

Análisis comparativo de metaheurísticas para calibración de localizadores de fallas en sistemas de distribución

Comparative analysis of metaheuristics optimization techniques to parameterize fault locators for power distribution systems.

Walter J. Gil-González*, Juan J. Mora-Flórez*§, y Sandra. Pérez-Londoño*

Grupo de Investigación en Calidad de Energía Eléctrica y Estabilidad (ICE³).

**Programa de Ingeniería eléctrica de la Universidad Tecnológica de Pereira (UTP)*

wjgil@utp.edu.co, §jjmora@utp.edu.co, saperez@utp.edu.co

(Recibido: Marzo 21 de 2012 - Aceptado: Mayo 16 de 2013)

Resumen

En este artículo se presenta un análisis comparativo de cuatro técnicas metaheurísticas que permiten la calibración óptima de un localizador de fallas basado en máquinas de soporte vectorial (MSV), con el objetivo de determinar cuáles técnicas presentan un mejor rendimiento para este problema. El localizador propuesto se prueba en un sistema de distribución estándar de 34 nodos de la IEEE, en el que se obtienen precisiones promedio para las mejores alternativas de 99%, utilizando una base de datos de 13824 registros de fallas monofásicas, bifásicas, bifásicas a tierra y trifásicas. La comparación de las alternativas de parametrización muestra que las técnicas metaheurísticas basadas en población presentaron un mejor rendimiento que aquellas basadas en trayectoria, en todos los casos estudiados.

Palabras Clave: *Calibración, localización de fallas, máquinas de soporte vectorial, sistemas de distribución, técnicas metaheurísticas.*

Abstract

In this paper, a comparative analysis on the use of four metaheuristics for obtaining an optimal adjustment of a fault locator based on support vector machines (SVM), is presented. This research is aimed to determine those techniques which help to obtain the best performance at the specific problem of fault location. The proposed fault locator is tested in the 34 nodes IEEE power distribution system where the average precision obtained considering the best alternatives is around 99%, using a database of 13824 registers from single phase, phase to phase, two phase to ground and three phase faults. The comparison of the parameterization alternatives shows how those metaheuristics based on population have better performance than those based on trajectory, having a good performance in all of the tested situations.

Keywords: *Calibration, fault locator, metaheuristics techniques, power distribution system, support vector machines.*

1. Introducción

La energía eléctrica constituye uno de los insumos fundamentales para el desarrollo económico y social de una región, dado que su disponibilidad determina en gran medida los niveles de productividad, las posibilidades de desarrollo agroindustrial y la calidad de vida de los pobladores. Por ello, el estudio de la calidad del suministro de energía eléctrica ha tomado mucha fuerza y dentro de esta temática, el problema de la localización de fallas para mantener altos índices de continuidad es uno de los más importantes, Short (2003), IEEE (2004).

Para la solución del problema de localización de falla en sistemas de distribución se han analizado diversos métodos, los cuales utilizan las medidas de tensión y corriente en la subestación. Los métodos basados en el modelo eléctrico de la red proporcionan información sobre la distancia asociada a la impedancia de falla. Debido al gran número de laterales, la distancia estimada se cumple para varios sitios del sistema, convirtiendo éste en un problema de múltiple estimación, que complica la ubicación del lateral bajo falla (Mora, 2006). Adicionalmente, se han utilizado métodos basados en minería de datos para solucionar el problema de múltiple estimación del sitio de falla, los cuales han sido probados con éxito en la localización de fallas, Morales et al. (2008), Thukaram et al. (2005).

Cada método debe ser parametrizado adecuadamente para asegurar un alto desempeño en la localización de fallas. Para encontrar los parámetros no existe una técnica exacta y una búsqueda exhaustiva representaría un esfuerzo computacional muy elevado que podría ser de muchos años, con la capacidad de cálculo actual. Por lo tanto, es necesario implementar técnicas que permitan obtener los parámetros que mejor se adapten al problema en estudio, ya que actualmente no hay una alternativa para determinar un valor genérico que se ajuste. Una de las primeras estrategias utilizadas fue la búsqueda en malla, que es una técnica poco eficiente y

con alto esfuerzo computacional, Morales & Gómez (2005). Otra alternativa para afrontar este problema es utilizando técnicas de optimización para reducir el alto esfuerzo computacional y mejorar la eficiencia en la parametrización, tal como se propone en Gutiérrez et al. (2010), donde se implementa una estrategia híbrida de bajo costo computacional fundamentada en el algoritmo genético de Chu Beasley.

En este artículo, y como alternativa de solución al problema de parametrización de los métodos basados en minería de datos aplicados a la localización de fallas, se propone un análisis comparativo amplio entre varias técnicas metaheurísticas mostrando que ventajas y desventajas tiene cada, ya que el rendimiento de cada técnica varía de acuerdo al problema bajo estudio, Rossi, O.D. et al. (2002).

