# A linear approach to determining an SVM-based fault locator's optimal parameters

## Enfoque lineal para determinar los parámetros óptimos de un localizador de fallas basado en MSV

Lucas Pérez Hernández[1], Juan Mora Flórez[2] y Juan Bedoya Cebayos[3]

**ABSTRACT**

The setting up process for support vector machines (SVM) is discussed in this paper. Such settings are normally obtained from exhaustive testing of SVM, settled on by using several configuration parameter values and evaluating performance by using techniques such as cross validation. The linear approach presented in this paper is based on redefining the classical SVM second-order objective function. Better setting parameters were obtained by using low computational cost methodology for resolving new linear optimisation. The proposed approach was applied to a typical classification problem regarding fault location in power distribution systems; the results so obtained were compared to those obtained using classical methodology. An 80% improvement was achieved in mean error when estimating fault location and 56% reduction in the computing time needed for obtaining the best results when using classical approaches.

**Keywords**: euclidean distance norm, linear programming, fault location, support vector machine.

**RESUMEN**

En este artículo se discuten los procesos de ajuste de las máquinas de soporte vectorial (SVM), los cuales normalmente se obtienen partir de un proceso de prueba exhaustivo de varios valores de los parámetros de configuración. Posteriormente, su desempeño se evalúa utilizando técnicas tales como la validación cruzada. El enfoque aquí presentado se fundamenta en la redefinición de la función objetivo de segundo orden de una máquina de soporte vectorial clásica utilizando una aproximación lineal. Como principales resultados obtenidos al resolver el nuevo problema de optimización lineal, se obtienen mejores parámetros de configuración utilizando una metodología de bajo costo computacional. La aproximación propuesta se aplica a un problema de clasificación típico de la localización de fallas en sistemas de distribución de energía eléctrica, donde los resultados son comparados con aquellos obtenidos usando la metodología clásica. Se presenta un mejoramiento en los resultados logrados en el error promedio de estimación de la localización de la zona en falla del 80%, y una reducción del tiempo computacional del 56% del requerido para alcanzar los mejores resultados con las alternativas clásicas.

**Palabras clave:** norma de la distancia euclidiana, programación lineal, localización de fallas, máquinas de soporte vectorial.

Recibido: abril 28 de 2008
Aceptado: marzo 2 de 2009

## Introduction

During the last decade several approaches have been proposed for solving some specific engineering problems, mainly due to the increasing use of the data mining techniques. The main approaches have been based on artificial neural networks (ANN), fuzzy logic, statistical classifiers and support vector machines (SVM) (Kecman, 2000). One of the drawbacks of these methodologies is the difficulty associated with the setting up process involved in the techniques being used. Such setting-up is normally empirically performed or exhaustive, high computational cost search strategies are used in a feasible space for configuration parameters.

The performance obtained when using some of the previously mentioned techniques varied according to the problem being considered. Performance has normally been low when using ANN if there is a strong relationship amongst different classes and also in such cases related to problems with huge databases (Purushothama *et al.*, 2001). Several references provide interesting proposals in the specific case of applying classifiers to fault location problems in power systems. ANN have been widely applied to resolving problems of fault location in transmission and distribution systems (Purushothama, *et al.*, 2001) (Mescal *et al.*, 2003). ANN- and fuzzy logic-based applications have also been used for locating faults in power systems, as presented in (Mora *et al.*, 2006). Statistically based classifiers such as the learning algorithm for multivariable data analysis have been used for determining fault location in power distribution systems, as presented in (Mora *et al.*, 2007). SVM have been recently tested in power distribution system fault location (Mora *et al.*, 2006).

