ determines the trade-off between the empirical error and the model complexity, leading to the dual problem

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^N \alpha_i - \tfrac{1}{2}\sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{s.t.} \quad & 0 \le \alpha_i \le C, \ \forall i \\ & \sum_{i=1}^N \alpha_i y_i = 1 \end{aligned} \tag{4}$$

where $a_i, i = 1, \cdots, N$ are the Lagrange multipliers and $K(a, b) = \varphi(a)^T \cdot \varphi(b)$ is the Mercer's inner product kernel. From the dual problem perspective, transform $\varphi(\cdot)$ does need not be specified at all but requires only knowledge of the kernel function that computes the inner products in the feature space. In this paper, we used the most widely used Gaussian RBF kernel function as follows:

$$K(a, b) = \exp(-\|a-b\|^2/\sigma). \tag{5}$$

From the solution of the quadratic optimization problem (Eq. (4)) $\hat{\alpha}_i, i = 1, \cdots, N,$ the optimal hyperplane $\hat{f}(x)$ can be constructed as follows:

$$\begin{aligned} \hat{f}(x) &= \hat{W}^T \cdot \varphi(x) + \hat{b} = \sum_{i=1}^N \hat{\alpha}_i y_i K(\mathrm{x}, x_i) + \hat{b} \\ \hat{b} &= y_{SV} - \sum_{i=1}^N \hat{\alpha}_i y_i K(x_i, x_{SV}) \end{aligned} \tag{6}$$

where (x_{SV}, y_{SV}) is any SV whose output class label satisfies $\alpha_{SV} > 0$.

3. PROPOSED METHOD

The proposed method is based on the distinguishing feature of an SVM, where only SVs are used to determine the optimal separating hyperplane (Vapnik, 1999). If only the training data are extracted in advance on the marginal boundary near the decision boundary, then an SVM model can be efficiently optimized with fewer training patterns, showing a decision boundary similar to that of large datasets containing many points outside the marginal boundary.

3.1 Phase I: Reducing the Dimension of the Phase Space

Given a nonlinear dynamic system described by Eq. (1), we first transform Eq. (1) into the following equivalent reduced system:

$$dx(t)/dt = G(x(t)) \tag{7}$$

where G: $R^m \rightarrow R^m$ is continuously differentiable with $m < n$.
Here, one essential requirement of this transformation is that, for each trajectory of Eq. (1), a trajectory of Eq. (7) should exist. One feasible condition for this situation is when the following relationship holds:

$$F(z) = AG(A^TA)^{-1}A^Tz) \tag{8}$$

where A is an n-by-m matrix of full rank m. Using transformation $z = Ax$, each trajectory of Eq. (1) can easily show that such $F(z)$ corresponds to a trajectory of Eq. (7). The main purpose of Phase I is to obtain a reliable set of sample points for the next Phase II because sampling in a lower dimensional space is easy.

3.2 Phase II: Extracting a Dataset Near the Basin-of-attraction boundary $\partial A(x_s)$

To identify the basin-of-attraction boundary $\partial A(x_s)$, the following result is adopted: under some mild conditions (Jardim *et al.*, 2000).

$$\partial\overline{\mathrm{A}(\mathrm{x_s})} = \bigcup_{i=1}^{N} \overline{W^s(x_1^i)} \tag{9}$$

where x_1^i, $i = 1, 2, \cdots, N$ are type-1 UEPs in the stability boundary and $W^s(x_1^i)$ is the stable manifold of x_1^i. ($\overline{A}$ denotes the closure of set A). The points on the marginal boundary near $\partial A(x_s)$ can be numerically obtained as follows:

Step 1: The type-1 equilibrium points on $\partial A(x_s)$, whose unstable manifolds contain trajectories approaches x_s., are located (Lee, 2003).
Step 2: For each type-1 equilibrium point (e.g., $\hat{x}$), its stable manifold is determined as follows:

i) A normalized stable eigenvector v of the Jacobian at $\hat{x}$ is determined.
ii) The intersection of this stable eigenvector with the boundary of an ε -ball of $\hat{x}$ is found (The intersection points are $\hat{x} + \varepsilon v$ and $\hat{x} - \varepsilon v$).
iii) The vector field from each of these intersection points is integrated after some specified time. If the trajectory remains inside this ε -ball, then the next step is processed; otherwise, the value of ε is replaced by $\alpha\varepsilon$, and the intersection points $\hat{x} \pm \varepsilon v$ are replaced by $\hat{x} \pm \alpha\varepsilon v$, where $0 < \alpha < 1$; this step is repeated.
iv) The vector field is numerically integrated backward (reverse time) starting from these intersection points.
v) The resulting trajectories are the stable manifold of $\hat{x}$.

For type-1 equilibrium point x_1^*, we extract the data points at uniform intervals from $W^s(x_1^*)$ (bold solid line), determined by Step 2, where four points are extracted. Next, to make a binary class dataset, we shift all the extracted points in the direction of stable equilibrium point x_s and also in the opposite direction as small intervals r_1 and r_2. Then, we assign a class label to each shifted point by numerically integrating the shifted point and checking whether or not it converges to x_s, that is, if it converges to x_s, then, a stable class label is assigned (denoted as-1); otherwise, unstable class label is assigned (denoted as+1). By repeating the above process for all type-1 equilibrium points, we can obtain a binary class dataset near $\partial A(\mathrm{x_s})$. Readers are referred to (Lee and Lee, 2006) for some illustrative examples.

3.3 Phase III: Constructing an SVM Model for Estimating the Basin of Attraction $A(x_s)$

The obtained optimal separating hyperplane can be considered as the estimated basin-of-attraction boundary (that is, $\hat{f}(x) = 0$). From this function, we can determine whether or not the current fault on the trajectory is inside $A(x_s)$, without using any numerical integration but by only checking if $\hat{f}(x) < 0$ (stable) or $\hat{f}(x) > 0$ (unstable). The constructed SVM classifier has the following desirable properties: first, it dramatically reduces the conservativeness of the estimated basin of attraction compared with those of the existing direct method or other neural network techniques, thereby enabling accurate assessment of TSA using the estimated optimal separating hyperplane. Second, applying labels to the output class for each extracted point whether it converges to x_s or not is time-consuming, and the computational burden for SVM quadratic optimization Increases proportionally with data size. Hence, reducing the data size without compensating for the generalization performance is very important. From this point of view, our scheme of extracting a dataset from potential SVs near $\partial A(x_s)$ helps overcome such computational complexity problems.

4. EXPERIMENTAL RESULTS

To illustrate the proposed method, some simulations were performed on the 3-MPS-i and 3-MPS-iisystems, considered in (Chiang *et al.*, 1995) and (Jung *et al.*, 2010), respectively. Phase-I has been applied to the original 4-D systems, and 2-D reduced gradient systems of three-machine power systems were obtained using machine number three as the reference. To demonstrate empirically the performance of the proposed method, we compared the accuracy of the proposed method with those of other existing methods on the above power systems. The performance comparison was conducted based on two factors: (i) assessment model (SVM, multilayer perceptron (MLP), and the C-UEP method) and (ii) extraction of a dataset (proposed method and random generation method). Taking these items into considerations, the proposed method was compared with the three other methods. In Table 1, C-UEP and Proposed represent the C-UEP and the proposed methods, respectively. R-MLP (random sampling+MLP) and R-SVM (random sampling +MLP) represent the MLP in (Jung *et al.*, 2010) and the SVM model in (Jung *et al.*, 2011), respectively. For fair comparison, the size of the randomly generated dataset was made equal to that of the dataset of the proposed method.