2. Aspectos teóricos básicos

2.1 Máquina de soporte vectorial (MSV)

La MSV es una técnica de minería de datos basada en los fundamentos de la teoría de aprendizaje estadístico desarrollada por Vapnik y Chervonenkis, Moguerza & Muñoz (2006), la cual, a diferencia de otros métodos, tiene la ventaja que no requieren ningún tipo de conjetura sobre la densidad de probabilidad de los datos. La arquitectura de la MSV sólo depende de un parámetro de penalización denotado como C y la función kernel (incluyendo sus parámetros). En el caso de la Función Base Radial (RBF), existe sólo un parámetro denotado como γ , como se presenta en la Ec. (1) Burges, (1998).

$$k(\vec{x}, \vec{y}) = e^{-\gamma \|\vec{x} - \vec{y}\|^2} \quad (1)$$

2.2 Técnicas metaheurísticas utilizadas

En esta sección se presentan brevemente los fundamentos de las técnicas metaheurísticas utilizadas en la calibración de las máquinas de soporte vectorial. Una descripción detallada de cada una de las técnicas está fuera del alcance de este artículo, pero puede obtenerse a partir de las

referencias citadas. Se utilizaron dos técnicas de basadas en trayectorias como la *búsqueda tabú* y *recocido simulado*, y dos técnicas basadas en población como la *optimización por enjambre de partículas* y *por colonia de hormigas*.

a. *Búsqueda tabú (TS)*. La idea básica de la búsqueda tabú es la utilización explícita de un historial de búsqueda (una memoria de corto plazo), tanto para escapar de los óptimos locales como para implementar una estrategia de exploración y evitar la búsqueda repetida en la misma región, Glover & Kochenberger (2002). Esta memoria de corto plazo se implementa como una lista tabú, donde se mantienen las soluciones visitadas más recientemente para excluirlas de los próximos movimientos. En cada iteración se elige la mejor solución entre las permitidas y ésta se añade a la lista tabú.

b. *Recocido simulado (SA)*. La idea del SA es simular el proceso de recocido del metal y del cristal, donde para evitar los óptimos locales, el algoritmo permite elegir una solución peor que la solución actual, Faber et al. (2008). En cada iteración se elige, a partir de la solución actual s , una solución s' del vecindario $N(s)$. Si s' es mejor que s (es decir, tiene un mejor valor en la función de *fitness*), se sustituye s por s' como solución actual. Si la solución s' es peor, entonces se acepta con una determinada probabilidad que depende de la temperatura actual T y de la variación en la función de *fitness* $f(s') - f(s)$ (caso de minimización). Esta probabilidad generalmente se calcula siguiendo la distribución de Boltzmann, que se presenta en la Ec. (2).

$$P = e^{\left(\frac{f(s') - f(s)}{T}\right)} \quad (2)$$

c. *Optimización por enjambre de partículas (PSO)*. Es una técnica metaheurística inspirada en el comportamiento social del vuelo de las bandadas de aves o el movimiento de los bancos de peces, Gómez (2008). Se fundamenta en los factores que influyen en la toma de decisión de un agente que forma parte de un conjunto de agentes similares. La toma de decisión por parte de cada

agente se realiza conforme a una componente social y una componente individual. De esa forma se determina el movimiento (dirección) de este agente, para alcanzar una nueva posición en el espacio de soluciones. Simulando este modelo de comportamiento se obtiene un método para resolver problemas de optimización.

d. *Optimización por colonia de hormigas (ACO)*. Es una técnica metaheurística inspirada en el comportamiento real de las hormigas cuando realizan la búsqueda de comida, Gallego et al. (2008). Este comportamiento tiene como etapa inicial la exploración aleatoria del área cercana al nido de la hormiga. Tan pronto como una hormiga encuentra la comida, la lleva al nido. Mientras que realiza este camino, la hormiga va depositando una sustancia química denominada feromona, que ayudará al resto de las hormigas a encontrar la comida. Esta comunicación indirecta entre las hormigas mediante el rastro de feromona las capacita para encontrar el camino más corto entre el nido y la comida. Esta funcionalidad es la que intenta simular este método para resolver problemas de optimización, donde el rastro de feromona se simula mediante un modelo probabilístico.

3. Metodología propuesta

La metodología general utilizada para la calibración o parametrización de la MSV con el propósito de la localización de fallas en sistemas de distribución, se presenta gráficamente en el esquema mostrado en la figura 1.

En la etapa 1 se implementan las técnicas metaheurísticas para el ajuste de parámetros óptimos en la MSV. En la etapa 2 se utilizan los parámetros encontrados (C, γ) en la etapa anterior, para el entrenamiento y la validación del método, y así localizar la falla. Estas dos etapas se aplican para cada una de las cuatro técnicas metaheurísticas que se analizan en este artículo.

3.1 Etapa 1: Parametrización de la MSV

En la etapa de calibración o parametrización, se

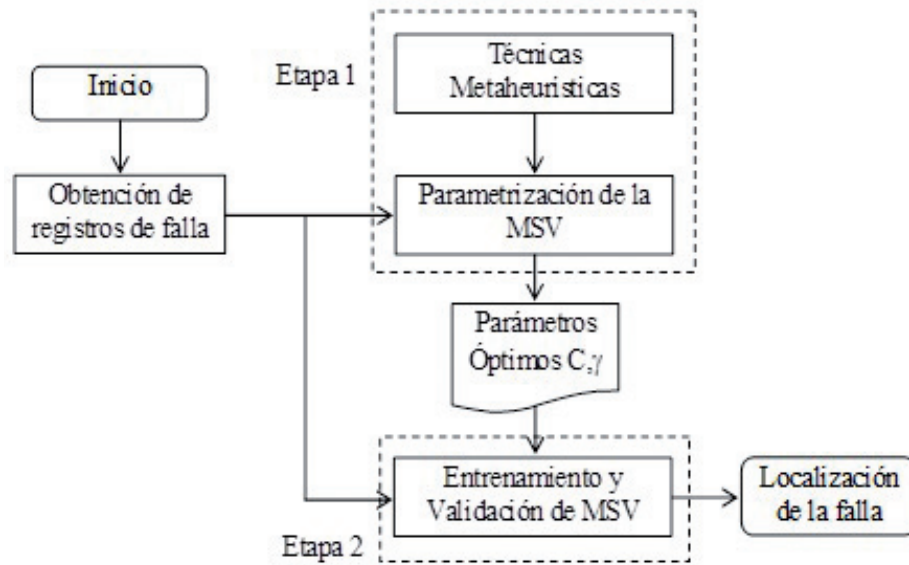


Figura 1. Esquema general del método de clasificación

implementa cada técnica metaheurística mediante la técnica de validación cruzada para 10 combinaciones de descriptores que se utilizaron como entradas, Gutiérrez et al. (2010). El error de validación cruzada se utiliza como la función objetivo de cada técnica, así el algoritmo basado en cada técnica metaheurística evoluciona hasta encontrar el menor error de validación.