Power distribution system fault location is clearly more difficult to resolve than that occurring in transmission systems. The associated difficulty is caused by several characteristics related to power systems, such as voltage and current being typically available only at

[1] Tecnólogo electricista, Ingeniero electricista y Estudiante, M.Sc., en Ingeniería eléctrica, Universidad Tecnológica de Pereira, Colombia. Profesor auxiliar, Universidad Tecnológica de Pereira, Colombia. lucasperez@upt.edu.co

[2] Ingeniero electricista y M.Sc., en Ingeniería eléctrica, Universidad Industrial de Santander, Colombia, M.Sc., en Tecnológias de la Información. Ph.D., en Ingeniería eléctrica, Universitat de de Girona, España. Profesor tiempo completo, Universidad Tecnológica de Pereira, Colombia. jjmora@utp.edu.co

[3] Ingeniero electricista y Estudiante, M.Sc., en Ingeniería eléctrica, Universidad Tecnológica de Pereira, Colombia. Analista transacciones del mercado, XM Compañía de expertos en mercados S.A. E.S.P. jcbedoya@xm.com.co

the distribution substation, the presence of single and double phase laterals, tapped loads (single or multi-phase, having unknown hourly variation) along the lines and laterals and inconsistent network development and loads being responsible for lines having heterogeneous sections (the presence of different conductor gauges, a combination of overhead lines and underground cables). Short circuit level variation caused by changes in system topology and the equivalent generation source implies measurement variation for each specific fault condition (Mora *et al.*, 2008).

SVM's great ability for resolving fault location is clear, according to the techniques used and the results obtained in power distribution systems. SVM settings are normally obtained from an exhaustive test of the classifier, settled by using several configuration parameter values and evaluating the results using cross validation. Consequently, it is supposed that if there is lower computational cost and the best setting-up adjustment, then better results will be obtained than those obtained using a classical SVM approach.

Considering the above, the approach presented in this paper is based on redefining the classical SVM second-order objective function. This is redefined as a linear function using a different distance paradigm. Constraint functions are also redefined as linear ones while kernels maintain the same structure defined in the classical references for SVM. This approach was then tested regarding fault location applications for power distribution systems.

## Classical support vector machines (SVM)

Support vector machines (SVM) were used in this research as a classification technique for assisting fault location because of the good results reported in diagnostic applications (Thukaram *et al.*, 2005).

SVMs are based on statistical learning theory and can be viewed for practical purposes as being a binary classification technique resulting from the development of ANN and its combination with optimisation, kernel theory and generalisation theories (Burges, 1998) (Vapnik, 2000). The following subsection briefly summarises SVM: dealing with separable linear data, using a soft margin constraint for dealing with noisy data and using kernel functions with non-linear separable data.

### Linear data

Suppose having *n* training elements, $\boldsymbol{x}_i$, in an $\boldsymbol{N}$ dimensional space. Each element has its respective *tag etiquette (y)* as presented in (1). This etiquette is used for labelling members of the same class (+1 or -1).

$$\boldsymbol{x}_i \in R^N \text{ and } \boldsymbol{y}_i \in \{+1,-1\} \quad (1)$$

The aim is to find an optimal hyper plane H: $y=\boldsymbol{w}\cdot\boldsymbol{x}-b=0$ which has the maximum distance to the nearest training pattern to force learning machine generalization (Vapnik, 2000). This distance is normally known as margin, as presented in Figure 1 for a two-dimensional space.

Weight ($\boldsymbol{w}$) and bias ($b$) are the only two parameters used for controlling function. Those data points which the margin pushes up against are called *support vectors*.

The optimisation constrained problem presented in (2) must be solved to find the optimum separation hyperplane (OSH) (margin is inversely proportional to $\sqrt{\boldsymbol{w}\cdot\boldsymbol{w}}$ )

$$\min_{w,b} \frac{1}{2}(\boldsymbol{w}\cdot\boldsymbol{w}) \quad (2)$$

Subject to:

$$\boldsymbol{y}_i(\boldsymbol{w}\cdot\boldsymbol{x}_i+b) \geq 1, \forall i \quad (2)$$

