

# Utilization of Support Vector Machine Classifiers to Power System Topology Verification

Robert Lukomski

Electrical Power Engineering Institute  
Wroclaw University of Technology  
Wroclaw, Poland  
robert.lukomski@pwr.wroc.pl

Kazimierz Wilkosz

Electrical Power Engineering Institute  
Wroclaw University of Technology  
Wroclaw, Poland  
kazimierz.wilkosz@pwr.wroc.pl

**Abstract**—Creating a correct power network connectivity model is very important stage of power system real-time modeling. Errors in this model may seriously degrade credibility of results of power system applications and implicate improper and potentially danger control actions. In this situation, checking correctness of power network topology model is of great importance. The aim of the paper is to present an original method for power system topology verification, using the knowledge resulting from relationships describing power system and Support Vector Machine classification technique. In the paper, theoretical background and principle of the proposed method are described. A case study shows utilization of the method for the verification of a topology model of the IEEE 14-bus test system. At the end of the paper, features of the described method are discussed.

**Keywords** – power system, topology, topology verification, state estimation, SVM

## I. INTRODUCTION

The credible information on power network connectivity and analog quantities are of great importance from viewpoint of real-time modeling of a power system. A power network topology model is created by processing of reported ON/OFF statuses of circuit breakers. Occurrence of errors in a topology model is very rare, however consequence of such errors is very serious.

The common topology errors (TEs) in bus-branch topology model are: (i) improper modeling of branch status (branch is considered as open but in fact it is in operation, branch is considered as closed but in fact it is out of operation), (ii) bus splitting, (iii) bus merging, (iv) bus reconfiguration. TEs are known to significantly affect the performance of real-time modeling of a power system. Necessity of identification of TEs in a power network connectivity model results in developing of numerous methods dealing with this problem. The studies have revealed that the TE identification is not trivial problem because of discrete nature breaker statuses and relatively complex relationships among analog quantities measured in a power system (active and reactive power flows and injections, voltage magnitudes).

The methods for topology verification (which allows detecting and identifying TEs) can be divided into two main groups: (i) *post-estimation* methods, (ii) *pre-estimation*

methods. In the case of the first group, verification is processed with the use of state estimation results. The verification methods of the second group are utilized before running state estimation and with use of raw measurement data.

Some methods belonging to first group use the measurement residuals obtained from state estimation [1], [2]. Modeling of circuit breakers statuses at bus section level by using extended state variable set to estimation is proposed in [3]-[7]. TE identification formulated as optimization problem with some operational and structural constraints is described in [8]-[10]. An iteratively re-weighted least square estimation for active and reactive power is developed in [11]. Utilization of robust M-Huber estimator enables detection of TEs by use of estimated branch power flows. An interval analysis and interval constraints propagation to handle TEs are proposed in [12].

Use of pre-estimation verification concept is presented in [13]. Rule based system for identification of connectivity errors is proposed in [14]. Power system topology verification is also recognized as pattern classification problem and artificial neural network technique is widely used [15]- [20].

It should be underlined that in [20] to solve the verification problem, additional knowledge on a power system is applied. As an effect, the learning of utilized artificial neural networks as well as classification process become more efficient and the verification method is more reliable. Additional knowledge on a power system is also utilized in the case of the method presented in this paper. To efficiently deal with single and multiple TEs the SVMs classifiers [21], [22] are incorporated into the verification method. SVMs are recognized as very useful and universal tool for classification and regression tasks. In distinction to artificial neural networks, learning of the SVMs relies on solution of constrained quadratic programming problem.

In the paper, SVM binary classifiers are proposed to classify preprocessed so-called unbalance indices [20] in order to detect bad modeling of power network branches in topology model. Architecture, parameter tuning and performance testing of SVM are described.

## II. TOPOLOGY VERIFICATION AS CLASSIFICATION PROBLEM

Aim of topology verification with respect to a separate branch is an ascertainment that a branch is correctly modeled or that modeling of a branch is incorrect. In other word, we can ascertain that each branch should be classily as a member of the class of correctly modeled branches or a member of the other class being the class of incorrectly modeled branches.