La validación cruzada es un método que consiste en dividir la base de datos de entrenamiento en n partes iguales. A continuación, el localizador basado en MSV se entrena con los datos contenidos en las $n-1$ partes de la base de datos y la parte restante para hallar el error de validación, calculada como se muestra en la Ec. (3). Este proceso se repite n veces, lo que permite utilizar todas las muestras para hallar un error de validación con esta base de datos. Por último, se deben promediar los n valores de error de validación encontrados, para obtener un solo error de validación asociado a los parámetros que se están utilizando, Morales & Gómez (2005).

La etapa de parametrización sigue varios pasos, los cuales se describen a continuación:

a. Zonificación de la red de distribución. El proceso de clasificación requiere que para cada uno de los datos de entrenamiento se le asigne una clase, con la cual se realizará la clasificación de un nuevo dato. Esta clase corresponde a una zona del sistema de distribución bajo análisis. Una zona no debe tener más de un lateral del circuito, para eliminar el problema de múltiple estimación, Mora (2006). Además, es posible hacer la zonificación con zonas grandes, si así lo requiere el operador de red, aunque también es posible reducir su tamaño y así dar importancia a aquellas zonas donde la probabilidad de ocurrencia de fallos sea mayor, o donde sea necesario restaurar el servicio de forma más rápida.

b. Adquisición y pre-procesamiento de los datos. Los datos utilizados para el entrenamiento fueron obtenidos a partir de la herramienta de simulación *Alternative Transients Program (ATP)*. Además, se utilizó la herramienta *simulacionRF* la cual

$$\text{Error de validación} = \frac{\text{Número de datos mal clasificados}}{\text{Número de datos de prueba}} \cdot 100\% \quad (3)$$

simula automáticamente condiciones de falla monofásica, bifásica y trifásica con diferentes valores de resistencia de falla.

El pre-procesamiento de los datos de falla está asociado con la extracción de información significativa denominada descriptores. En este artículo, los descriptores utilizados corresponden a las variaciones de la tensión de fase (dV), de la corriente (dI), del ángulo de tensión de fase ($d\theta_v$), del ángulo de corriente de fase ($d\theta_i$), de la tensión de línea (dV_L), de la corriente de línea (dI_L), del ángulo de tensión de línea ($d\theta_{vL}$) y del ángulo de corriente de línea ($d\theta_{iL}$). A cada descriptor se le asigna una etiqueta relacionada con la zona en la cual ocurrió la falla, Gutiérrez et al. (2010). Adicionalmente, para evitar que descriptores con magnitudes más altas dominen en el cálculo de la zona, los datos se normalizan en el intervalo [0, 1], tal como se propone en la Ec. (4).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4)$$

Donde X es el descriptor, X_{min} y X_{max} son los valores mínimo y máximo que existen para cada descriptor en la base de datos de falla, respectivamente.

c. *Búsqueda de parámetros óptimos para cada combinación de descriptores de entrada.* Con los descriptores mencionados en el paso b es posible realizar diferentes combinaciones para ajustar el clasificador. En total se utilizaron 10 combinaciones que sirven como entrada del clasificador MSV.

En las siguientes secciones se muestra cómo se implementaron cada una de las técnicas de parametrización del localizador basado en MSV. El criterio de parada para las técnicas utilizadas son un límite de iteraciones y un valor mínimo de la función objetivo

3.1.1 Parametrización de la MSV con búsqueda tabú (TS)

Como se presenta en la figura 2, la técnica basada en TS comienza con un punto inicial aleatorio, del

cual se calcula la función objetivo y su resultado se almacena como incumbente. Inicialmente, se crea la vecindad alrededor del punto inicial, determinada como se muestra en la Ec. (5) y se evalúa la función objetivo de cada punto. El punto actual se guarda en la lista tabú de tamaño M y queda restringido por las próximas M iteraciones (donde M es igual a 10, debido a que con este valor se han obtenido buenos resultados para este problema). Se elige la mejor solución de la vecindad y se compara con la solución global, y si ésta es menor se guarda como la nueva incumbente.

$$C_{i+1} = R_c \cdot \sin \theta + C_i \text{ con } R_c = \text{limite máximo de } C/C'$$

$$\gamma_{i+1} = R_\gamma \cdot \cos \theta + \gamma_i \text{ con } R_\gamma = \text{limite máximo de } \gamma/\gamma' \quad (5)$$

Los parámetros C' y γ' tienen que ver con la vecindad del punto actual, que para este problema se escogió $C' = 250$ y $\gamma' = 33$, para que las nuevas soluciones sean diferentes a la actual. El valor de θ está espaciado en 45° entre una solución y otra, para que se explore uniformemente toda la vecindad.