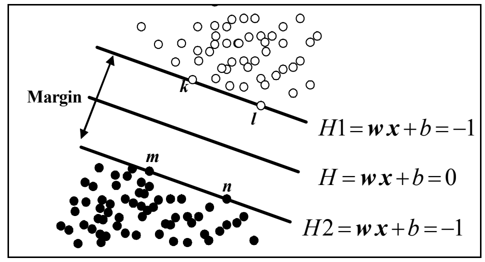


Figure 1. Separating hyper planes

### Soft margin

The previously presented methodology is based on the absence of mixed classes. The previous strategy was reformulated to cope with this common problem, by considering relaxing optimization to define what is known as a *"soft margin."* The optimization problem presented in (2) is now thus given as (3).

$$\min_{w,b} \frac{1}{2}(\boldsymbol{w}\cdot\boldsymbol{w})+C\sum_{i=1}^{n}\xi_i$$

$$\boldsymbol{y}_i(\boldsymbol{w}\cdot\boldsymbol{x}_i+b) \geq 1-\xi_i, \forall_i \quad (3)$$

Subject to:

$$\xi_i \geq 0, \forall i$$

Figure 2 shows this situation with nonlinearly separable classes. The slack variables $\xi_i$ mean that the classification error can be measured as a function distance to the hyperplane.

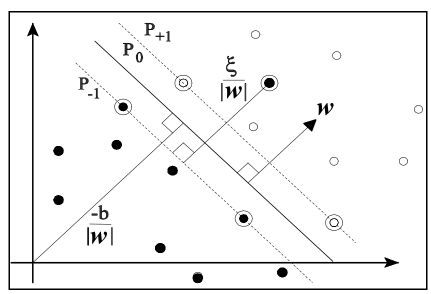


Figure 2. Non-separable case using a linear hyperplane

Parameter *C* is denoted as "error penalisation constant" and has to have a priory fixed by the user. A high *C* value means a high penalisation error.

### Kernel-based SVM

In the case of a non-linear separable dataset, the dataset can be transformed into a new high dimension space where the data is linearly separable. Figure 3 presents the intuitive idea of this transformation. The transformation function *Φ(.)* is defined in terms of input data scalar products in the original classification space. It is thus not necessary to specify *Φ(.)*; instead, kernel functions are used since they transform the scalar product in the transformed space in a single step.

Several kernel functions may be used in defining the new classification space, as presented in (Burges, 1998).

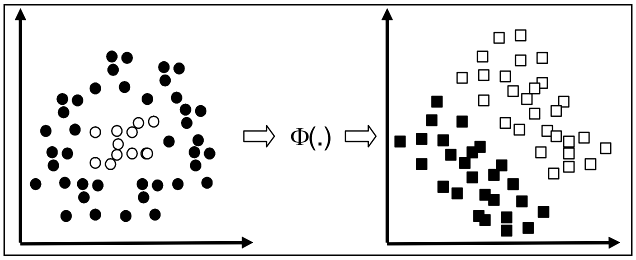


Figure 3. Data transformation in a new space where classes are linearly separable

Using an appropriate kernel function, SVM can separate data from different classes in this new space. Linear classification algorithms can thus be extended to non-linear cases by using an appropriate kernel function.

When an RBF is selected as kernel function (as in section five), two parameters (error penalisation constant $C$ and kernel parameter $\sigma$) have to be set to tune the classification algorithm. More extended information about SVM and kernel-based methods can be consulted in Burges (1998) and Vapnik (2000).

## The linear model proposed for support vector machines (linear programming support vector machine - LPSVM)

Suitable SVM configuration parameter values (error penalisation constant $C$ and kernel parameter $\sigma$) must be determined to solve the problem proposed in (3) using non-linear or quadratic programming techniques. A commonly used strategy for setting $C$ and $\sigma$ is based on an exhaustive search and classical techniques such as cross validation where the main idea is to perform an extensive simulation and test for determining such configuration parameters' best values (Chih *et al.,* 2003). This strategy is not always useful for finding optimal parameters and has a very high computational cost (Mora *et al.,* 2006).