### A. Power System Model

The used power system model is based on the assumption that every branch in a power network (i.e. a power line, a transformer) is modeled as the  $\pi$ -equivalent circuit (Fig. 1). It is assumed that there is an accessible and credible measurement data set of such quantities as active and reactive power flows at the ends of each branch, power injections, loads and voltage magnitudes at each node. Usually, if a branch is not included in a power system model the measurement data related to the branch are considered as irrelevant.

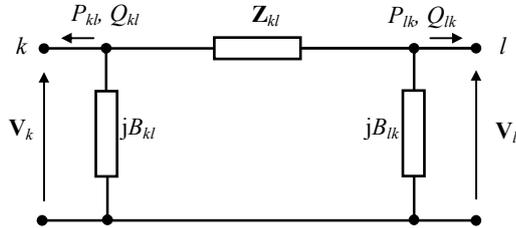


Figure 1. The assumed  $\pi$  model of the branch.

$$\mathbf{Z}_{kl} = R_{kl} + jX_{kl}, B_{kl} = B_{lk} = B.$$

$B$  is a half of the capacitive susceptance of the branch.

### B. Unbalance Indices

A power system can be described by many relationships among measured quantities. Relationships among these quantities are determined by the Kirchhoff's and Ohm's Laws. In this paper the following relationships among measured quantities in a power system are taken into consideration:

- a) For a node of a power system,

$$\sum_{i \in I_k} P_{ki} = 0, \quad (1)$$

$$\sum_{i \in I_k} Q_{ki} = 0, \quad (2)$$

where:  $P_{ki}$ ,  $Q_{ki}$  – respectively active and reactive power flows at the node  $k$  on the branch connecting the nodes  $k$  and  $i$ ,  $I_k$  – the set of nodes connected with the node  $k$ .

- b) For a branch of a power system,

$$P_{kl} + P_{lk} + R_{kl}W_{xy} = 0, \quad (3)$$

$$Q_{kl} + Q_{lk} + X_{kl}W_{xy} - W_{xy}^V = 0, \quad (4)$$

where:  $x = k$  and  $y = l$  or  $x = l$  and  $y = k$ ,

$$W_{xy} = \frac{P_{xy}^2 + (Q_{xy} - B_{kl}V_x^2)^2}{V_x^2},$$

$$W_{xy}^V = B_{kl}(V_x^2 + V_y^2),$$

$V_k, V_l$  –voltage magnitudes at the nodes  $k$  and  $l$  respectively;  $R_{kl}, X_{kl}, B_{kl}$  – parameters of the  $\pi$  model of the branch (Fig. 1).

If there are no TEs, all these relationships are fulfilled. When TE occurs some of the relationships become unfulfilled. It should be underlined that if a branch is not included in the power system model, the relationships for this branch are not considered, because measurement data for it are not taken into account.

The balances of active and reactive powers at the terminal nodes of the branch  $k-l$  in the model for different cases of modeling the branch  $k-l$  are shown in Tab. I, when at the most the branch  $k-l$  is incorrectly modeled.

If power flow balances at the nodes, between which there is a branch, instead of the power flow measurement data on the branch are utilized for testing fulfillment of the relationships (3) and (4), then the left sides of these relationships for different possible cases under assumption that at the most the branch  $k-l$  is incorrectly modeled are as it is shown in Tab. II.

TABLE I. THE BALANCES OF ACTIVE AND REACTIVE POWERS AT THE TERMINAL NODES OF A BRANCH  $k-l$  IN THE MODEL, WHEN AT THE MOST THE BRANCH  $k-l$  IS INCORRECTLY MODELED

		A state of the branch $k-l$ in a system	
		ON	OFF
A state of the branch $k-l$ in the model	ON	$W_{Pk} = 0, W_{Qk} = 0,$ $W_{Pl} = 0, W_{Ql} = 0$	$W_{Pk} = 0, W_{Qk} = 0,$ $W_{Pl} = 0, W_{Ql} = 0$
	OFF	$W_{Pk} = -P_{kl},$ $W_{Qk} = -Q_{kl}$ $W_{Pl} = -P_{lk},$ $W_{Ql} = -Q_{lk}$	$W_{Pk} = 0, W_{Qk} = 0,$ $W_{Pl} = 0, W_{Ql} = 0$
ON – for the system: the branch is in operation; for the model: the branch is included in the model, OFF – for the system: the branch is out of operation; for the model: the branch is not included in the model, $W_{Px}, W_{Qx}$ – the left sides of relationships (1) and (2) for the node $x$ respectively, $x \in \{k, l\}$ .			