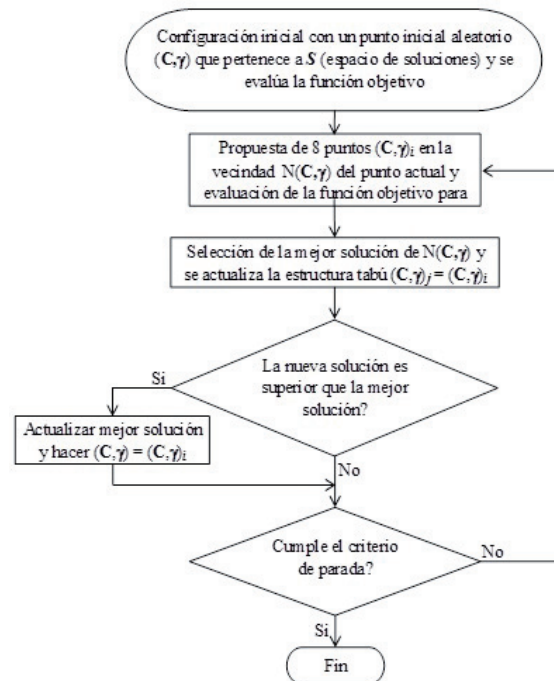


Figura 2. Diagrama de flujo básico para el algoritmo TS

3.1.2 Parametrización de la MSV con recocido simulado (SA)

Como se presenta en la figura 3, el SA primero calcula la temperatura inicial, tal como se presenta en forma detallada en la figura 4, y se guarda la mejor solución encontrada. Luego de tener la temperatura inicial, se crea un punto aleatorio que pertenece al espacio de soluciones y se evalúa en la función objetivo. Posteriormente se compara la nueva solución con la incumbente, si ésta es menor, se guarda como la nueva incumbente; en caso contrario, la nueva solución se guarda como la actual si ésta supera un criterio probabilístico. El proceso realizado anteriormente representa una alternativa de la temperatura actual.

El número de alternativas se calcula como se muestra en la Ec. (6), donde se utilizó un $h = 1$ para que el número de alternativas fuera constante. El decrecimiento de la temperatura se calcula como se muestra en la Ec. (7) y con un valor de $\beta = 0.9$, tal como se recomienda en Eberhart & Yuhui (2001).

En la figura 3, el valor de N_0 depende de dos constantes (α, k), con $\alpha = 2$, ya que depende del número de variables del problema (C), y con $k = 6$ debido a que este problema es de alta complejidad, Gallego et al. (2008). En el algoritmo de la figura 4 se crea un punto aleatorio y se evalúa la función objetivo. Se crea otro punto aleatorio, se evalúa la función objetivo y se compara con la solución

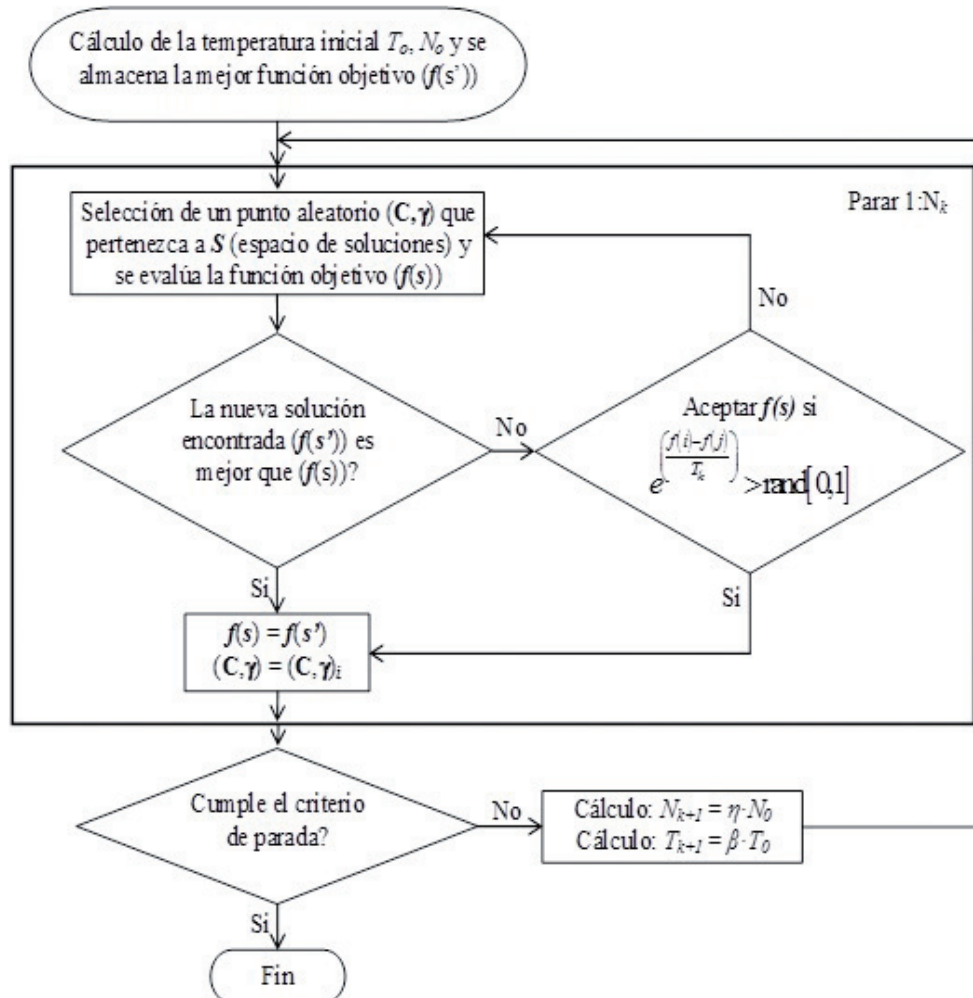


Figura 3. Diagrama de flujo básico para el algoritmo SA

anterior. Si ésta es menor, se aumenta m_1 y en caso contrario se aumenta m_2 ; esto se hace hasta que se cumpla la Ec. (8). Posteriormente se calcula la temperatura inicial T_0 como se muestra en la Ec. (9), donde se utilizó $X = 0.85$, tal como se