The proposed solution is aimed at obtaining optimal SVM parameters based on redefining the quadratic optimisation problem presented in (3), as a new linear programming problem to that proposed in (4).

$$\begin{aligned} \|w\|_2^2 &= w_1^2 + w_2^2 + w_3^2 + \ldots + w_\ell^2 \\ \|w\|_1 &= |w_1| + |w_2| + |w_3| + \ldots |w_\ell| \end{aligned} \quad (4)$$

From (4), the new objective function in 1-norm and the set of constrains are redefined, as presented in (5).

$$\begin{aligned} &\min \quad Q(w,\xi) = \sum_{i=1}^{\ell} |w_i| + C\sum_{i=1}^{M} \xi_i \\ &s.t: \\ &\quad y_i\left(w^T g(x_i) + b\right) \geq 1 - \xi_i \end{aligned} \quad (5)$$

Where $M$ is the original data set dimension. From (4), it can be noticed how an alternative solution may be proposed using linear programming by using a linear kernel such as ($g(x) = x$ ). In the case of not using linear kernel functions, (5) must be redefined from such dual representation, as presented in (6).

$$D(x) = \sum_{i=1}^{M} \alpha_i H(x, x_i) + b \quad (6)$$

$\alpha_i$ and $b$ are real values and $H(x,x_i)$ is a kernel function defined as (7).

$$H(x, x_i) = g(x)^T g(x') \quad (7)$$

$g(x)^T$ is the function defining a new representation space where data is linearly separable using linear programming and $g(x')$ is the function used for transforming the data set. These two functions perform the same work as $\omega$ and $g(x)$ in equation constrains (5).

Considering the foregoing, SVM (LPSVM) linear formulation is now given as (8).

$$\begin{aligned} &\min \quad Q(w,\xi) = \sum_{i=1}^{M} \left(|\alpha_i| + C\xi_i\right) \\ &s.t: \\ &\quad y_i\left(\sum \alpha_i H(x, x_i) + b\right) \geq 1 - \xi_i \end{aligned} \quad (8)$$

Where $\alpha_i$ contains orthogonal vectors from the original representtation space. $\alpha$ and $b$ must now be redefined to solve the linear programming problem proposed in (8), as proposed in (9).

$$\begin{aligned} \alpha_i &= \alpha_i^+ - \alpha_i^- \\ b &= b^+ - b^- \end{aligned} \quad (9)$$

From the proposed definition of $\alpha$ and $b$, the idea of a linear distance between classes is considered. Having redefined the classic SVM problem as a linear problem, special considerations must be adopted to avoid degenerate solutions (those giving more than one hyperplane generated by the active constraint). Such new formulation of the optimisation problem could give degenerated solutions, especially in the case of small values of $C$, as explained in appendix A.

Now, the problem proposed in (8) is presented as (10), by using (9).

$$\begin{aligned} &\min \quad Q = \sum_{i=1}^{M} \alpha_i^+ + \alpha_i^- + C\xi_i \\ &s.t: \\ &\quad y_i\left(\sum\left(\alpha_i^+ - \alpha_i^-\right) H(x, x_i) + b^+ - b^-\right) + \xi_j = 1 + u_j \end{aligned} \quad (10)$$

The decision function for (10) is now presented as (11).

$$D(x) = \sum_{i=1}^{M} \left(\alpha_i^+ - \alpha_i^-\right) H(x, x_i) + b^+ - b^- \quad (11)$$

In (11), adding all $\alpha_i$ $g(x_i)$, for $i = 1, \ldots, M$, minimisation of $\sum|\alpha_i|$ does not causes the immediate maximisation of the separating margin estimated using a 1-norm. That is why the dual problem of (10) must be evaluated, as presented in (12).

$$\begin{aligned} &\max \quad z_i \\ &s.t: \\ &\quad \sum_{i=1}^{M} y_i H(x_i, x_j) z_i + v_j^+ = 1 \quad j = 1 \ldots M \\ &\quad \sum_{i=1}^{M} y_i H(x_i, x_j) z_i = v_j^- - 1 \quad j = 1 \ldots M \\ &\quad \sum_{i=1}^{M} y_i z_i = 0 \\ &\quad z_j + w_j = C \qquad j = 1 \ldots M \end{aligned} \quad (12)$$

All the constraints presented in (12) are simplified by using slack variables ($v$ and $w$), and the dual variables are represented by $z_i$, redefining the maximum separating margin.