Table 1. Error rate (%) on the Test Dataset. F. Alarm and F. Dismissal Represent the False Alarm and False Dismissal Rates, Respectively

	3-MPS-i		3-MPS-ii	
	F. Alarm	F. Dismissal	F. Alarm	F. Dismissal
C-UEP	18.0	6.0	17.5	32.5
R-MLP	2.0	16.0	2.5	10.0
R-SVM	2.0	10.0	2.5	20.0
Proposed	**2.0**	**0.0**	**0.0**	**0.0**

The training and test data sizes are 70 and 50on 3-MPS-i, and 50 and 40 on 3-MPS-ii, respectively. The same test dataset was used for the performance comparison of the four methods. The criteria were the false dismissal rate (the rate of unstable points classified as stable class) and the false alarm rate (the rate of stable points classified as unstable class) on the test dataset. The parameter tuning of the proposed method was relatively easier than those of the other existing methods. The shift amount parameters r_1 and r_2 were determined according to our desired conservativeness. If we want to avoid an unnecessary interruption (false alarm), then we set r_1 to a smaller value. As a result, larger estimated basins of attraction are obtained. On the other hand, if we want to avoid false dismissals, then we make the basin of attraction more conservative (on the extreme, a subset of $A(x_s)$). To accomplish this objective, we set r_2 to a smaller value and r_1 to a relatively larger value. As shown in Figure 1(a) and Figure 1(b), the estimated basin of attraction of the proposed method is nearly the same as that of the true basin of attraction $A(x_s)$. The C-UEP method has a very conservative smaller estimate, and the R-MLP and R-SVM methods have larger estimates, in addition to suffering from non-zero false dismissal rate problem by incorrectly classifying unstable states as stable state, which is often worse than the direct methods and resulting in very large operating costs. The simulation results demonstrated that the proposed method dramatically reduces not only the computational burden but also the conservativeness of the estimated basin of attraction.

5. CONCLUSIONS

In this paper, a new hybrid method for TSA has been developed. The proposed method first extracts the data points from the union of stable manifolds of type-1 equilibrium points on the basin-of-attraction boundary and obtains the binary class dataset after some perturbation of the extracted points in which the SVM model is trained. The extracted dataset is designed to compared with the other be distributed near the basin-of-attraction boundary so that it assessment for TSA. Benchmark results demonstrated that the proposed method is computationally efficient and reliable becomes similar to SVs for the constructed decision boundary. Therefore, the proposed method dramatically reduces not only the computational burden but also the conservativeness of the estimated basin of attraction. Ultimately, it enables an online existing technique. Application of the proposed method to more large-scale and high-dimensional power system problems has yet to be investigated.