$$N_{k+1} = \eta N_k \quad (6)$$

$$T_{k+1} = \beta T_k \quad (7)$$

$$m_1 + m_2 = N_0 \quad (8)$$

$$T_0 = \frac{\Delta F +}{\ln\left(\frac{m_2}{m_2 \cdot X - m_1(1-X)}\right)} \quad (9)$$

Donde:

$$\Delta F + = \frac{\sum_{i=1}^{m_2} \Delta F_i}{m_2} \quad (10)$$

3.1.3 Parametrización de la MSV con optimización por cúmulo de partículas (PSO)

Como se muestra en Gómez (2008), PSO primero define un número de partículas ($P = 20$) (según recomendación de Gómez (2008)) y la variación máxima de velocidad ($V_{\text{máx}}$) para la primera iteración está definida por la Ec. (11). Se crean P partículas aleatoriamente, cada una de las cuales es una solución (C, γ) del problema que se está tratando. Se evalúa la función en cada partícula y se guarda la mejor solución encontrada y su posición. En la segunda iteración se generan las velocidades para cada punto aleatoriamente, en las siguientes iteraciones la velocidad se calcula de acuerdo a la Ec. (12). Se actualizan las partículas

$$-500 \leq V_{\text{máx}} C \leq 500 \quad (11)$$

$$-2 \leq V_{\text{máx}} \gamma \leq 2$$

$$v^t \leftarrow \omega^{t-1} \cdot v^{t-1} + \phi_1 \cdot \text{rand}_1 \cdot (\text{mejorpos}_i - x_i^{t-1}) + \phi_2 \cdot \text{rand}_2 \cdot (\text{mejorpos} - x_i^{t-1}) \quad (12)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (13)$$

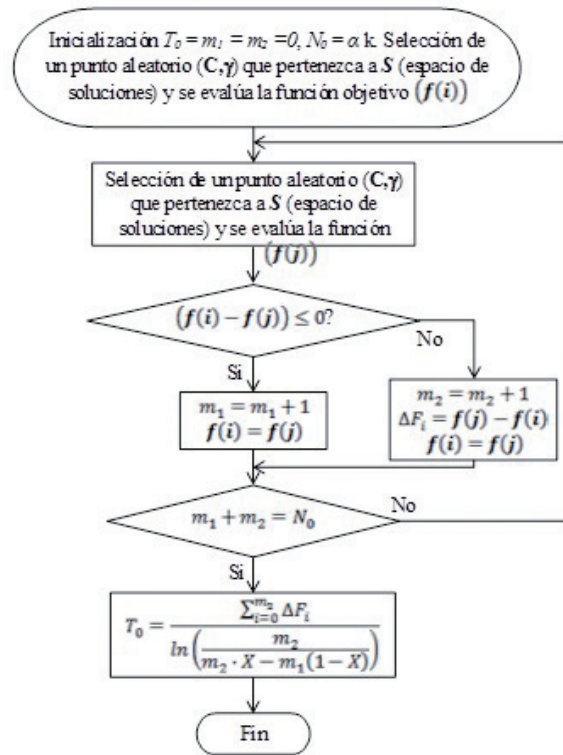


Figura 4. Diagrama de flujo para obtener la temperatura inicial para el SA

como se muestra en la Ec. (13) y se evalúa la función en cada punto. Se compara la solución de cada partícula con la iteración anterior de la misma partícula, y si ésta es mejor se actualiza. Se escoge la mejor solución actual y se compara con la incumbente, y si ésta es mejor se actualiza la incumbente y se verifica si cumple un criterio de parada.

Los límites propuestos en la Ec. (11) son los que presentaron mejores resultados en la etapa de parametrización.

En este artículo se utilizó $\phi_1=2$ y $\phi_2=2$ para la Ec. (12), los cuales son valores típicos de ajuste usados en anteriores investigaciones, Gómez (2008), Kennedy et al. (2001). También se utilizó $\omega_{\max}=0.9$ y $\omega_{\min}=0.4$ para la Ec. (14), ya que estos valores han proporcionado buen desempeño en varias aplicaciones, Eberhart & Shi (2001).

$$\omega^t = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{t_{\max}} \cdot t \quad (14)$$

3.1.4 Parametrización de la MSV con optimización por colonia de hormigas (ACO)

Como se presenta en Gil (2011), el ACO primero define los números de hormigas n , también se define el número de paso de cada hormiga m (el valor de n se varió en 4 y 13 y el de m entre 15 y 26, al final se utiliza un valor igual a 10 y 20, respectivamente, ya que con éstos valores el algoritmo presentó buen rendimiento para este problema). Se crean n hormigas aleatoriamente y cada una es una solución (C, γ) del problema que se está tratando. Se evalúa la función objetivo de cada hormiga y se guarda la mejor solución encontrada. Posteriormente se genera una vecindad para cada hormiga y se evalúa la función objetivo en la vecindad, y la nueva solución se escoge de acuerdo a las Ec. (15) y (16). El espacio de soluciones se discretizó con intervalos de 10^4 para el parámetro C , e intervalos de 8 para el parámetro γ . Después de que las hormigas cumplen con los pasos se escoge la mejor solución y se compara con la incumbente. Si ésta es mejor se actualiza la incumbente. Se verifica el criterio de parada, si no cumple se generan otra vez las hormigas cerca de la mejor solución. Los valores de la Ec. (16) se recomiendan en Gallego et al. (2008).