Solving the new problem leads to solutions for dual (12) and primal (10) formulations. If these solutions are optimal, they meet the complementarily conditions presented in (13), according to classic linear and non-linear programming theory (Bazaraa *et al.,* 1990) (Bazaraa and Shetty, 1993).

$$\begin{aligned} \alpha_i^+ v_i^+ &= 0 \quad for \quad i = 1 \ldots M \\ \alpha_i^- v_i^- &= 0 \quad for \quad i = 1 \ldots M \\ \xi_i w_i &= 0 \quad for \quad i = 1 \ldots M \\ u_i z_i &= 0 \quad for \quad i = 1 \ldots M \end{aligned} \qquad (13)$$

The training set $x_i$, where the constraints presented in (14) are not satisfied, forms the support vectors.

$$\begin{aligned} \alpha_i &= \alpha_i^+ - \alpha_i^- = 0 \\ \xi_i &= 0 \\ z_i &= 0 \end{aligned} \qquad (14)$$

Equation (14) originates due to the following two conditions of a classic linear problem being complied with (Bazaraa *et al.,* 1990) (Bazaraa and Shetty, 1993):

a. If a hyperplane generated by one constraint is active, then its dual variable is higher than zero and its slack variable equals zero; and

b. If a hyperplane generated by one constraint is not active, its dual variable equals zero.

## Tests and results

This section is devoted to presenting the effectiveness of how the proposed approach was tested by comparing LPSVM performance with that of a classical SVM. The proposed problem was related to determining the fault area in a power distribution system.

### Power system used for tests

A 123-node 4.16 kV power distribution system proposed by the IEEE (Figure 4) has been used for testing the fault location approach, as in (IEEE, 1993). A complete dataset have been obtained from simulation of faults in each node under different conditions. Single-, double- and three-phase faults have been simulated by using 21 different fault resistance values ranging from 0.5 to 40 ohms (Dagenhart, 2000). The alternative transient program (ATP) has been used in an integrated environment linking ATP and Matlab to automate fault generation when simulating a power system (Mora *et al.,* 2006). The fault database consisted of 4,495 records of voltage and current.

### Power system characterisation – features or attributes

Characterising voltage and current signals measured at the substation before and during a fault is presented in this section. Such characterisation has been applied to transient and steady state signals (Mora, 2006).

Characterisation has resulted in obtaining a set of attributes for each fault situations which have been used as inputs for the statistical learning-based classifier. The classifier can be trained to recognise the faulty area in a power distribution system by means of such attributes.

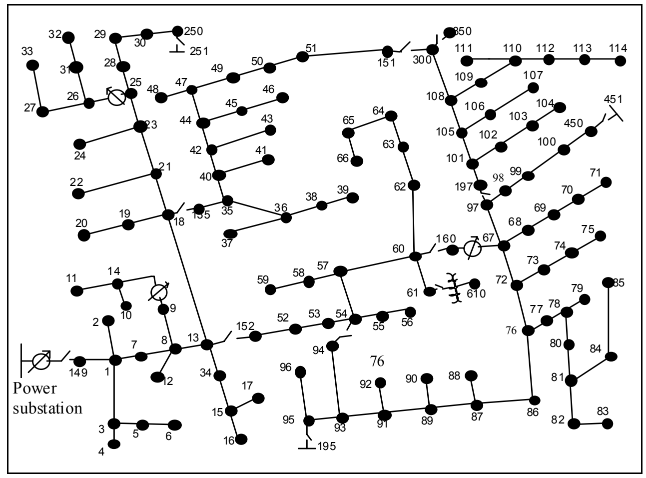


Figure 4: IEEE 123 node test distribution feeder simulated for locating faults

The attributes used for characterising the faults were voltage magnitude variation ( $\Delta V$ ) associated with variation in the effective value of the steady states before and during a fault, current magnitude variation ( $\Delta I$ ) as effective current variation value is defined as being the variation between pre-fault and fault steady states, apparent power variation ( $\Delta S$ ) related to load variation during the fault situation, fault reactance ( $X_f$ ) and voltage transient frequency caused by the fault ( $f$ ).