In the described approach to have possibility of examination of relationships for all nodes and all branches independently of their correct or incorrect inclusion in a power system model the so-called unbalance indices for nodes and branches are introduced.

TABLE II. THE BALANCES OF ACTIVE AND REACTIVE POWERS FOR THE BRANCH K-L IN THE MODEL, WHEN AT THE MOST THE BRANCH K-L IS INCORRECTLY MODELED.

		A state of the branch $k-l$ in a system	
		ON	OFF
A state of the branch $k-l$ in the model	ON	$W_{Pxy} = R_{kl}W_{xy}^0$ $W_{Qxy} = X_{kl}W_{xy}^0 - W_{xy}^V$	$W_{Pxy} = R_{kl}W_{xy}^0$ $W_{Qxy} = X_{kl}W_{xy}^0 - W_{xy}^V$
	OFF	$W_{Pxy} = 0$ $W_{Qxy} = 0$	$W_{Pxy} = R_{kl}W_{xy}^0$ $W_{Qxy} = X_{kl}W_{xy}^0 - W_{xy}^V$
$W_{Pxy}, W_{Qxy}$ – the left sides of relationships (3) and (4) for the branch $k-l$ respectively, $W_{xy}^0 = \frac{B_{kl}^2 V_x^4}{V_x^2}$ .			

Unbalance indices are the left-hand sides of the appropriate relationships, considered in the form in which their right-hand sides are equal to zero and are defined as follows:

$$W_{Pk} = \sum_{i \in I_k} P_{ki}, \quad (5)$$

$$W_{Qk} = \sum_{i \in I_k} Q_{ki}, \quad (6)$$

$$W_{Pxy} = -W_{Px} - W_{Py} + R_{lk}W_{xy}^u, \quad (7)$$

$$W_{Qxy} = -W_{Qx} - W_{Qy} + X_{kl}W_{xy}^u - W_{xy}^V, \quad (8)$$

where:

$$W_{xy}^u = \frac{W_{Px}^2 + (W_{Qx} + B_{kl}V_x^2)^2}{V_x^2}.$$

It should be noted that the nodal unbalance indices instead of power flow measurement data are taken into account when branch unbalance indices are calculated. This fact enables considering branch unbalance indices independently of correct or incorrect inclusion of branches in the power system model.

Unbalance indices create characteristic sets of values for different cases of modeling a power system. If the topology model is correct and there are no errors burdening measurement data, all nodal unbalance indices are equal to zero and branch unbalance indices are near to zero, as well. The same situation is, when there is a branch that is actually out of operation but it is included in the topology model (the inclusion error). If a branch is actually in operation in a power system but it is not included in the topology model (the exclusion error), then:

a) the unbalance indices for terminal nodes of this branch considerably differ from zero,

b) the unbalance indices for the considered branch are equal to zero,

c) absolute values of the unbalance indices for other branches, that are incident to the nodes mentioned under point a), have especially large values.

It should be stressed that the behavior of unbalance indices for active power and for reactive power is generally the same for the same TE.

Analyzing unbalance indices for nodes and branches one can observe that the exclusion error of the branch  $j$  has no influence on:

- unbalance indices for nodes, that are not terminal nodes of the branch  $j$ ,
- unbalance indices for branches that are not incident to the terminal nodes of the branch  $j$ .