$$P_{ij}^k(t) = \begin{cases} \text{máximo } \{\tau_{ij}^a, \eta_{ij}^b\} & \text{si } q > q_0 \\ j \text{ caso contrario} & \end{cases} \quad (15)$$

$$j = \begin{cases} \text{Ruleta si } 0.5 \leq x \\ \text{Torneo si } 0.1 \leq x < 0.5 \\ \text{Aleatorio si } 0.0 \leq x < 0.1 \end{cases} \quad (16)$$

3.2 Etapa 2: Entrenamiento y validación de la MSV

El desarrollo de esta etapa se hace con todos los datos de fallas registrados en la subestación. A partir de esos registros se propone una validación cruzada ($n = 10$), ya que con este valor se pretende que la base de entrenamiento sea grande y la de validación pequeña, lo cual ocurre en la etapa de aplicación para un registro real. Para cada subconjunto de validación cruzada se halla la precisión según lo propuesto con la Ec. (17). Posteriormente se obtiene la media de los subconjuntos, a partir de los parámetros obtenidos para cada una de las combinaciones de datos de entrada.

$$\text{Precisión} = \frac{\text{Número de datos bien clasificados}}{\text{Número de datos de prueba}} \cdot 100\% \quad (17)$$

4. Aplicación de la metodología propuesta en un sistema de distribución prototipo

Esta metodología se desarrolló utilizando el software Matlab® para elaborar los algoritmos necesarios, el software de simulación de circuitos eléctricos ATP y el software *simulaciónRF*, propio del equipo de trabajo, Pérez et al. (2010), el cual simula automáticamente unas condiciones de falla deseadas y permite una rápida obtención de las bases de datos necesarias.

El problema cuadrático de la MSV se resuelve con el optimizador que se presenta en Ma et al. (2002).

4.1 Descripción del sistema de prueba y escenarios

El sistema de prueba es una modificación del IEEE 34 nodos con una alimentación de 24,9 kV, IEEE (2004). La modificación consiste en concentrar los nodos 888 y 890 como una carga en el nodo 832. En la figura 5 se muestra el sistema de prueba modificado y la zonificación (11 zonas).

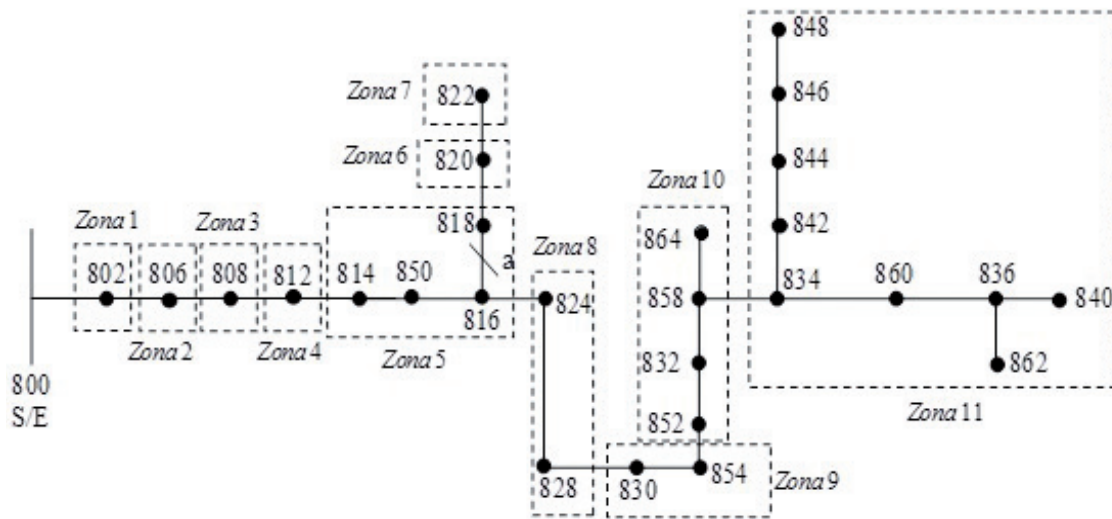


Figura 5. Sistema de pruebas IEEE de 34 nodos y zonificación para localización de fallas

4.2 Aplicación de la MSV a la localización de fallas

a. *Zonificación de la red.* La división del circuito se realiza en 11 zonas como se muestra en la figura 5.

b. *Adquisición de la base de datos de falla.* Se realizaron simulaciones en cada uno de los nodos considerando resistencias de falla que corresponden a los siguientes valores 1Ω , 5Ω , 10Ω , 20Ω , 30Ω , y 40Ω , los cuales son valores comúnmente utilizados en este tipo de pruebas, Dagenhart (2000). Se simularon 4 tipos de fallas diferentes, una monofásica (a-g), una bifásica (a-b), una bifásica a tierra (a-b-g) y una trifásica (a-b-c). Se realizó además la simulación para cuatro variaciones de carga diferentes, entre 10%-30%, 40%-70%, 80%-120%, y 130%-150% de la carga nominal. Para cada variación de carga se realizaron variaciones de la longitud de la línea y nivel de tensión de la fuente, entre 95-105% y 95-105% de la nominal, respectivamente. Todos los datos fueron utilizados ya sea en la parametrización, el entrenamiento o la validación en distintas pruebas. En total se obtuvieron 13824 registros de falla.

c. *Pre-procesamiento de la señal.* El conjunto de entrenamiento se normaliza para evitar

desequilibrios causados por variaciones extremas en los valores de descriptores. Esta normalización se encuentra en el intervalo $[0, 1]$ tal como se propone en la Ec. (4), Mora et al. (2008). Adicionalmente, a cada conjunto de descriptores se le asocia una etiqueta dependiendo de la zona en la cual ocurrió la falla.