### Estimating parameters $\sigma$ and C, using classical SVM

The classical strategy for obtaining good parameter combination (low classification error) is given by the net search, basically consisting of proposing a variation range where paired parameter values ($\sigma$, $C$) are tested. A nonlinear optimisation algorithm has been used in extensive testing using cross validation for determining classification error (sequential minimal optimisation algorithm- SMO), (Chih *et al.,* 2003).

### Solving the linear SVM model

The mathematical model proposed in this paper is given by (10), which is known as a linear model constrained by equalities and continuous variables. This model is solved by using the simplex method, coded in a straightforward and efficient routine (Vapnik, 2000). This routine was implemented by using Matlab.

By solving the proposed model, a partial solution is always feasible but not optimal and that is why a complete solution should be achieved to evaluate all possible partial solutions. Parameter $\sigma$ and $C$ optimal values are contained in the proposed linear model's constraints if it is optimal and feasible.

### Results

Classification tests were applied to recognise faults in a power distribution system using voltage and current signal attributes as input. Table 1 gives the results obtained with classical SVM while Table 2 gives corresponding LPSVM results. The number of classification events, those used for validation and the optimal parameters obtained in each tests are given in both Tables. Table 1 gives the number of SVM needed to solve the problem; however, this column is not given in Table 2 because the problem was solved by using a linear programming algorithm.

Fault locator performance or level of confidence for both SVM was based on the difference of the total data used for the test and well-classified registers, as presented in (15).

$$Confidence\,[\%] = \frac{Number\ of\ \ faults\ well\ located}{Total\ \ number\ of\ \ faults} \times 100 \qquad (15)$$

Tables 1 and 2 present the results while a comparative graph of errors is presented in Figure 5.

Table 1. Results obtained using classical SVM

| *Fault type* | | *Training data* | | *Test data* | | *SVM adjustment* | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Number of faults | Confidence [%] | Number of faults | Confidence [%] | Number of SVMS | Number of vectors | Parameters $\sigma$ | Parameters $C$ |
| Single phase fault | A-Ground | 412 | 97.2 | 563 | 95.6 | 190 | 4 | 5.8 | 0.92e5 |
| | B-Ground | 352 | 96.3 | 462 | 94.5 | 136 | 7 | 0.56 | 1.34e3 |
| | C-Ground | 392 | 96.8 | 532 | 94.8 | 136 | 9 | 0.43 | 4.92e4 |
| Double phase fault | A-B, A-B-G | 584 | 98.3 | 718 | 98.0 | 91 | 4 | 5.5 | 5.0e4 |
| | B-C, B-C-G | 568 | 98.4 | 698 | 97.8 | 78 | 5 | 3.3 | 1.25e5 |
| | C-A, C-A-G | 592 | 98.1 | 728 | 98.0 | 91 | 4 | 5.6 | 5.01e4 |
| Three phase fault | A-B-C, A-B-C-G | 568 | 99.2 | 698 | 99.0 | 78 | 4 | 5.0 | 1e3 |