This observation shows existence of the local effect of TE. In this situation one can conclude about correctness of modeling the distinguished branch  $j$  on the basis of investigations of unbalance indices for certain areas of the power network:  $A_j^k, A_j^l$ , where:  $k, l$  are numbers of the terminal nodes of the branch  $j$ .  $A_j^x, x \in \{k, l\}$  is the area, in which the branch  $j$  exists with the central node  $x$ . The area  $A_j^x$  comprises:

- the node  $x$  (which is one of the terminal nodes of the branch  $j$ ),
- the branch  $j$  and all other branches incident to the node  $x$ ,
- all nodes which are connected with the node  $x$  by the earlier-mentioned branches.

An investigation of unbalance indices is much more complicated in the case of multiple TEs. Characteristic patterns, which are created by unbalance indices for different cases of correctness of modeling, overlap each other. However, using proper classification techniques, these patterns can be recognized and topology verification decisions can be taken.

### C. Principle of Topology Verification Method

The proposed topology verification method consists of the following steps:

- calculation of unbalance indices for nodes and branches,
- pre-processing of the unbalance indices (pre-processing standardization),
- local classification,
- global classification.

The pre-processing standardization of each unbalance index is realized using Radial Basis Function (RBF) unit with Gaussian transfer function:

$$f(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (9)$$

where:  $\sigma$  - width parameter.

The pre-processing standardisation allows keeping input values for RBF network in the range (0; 1].

The  $\sigma^2$  parameter for the nodal index  $W_{Pk}$   $k \in \{1, 2, \dots, n\}$  is calculated as follows (errors are assumed to be independent):

$$\sigma_{W_{Pk}}^2 = a \sum_{l \in I_k} \sigma_{P_{kl}}^2, \quad (10)$$

where:  $\sigma_{P_{kl}}$  - standard deviation of data of the active power flow  $P_{kl}$ ;  $a$  - correction coefficient selected in an experimental way.

The  $\sigma^2$  parameter for the branch unbalance index  $W_{Pkl}$   $k, l \in \{1, 2, \dots, n\}$  is given by:

$$\sigma_{W_{Pkl}}^2 = a(\sigma_{W_{Pk}}^2 + \sigma_{W_{Pl}}^2). \quad (11)$$

The  $\sigma^2$  parameters for unbalance indices for reactive power are calculated in the similar way. It was found that  $a=2$  for active power unbalance indices and  $a=1,8$  for reactive power unbalance indices.

For the branch  $k-l$ , two decision regarding correctness of modelling this branch can be taken. One decision is taken by the SVMs related to the area  $A_{kl}^k$ . The second decision is taken by the SVMs related to the area  $A_{kl}^l$ .

Local classification step is aimed at assignment of input patterns to one of the classes corresponding to the correctness of modeling branches in a power system. Each local classifier corresponds to one node of a considered power network and consists of  $m$  SVMs, where  $m$  is the number of branches connecting the node  $k$  with the nodes having numbers from the set  $I_k$ . Output of SVM is a base for taking a verification decision on correctness of modeling a branch with number belonging to  $I_k$ . For the node  $k$ , inputs for a local classifier corresponding this node are the results of the pre-processing of active and reactive power unbalance indices for: the node  $k$ , the nodes having numbers from the set  $I_k$ , each branch connecting the node  $k$  and the node  $l$ , under assumption  $l \in I_k$ . The criterion for taking a decision on correctness of modeling a branch is as follows.

$$D_{kl} = \begin{cases} \text{the branch } k-l \text{ is incorrectly modelled} & \text{when } Y_{kl} \leq -1 \\ \text{the neutral decision} & \text{when } Y_{kl} \in (-1, 1) \\ \text{the branch } k-l \text{ is correctly modelled} & \text{when } Y_{kl} \geq 1 \end{cases} \quad (12)$$

where:  $D_{kl}$  is a decision,  $Y_{kl}$  is an output value corresponding to the branch between the node  $k$  and the node  $l$ .

The global classification step produces final decisions on correctness of modeling of power network branches. It processes results of local classifications. To take the final decision on correctness of modeling a selected branch the outputs of two local classifiers are considered. These classifiers are assigned to the terminal nodes of the considered branch.