4.3 Parametrización y pruebas del algoritmo de clasificación

En la etapa de parametrización sólo se utilizaron datos con variación entre 80%-120% de la carga nominal (que incluyen variaciones de la longitud de la línea y nivel de tensión de la fuente) y con resistencias de falla 1Ω , 20Ω y 40Ω . Los resultados obtenidos de cada técnica utilizada en la parametrización para fallas monofásicas se muestran en la tabla 1, donde se muestra el valor del parámetro C de la MSV y el valor del parámetro γ del kernel. Para los otros tipos de fallas también se determinaron los parámetros óptimos utilizando las técnicas propuestas.

En las tablas 2 y 3 se presenta el error de validación de la parametrización y la precisión del método para cada técnica utilizada, respectivamente. En la tabla 3 se presenta la precisión para cada tipo de falla, con cada una de las técnicas y utilizando

Tabla 1. Resultados de C y γ para fallas monofásicas a-g

Descriptores	TS		SA		PSO		ACO	
	C	γ	C	γ	C	γ	C	γ
$dV, dI, d\theta_v$	707539,5	19,4	432721,0	140,1	52770,6	99,4	4141000,0	40,0
$dV, dI, d\theta_l$	7142192,7	72,3	2599,6	89,7	3658902,8	54,4	5011000,0	136,0
$dV, d\theta_v, d\theta_l$	9992151,9	32,0	4140351,0	78,9	5526,2	22,1	1831000,0	136,0
$dI, d\theta_v, d\theta_l$	26631,9	81,5	9995482,1	82,1	33932,0	79,4	1971000,0	72,0
$dV, dI, d\theta_v, d\theta_l$	106623,9	75,4	448365,2	28,6	9945082,4	9,2	7301000,0	16,0
$dV_l, dI_l, d\theta_{vl}$	170061,6	71,7	4852,3	569,3	3684,2	175,5	591000,0	648,0
$dV_l, dI_l, d\theta_{ll}$	9965575,4	137,4	85740,3	534,0	5276652,6	190,0	6221000,0	272,0
$dV_l, d\theta_{vl}, d\theta_{ll}$	206513,7	12,7	2935013,7	39,6	8635717,6	29,7	8721000,0	40,0
$dI_l, d\theta_{vl}, d\theta_{ll}$	25543,1	14,2	1200607,2	556,8	592476,7	199,3	5501000,0	696,0
$dV_l, dI_l, d\theta_{vl}, d\theta_{ll}$	14571,1	97,6	8948950,3	178,3	2384015,2	128,9	3131000,0	120,0

Tabla 2. Error de validación cruzada para los localizadores ante diferentes técnicas de parametrización

Descriptores	Falla monofásica				Falla bifásica				Falla bifásica a tierra				Falla trifásica			
	TS	SA	PSO	ACO	TS	SA	PSO	ACO	TS	SA	PSO	ACO	TS	SA	PSO	ACO
$dV, dI, d\theta_v$	0,00	0,00	0,00	0,21	2,16	0,73	0,96	3,14	0,96	0,00	1,21	0,00	6,80	2,66	3,64	3,39
$dV, dI, d\theta_l$	0,00	0,00	0,00	0,00	0,24	0,00	0,24	0,96	0,73	0,73	0,24	0,97	1,19	1,91	0,00	0,97
$dV, d\theta_v, d\theta_l$	0,61	0,82	0,00	0,00	1,43	1,92	1,92	2,89	0,00	0,24	0,00	0,00	1,46	0,00	0,00	0,00
$dI, d\theta_v, d\theta_l$	0,41	0,20	0,41	0,41	0,73	0,48	0,97	1,45	0,00	0,00	0,00	0,00	0,24	0,00	0,96	0,00
$dV, dI, d\theta_v, d\theta_l$	0,00	0,41	0,00	0,00	0,97	0,00	0,97	1,21	0,49	0,00	0,00	0,00	0,00	0,00	0,00	0,71
$dV_l, dI_l, d\theta_{vl}$	1,85	1,65	1,23	2,27	3,37	5,53	1,92	0,72	1,70	2,63	0,97	2,17	0,49	1,45	1,21	0,49
$dV_l, dI_l, d\theta_{ll}$	2,06	2,47	3,50	1,86	0,72	2,18	0,73	0,97	1,46	0,73	0,49	0,49	0,97	0,72	0,97	0,24
$dV_l, d\theta_{vl}, d\theta_{ll}$	0,41	1,24	0,00	0,00	0,97	0,00	0,49	0,73	0,49	0,73	0,48	0,48	0,00	0,00	0,00	0,00
$dI_l, d\theta_{vl}, d\theta_{ll}$	0,21	0,00	0,82	0,62	1,69	0,72	0,73	0,49	0,96	0,73	0,95	0,72	0,73	0,00	0,00	0,00
$dV_l, dI_l, d\theta_{vl}, d\theta_{ll}$	0,82	0,83	0,62	0,21	1,44	0,72	0,73	0,73	0,97	0,48	0,24	0,73	0,00	0,24	0,00	0,00

10 combinaciones de descriptores En la precisión se utilizaron todas las variaciones de carga, línea y fuente del sistema. El número de registros total para falla monofásica es de 3888 y para cada tipo de fallas restantes es 3312.