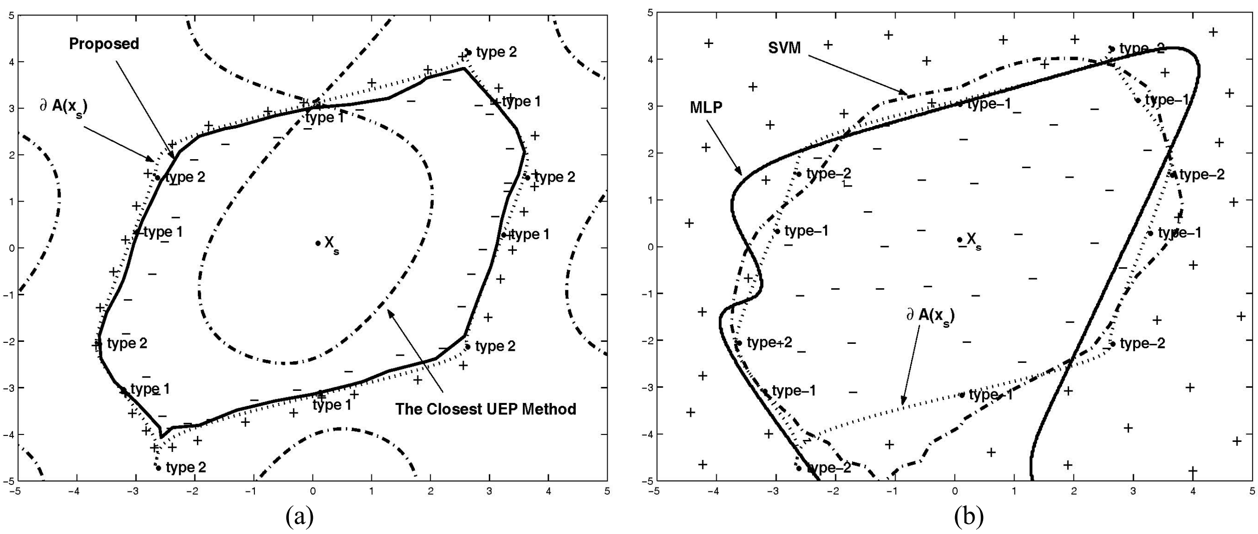

Figure 1. Illustration of the Estimated Basins of Attraction. The normal dotted line is the exact basin-of-attraction boundary $\partial A(x_s)$. (a) The bold solid line and the bold dashed-dotted line are the estimates of the proposed method and the C-UEP, respectively, where the symbols '-' (stable) and '+' (unstable) denote the dataset extracted by the proposed method. (b) The bold solid line and bold dashed–dotted line are the estimates of R-MLP and R-SVM, respectively, where the plotted dataset is randomly generated for training of these methods

ACKNOWLEDGMENTS

This work was supported by the 2010 Research Fund of the University of Ulsan. Part of the results of this paper was presented in the WCCI 2006 Conference.

REFERENCES

Chiang, H. D., C.-C. Chu, and G. Cauley, "Direct stability analysis of electric power systems using energy functions: Theory, applications, and perspective," *Proc. IEEE* 83 (1995), 1497-1529.

Jardim, J. L., C. A. da S. Neto, A. P. A. da Silva, A. C. Zambronide Souza, D. M. Falc˜ao, C. L. T. Borges, and G. N. Taranto, "A unified online security assessment system," *In Proc. CIGRE*, Paris, France, 2000.

Jung, K.-H., D. Lee, and J. Lee, "Fast Support-based Clustering for Large-scale Problems," *Pattern Recognition* 43 (2010), 1975-1983.

Jung, K.-H., N. Kim, and J. Lee, "Dynamic pattern denoising method using multi-basin system with kernels," *Pattern Recognition* 44, 8 (2011), 1698-1707.

Kim, H.-C. and J. Lee, "Clustering Based on Gaussian Processes," *Neural Computation* 19, 11 (2007), 3088-3107.

Lee, J., "Dynamic gradient approaches to compute the closest unstable equilibrium points for basin of attraction estimate and their computational limitations," *IEEE Trans. on Automatic Control*, 48, 2 (2003), 321-324.

Lee, J. and H.-D. Chiang, "A Dynamical Trajectory-Based Methodology for Systematically Computing Multiple Optimal Solutions of General Nonlinear Programming Problems," *IEEE Trans. on Automatic Control* 49, 6 (2004), 888-899.

Lee, J. and D. Lee, "An Improved Cluster Labeling Method for Support Vector Clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence* 27, 3 (2005), 461- 464.

Lee, J., "Pseudo-Basin of Attraction for Combinatorial Dynamical Systems: Theory and Its Application to Combinatorial Optimization," *IEEE Trans. on Circuits and Systems-Part II* 52, 4 (2005), 189-193.

Lee, D. and J. Lee, "Support Vector Classifier Using Basin-Based Sampling for Security Assessment of Nonlinear Power and Control System," *IEEE International Joint Conference on Neural Networks*, 2006.

Lee, J. and D. Lee, "Dynamic Characterization of Cluster Structures for Robust and Inductive Support Vector Clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence* 28, 11 (2006), 1869-1874.

Lee, D., K.-H. Jung, and J. Lee, "Constructing Sparse Kernel Machines Using Attractors," *IEEE Trans. on Neural Networks* 20, 4 (2009), 721-729.

Lee, D. and J. Lee, "Dynamic Dissimilarity Measure for Support-based Clustering," *IEEE Trans. on Knowledge and Data Engineering* 22, 6 (2010), 900-905.

Moulin, L. S., A. P. Alves da Silva, M. A. El-Sharkawi, and R. J. MarksII, "Support vector machines for transient stability analysis of large-scale systems," *IEEE Trans. on Power Systems* 19 (2004), 818-825.

Vapnik, V. N., "An overview of statistical learning theory," *IEEE Trans. on Neural Networks* 10 (1999), 988-999.