Decisions given by local classifiers may be neutral or in contradiction, therefore some simple decision rules are used. The final decision is the neutral one in the following cases: (i) decisions of local classifiers are different and none of them is the neutral decision, (ii) each of the local classifiers produces the neutral decision. In other cases, the final decision is different from the neutral one.

### III. IMPLEMENTING SUPPORT VECTOR MACHINES CLASSIFIERS FOR TOPOLOGY VERIFICATION

SVMs enable minimizing classification error and maximize the geometric margin between separated classes [21], [22]. This type of classifier maps input vector to a higher dimensional space where separating hyperplanes are constructed. It can be stated that the larger the margin between separating hyperplanes the better generalization performance of the classifiers.

Parameters of the SVM classifier are found in the supervised learning process. Each instance in the training set contains one output value corresponding to class label and many input patterns.

Determination of parameters of the SVM classifier for a training set containing pairs  $D = \{(\mathbf{x}_i, y_i)\}$ ,  $i = 1, 2, \dots, s$ , where  $\mathbf{x}_i \in \mathbf{R}^n$  and  $y_i \in \{-1, 1\}$ , requires solution of the following optimization problem:

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^s \xi_i, \quad (13)$$

$$\text{subject to } y_i (\mathbf{w}^T \Phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0,$$

where:  $\mathbf{w}$  - vector of weights,  $C$  - cost parameter,  $y_i$  - target values,  $b$  - bias,  $\xi_i$  - weakening variables,  $\Phi(\mathbf{x}_i)$  - function mapping the vector  $\mathbf{x}_i$ .

An essential element of the optimization process is choice of the kernel function. This function is defined as  $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi^T(\mathbf{x}_i)\Phi(\mathbf{x}_j)$ . The main kernel functions, which are usually used in practice, are as follows:

- linear kernel:  $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$ ,
- polynomial kernel:  $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0$ ,
- radial basis function (RBF) kernel:  
 $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \gamma > 0$ ,
- sigmoid kernel:  $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r)$ .

where:  $\gamma, d, r$  are kernel parameters.

Weights and bias are found by solution of constrained quadratic programming problem. Many quadratic-optimization algorithms are available and the selection of suitable optimization method depends usually on the size of the problem and computational burdens.

It should be noted that SVMs originally are proposed for binary classification. In the presented topology verification method the local classifier contains  $m$  SVMs with single binary output, where  $m$  is number of branches adjacent to the node

with corresponding local classifier. Each SVM output corresponds the one branch and gives the value being a base for taking decision on branch modeling correctness. For all SVM of one local classifier the input vector  $\mathbf{x}$  is the same (Fig. 2).

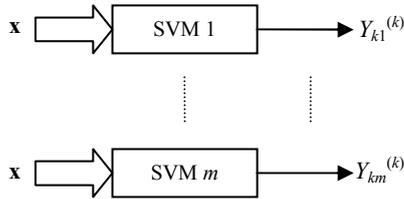


Figure 2. Structure of the local classifier for node  $k$ .  
 $\mathbf{x}$  - an input pattern;  $m$  - a number of adjacent branches to the considered node (a number of elements in the set  $I_k$ );  
 $Y_{ki}^{(k)}, i \in \{1, 2, \dots, m\}$  - a classifier output for the branch  $k-i$ .

#### IV. EXPERIMENTAL RESULTS

The presented method was implemented in the MATLAB environment. The method has been tested using the IEEE 14-bus test system (Fig. 3). It has been assumed that:

- all branches are actually in operation,
- single and multiple TEs are considered,
- measurement data are burdened with small errors,
- wide range of load curve changes is taken into account.

Learning sets containing several hundreds learning patterns were created separately for each SVM classifier. They comprise results of pre-processing unbalance indices and appropriate verification decisions. The test system contains 20 branches and therefore 40 SVMs (two classifiers belonging to terminal nodes of one branch) are learned. After training phase the performance test was carried out with use of patterns outside the training set. To assure the best performance linear and RBF kernel function with wide range of their parameters were tested. The obtained structures were presented in Tab. III. The number of support vectors determining the separation hypersurface (a number of hidden units) decides about the classifier complexity. One can be ascertain that for all cases the linear kernel function is suitable and the cost parameter  $C$  depends on the number of input variables -  $C$  is "compromise" between misclassification possibilities and complexity of a classifier.