Table 2. Results obtained using proposed LPSVM

| *Fault type* | | *Training data* | | *Test data* | | *SVM adjustment* | | |
|---|---|---|---|---|---|---|---|---|
| | | Number of faults | Confidence [%] | Number of faults | Confidence [%] | Number of vectors | Parameters $\sigma$ | Parameters $C$ |
| Single phase fault | A-Ground | 412 | 99.6 | 563 | 99.1 | 19 | 5.0 | 1.0e5 |
| | B-Ground | 352 | 99.7 | 462 | 99.2 | 15 | 0.5 | 1.0e3 |
| | C-Ground | 392 | 99.5 | 532 | 98.0 | 18 | 0.5 | 5.1e4 |
| Double phase fault | A-B, A-B-G | 584 | 99.8 | 718 | 99.7 | 12 | 5.0 | 5.1e4 |
| | B-C, B-C-G | 568 | 100.0 | 698 | 99.8 | 19 | 2.8 | 1.0e5 |
| | C-A, C-A-G | 592 | 100.0 | 728 | 100.0 | 19 | 5.0 | 5.1e4 |
| Three phase fault | A-B-C, A-B-C-G | 568 | 99.8 | 698 | 99.6 | 17 | 5.0 | 1e3 |

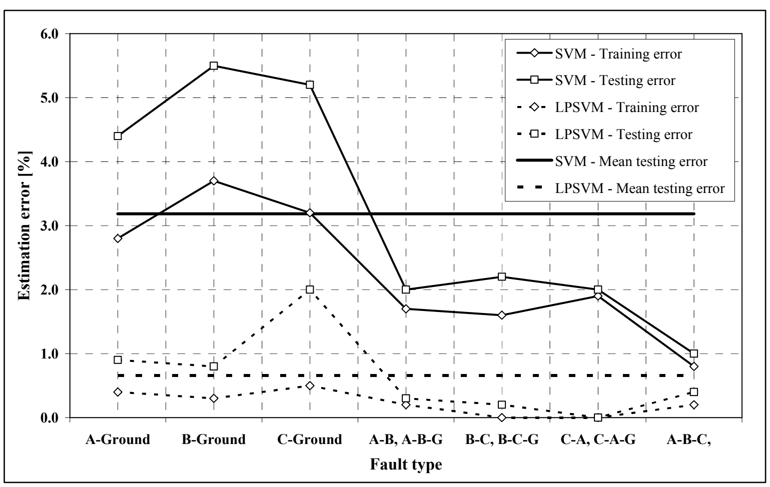


Figure 5. Comparative graph of errors for SVM and the proposed LPSVM

## Analysis of results

The results given in Tables 1 and 2 show how the solutions were good by using the net search; however, the solutions were better than those obtained using the classical approach when using the proposed approach. This is better presented in Figure 5.

In addition to the above, computational cost was comparatively low for the proposed approach where solutions were obtained in only 2,400 seconds whilst 5,800 seconds were necessary to resolve the complete problem using the classical approach.

The absence of degenerated solutions could be determined by considering the high values of *C* in both cases.

As an interesting comparison between LPSVM and SVM models, LPSVM input and output variables gave feasible but not optimal solutions, similarly to the net search performed by the classical SVM. The difference was associated with the best process associated with the simplex optimisation method used in LPSVM solution.

## Conclusions

Fault location in power distribution systems is an interesting practical problem for most utili-ties. The classification approach used for deter-mining faulty areas is one of the commonly applied alternatives taking advantage of fault databases.

Classical SVM approaches (based on a quadratic optimisation problem) have given very good fault location results; however, high computational efforts are required for adjusting the SVM-based locator. The proposed approach (based on a linear programming optimisation strategy) gives an optimal, low computational cost solution to fault location because SVM parameters are also included in the optimisation problem, thereby reducing the time taken for extensive tests and cross validation.

The proposed approach was applied to typical fault location classification in a 128 node power distribution system; the results were compared to those obtained using classical methodology. An 80% improvement was achieved regarding mean fault location estimation error in testing and 56% reduction in the computing time needed for obtaining the best results using classical approaches. LPSVM performance was better than classical SVM, considering fault location error and computational cost.

This approach contributes towards improving power continuity indexes in distribution systems by opportune area-specific fault location. It helps reduce the impact of such faults on duration- and frequency-associated indexes in three ways; fault location helps speed up service restoration, switching operations can reduce the faulty area once the fault has been located and scheduled maintenance tasks can be performed by locating non-permanent faults to avoid future faults.

## Bibliography

Bazaraa M., Sherali H., Jarvis J., Linear Programming and Network Flows., Second Edition, John Wiley and Sons. 1990.