The accuracy rate measured as probability of taking the neutral decision for the case of double TEs is shown in Tab. IV. It should be underlined that in the case of single TEs only correct classification were obtained. The misclassifications for both TE cases were not detected. Some doubtful cases have occurred and the neutral decisions have been taken for the branches with numbers 19 and 20. In these cases there is no possibility to state the correctness of the considered branch in the test system. A reason is relatively small level of power flows in the mentioned branches. The obtained results show that the efficiency of the SVM classifiers is relatively high.

TABLE III. STRUCTURE AND PARAMETERS OF LOCAL SVM CLASSIFIERS.

$k$	SVM parameters				
1	$\frac{g=1}{10-7}$ L(10)	$\frac{g=2}{10-7}$ L(10)			
2	$\frac{g=1}{18-11}$ L(10)	$\frac{g=3}{18-11}$ L(10)	$\frac{g=4}{18-12}$ L(10)	$\frac{g=5}{18-7}$ L(10)	
3	$\frac{g=3}{10-6}$ L(10)	$\frac{g=6}{10-6}$ L(10)			
4	$\frac{g=4}{22-14}$ L(15)	$\frac{g=6}{22-15}$ L(10)	$\frac{g=7}{22-13}$ L(15)	$\frac{g=8}{22-15}$ L(10)	$\frac{g=9}{22-18}$ L(10)
5	$\frac{g=2}{18-11}$ L(10)	$\frac{g=5}{18-15}$ L(10)	$\frac{g=7}{18-12}$ L(10)	$\frac{g=10}{18-12}$ L(10)	
6	$\frac{g=10}{18-11}$ L(10)	$\frac{g=11}{18-10}$ L(10)	$\frac{g=12}{18-9}$ L(10)	$\frac{g=13}{18-6}$ L(10)	
7	$\frac{g=8}{14-10}$ L(20)	$\frac{g=14}{14-9}$ L(20)	$\frac{g=15}{14-10}$ L(20)		
8	$\frac{g=14}{8-5}$ L(10)				
9	$\frac{g=9}{18-11}$ L(10)	$\frac{g=15}{18-12}$ L(15)	$\frac{g=16}{18-12}$ L(10)	$\frac{g=17}{18-12}$ L(10)	
10	$\frac{g=16}{10-7}$ L(10)	$\frac{g=18}{10-5}$ L(10)			
11	$\frac{g=11}{10-6}$ L(10)	$\frac{g=18}{10-6}$ L(10)			
12	$\frac{g=12}{10-6}$ L(10)	$\frac{g=19}{10-6}$ L(10)			
13	$\frac{g=13}{14-5}$ L(10)	$\frac{g=19}{14-6}$ L(10)	$\frac{g=20}{14-6}$ L(10)		
14	$\frac{g=17}{10-5}$ L(10)	$\frac{g=20}{10-6}$ L(10)			

$k$  - a node number;  $g$  - a branch number, to which a SVM classifier is assigned;  $A-B$  L( $C$ ) - SVM with  $A$  inputs,  $B$  hidden units, the linear (L) kernel function and with the cost parameter  $C$

TABLE IV. A PROBABILITY OF TAKING THE NEUTRAL DECISION  $P_N$  IN THE VERIFICATION PROCESS FOR THE BRANCHES OF THE TEST SYSTEM IN CASE OF DOUBLE TOPOLOGY ERRORS.

$g$	1 - 18	19	20
$P_n$	0	0,08	0,06

#### V. CONCLUSION

The proposed topology verification method with the use of prior knowledge on a power system and the SVM classifiers gives very promising results. The method uses raw measurement data and local effect of topology. The verification is performed by many local processes realized with the use of the classifiers assigned to the nodes of a power network. The used classification structure assigns two SVMs to one branch.