Bazaraa M., Shetty C., Nonlinear Programming: Theory and Algorithms., Second Edition, John Wiley and Sons., 1993.

Burges, C., A tutorial on support vector machines for pattern recognition., Data Mining and Knowledge Discovery, 1998, pp. 121-167.

Chih, H., Chung, C., Jen, L., Practical guide to support vector classification., Reporte técnico, Department of Computer Science, National Taiwan, Tech. Rep., 2003.

Dagenhart, J., The 40- Ground-Fault Phenomenon., IEEE Transactions on Industry Applications, Vol. 36, No. 1, 2000, pp. 30-32.

IEEE Distribution System Analysis Subcommittee., Radial Test Feeders., IEEE Standards Board. 1993. (http://www.ewh.ieee.org/soc/pes/dsacom/testfeeders.html)

Kecman, V., Learning and Soft Computing: Support Vector machines., Neural Networks, and Fuzzy Logic Models, MIT Press., 2001.

Mescal, A., Al-shaher, A., Manar, S., Fault location in multi-ring distribution network using artificial neural network., Electric Power Systems Research, 2003, pp. 87-92

Mora, J, Carrillo, G., Pérez, L., Fault location in power distribution systems using ANFIS nets and current patterns., 2006 IEEE PES Transmission and Distribution Conference and Exposition, Caracas, August, 2006.

Mora, J, Melendez, J., Carrillo, G. Fault Location in Power Distribution Systems Based on Signal Descriptors., 2006 IEEE 12th International Conference on Harmonics and Quality of Power, Porto, October, 2006.

Mora, J., Bedoya, J., Melendez, J., Extensive Events Database Development using ATP and Matlab to Fault Location in Power Distribution Systems., Transmission & Distribution Conference and Exposition: Latin America, TDC '06. IEEE/PES, 2006.

Mora, J., Localización de Fallas en Sistemas de Distribución de Energía Eléctrica usando Métodos Basados en el Modelo y Métodos Basados en el Conocimiento., Tesis Doctoral. University of Girona, España. 2006.

Mora, J., Carrillo, G., Barrera, B. Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis., IEEE Transaction on Power Delivery, Vol. 22, No. 3, July, 2007.

Mora, J., Carrillo, G., Meléndez, J., Comparison of impedance based fault location methods for power distribution systems., Electric Power Systems Research. Doi:10.1016/j.epsr.2007.05.010, availabe online 20 July 2007.

Purushothama, G., Narendranath, A., Parthasarathy, K., ANN applications in fault locators., International Journal of Electrical Power and Energy Systems, 2001, pp. 491-506.

Thukaram, D., Khincha, H., Vijaynarasimha, H., Artificial Neural Network and Support Vector Machine Approach for Locating Faults in Radial. Distribution Systems., IEEE Transactions on Power Delivery, Vol. 20, No. 2, April, 2005, pp. 710 721.

Vapnik. V., The nature of Statistical Learning Theory., Second Edition, Springer Verlag, 2000.

## Appendix A: The presence of degenerated solutions

### Proposed theorem

There was a positive $C_0$ value with LPSVM, where there was a degenerated solution for a $C$ value between 0 and $C_0$ including the borders.

### Test

Considering a feasible solution to $\xi$ slack variables in the case of high C values, the problem proposed in (8) had an optimal solution for positive values of $a_i$. Variations in objective function and constrains in the case of optimality conditions being accomplished were as follows:

a. For $a_i = 0$, and $i = 1, \ldots, M$, constrains presented in (8) were rewritten as (A.1).

$$y_i b \geq 1 - \xi_0 \qquad (A.1)$$

b. For $b = 0$, the objective function was satisfied if $\xi = 1$, as given by (A.2).

$$Q = MC \qquad (A.2)$$

Thus, if the C value decreased, it was possible to find a maximum $C = C_0$, where the objective function had a minimum in $a_i = 0$. If there were a basis variable and its value equalled zero, then there was a degenerated solution.