Se puede notar en la tabla 3, que para fallas monofásicas la mayoría de las combinaciones de descriptores utilizadas proporcionaron buenos resultados. Adicionalmente, cada técnica metaheurística obtuvo una precisión del 100% en

al menos cuatro combinaciones de descriptores en este tipo de falla.

Se puede observar que para fallas bifásicas las combinaciones de descriptores de fase tuvieron mejor rendimiento que los de línea.

Finalmente, se observa que para fallas trifásicas, ACO obtuvo, en al menos ocho combinaciones de descriptores, precisiones mayores del 99.28%, mientras que el resto de técnicas utilizadas no

Tabla 3 Precisión obtenida con los localizadores ante diferentes técnicas de parametrización

Descriptores	Falla monofásica a-g				Falla bifásica a-b				Falla bifásica a tierra a-b-g				Falla trifásica a-b-c			
	TS	SA	PSO	ACO	TS	SA	PSO	ACO	TS	SA	PSO	ACO	TS	SA	PSO	ACO
$dV, dI, d\theta_v$	100,00	100,00	100,00	100,00	100	99,85	99,94	99,85	88,16	99,03	99,58	99,49	70,77	96,26	95,59	95,09
$dV, dI, d\theta_l$	99,97	100,00	98,97	100,00	90,37	82,31	89,67	99,82	92,78	89,10	88,13	94,23	70,23	60,78	78,86	99,55
$dV, d\theta_v, d\theta_l$	100,00	99,95	100,00	100,00	73,61	82,34	81,22	96,26	99,19	95,11	93,27	93,57	60,93	99,94	81,82	99,88
$dI, d\theta_v, d\theta_l$	100,00	100,00	100,00	100,00	85,9	95,9	99,82	99,15	81,64	85,48	83,33	85,11	64,61	99,52	64,16	99,7
$dV, dI, d\theta_v, d\theta_l$	99,69	99,77	99,95	99,95	99,88	99,76	99,52	99,79	94,02	99,94	97,89	84,15	99,40	74,00	99,55	99,61
$dV_l, dI_l, d\theta_{vl}$	100,00	99,36	99,56	99,43	66,31	96,47	90,82	99,52	92,45	99,7	99,55	99,55	96,35	97,19	97,62	97,13
$dV_l, dI_l, d\theta_{ll}$	99,97	100,00	98,59	99,19	60,78	60,81	60,81	60,81	54,92	93,90	90,61	86,26	65,79	99,43	76,15	99,67
$dV_l, d\theta_{vl}, d\theta_{ll}$	100,00	100,00	100,00	99,90	60,87	71,86	64,31	60,87	98,49	99,67	99,76	99,85	97,28	99,79	99,52	99,28
$dI_l, d\theta_{vl}, d\theta_{ll}$	100,00	100,00	100,00	99,95	71,29	61,56	64,4	60,87	73,4	99,97	99,76	99,88	99,73	99,64	97,56	99,61
$dV_l, dI_l, d\theta_{vl}, d\theta_{ll}$	99,69	100,00	99,97	99,92	60,87	60,87	93,69	60,87	99,52	99,61	99,34	96,86	99,61	96,98	99,73	99,61

tuvieron un buen rendimiento en la mitad o más combinaciones de descriptores utilizadas.

5. Conclusiones

Los mejores resultados de cada técnica metaheurística se presentan para la falla monofásica, lo cual es una ventaja debido a que esta falla se presenta con mayor frecuencia en los sistemas de distribución de energía eléctrica. Adicionalmente, la técnica metaheurística ACO mostró un mejor rendimiento en todos los tipos de falla en comparación con otras técnicas metaheurísticas. En todos los casos estudiados (combinaciones de descriptores con tipos de falla) el 70% de éstos presentó buenos resultados, con precisión superior al 98%.

Las técnicas metaheurísticas basadas en población utilizadas en este artículo presentaron en general un mejor desempeño en todos los casos estudiados con un 60% de buenos resultados (precisión > 98%), con respecto a las técnicas metaheurísticas basadas en trayectoria. Esto es debido a que las técnicas basadas en población pueden evaluar mucho más el espacio de soluciones que las técnicas basadas en trayectoria. La técnica metaheurística TS no presentó en general buen

desempeño, pues tan sólo se obtuvieron un 45% de éxito en todos los casos estudiados (precisión > 98%).

Finalmente, una precisión del 100% de la MSV ante una condición específica del sistema de distribución, no garantiza que el método no pueda tener clasificaciones erróneas ante otras condiciones de operación. Para estar seguro que el algoritmo tiene un alto desempeño, es necesario evaluar el clasificador con todas las posibles variaciones que se puedan presentar en un circuito real, como se propuso en este artículo para variaciones de la carga, de la tensión en la fuente y la longitud de línea. Sin embargo, la MSV tiene un buen coeficiente de generalización lo que implica que provee buenos resultados ante datos desconocidos, el cual lo hace útil en sistemas de distribución reales.

6. Referencias bibliográficas

Burges, C.J.C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*. 2(2), 121-127.

Dagenhart, J. (2000). "The 40-Ω ground-fault phenomenon" *IEEE Transactions on Industry Applications*, 36, 1, 30-32.

- Eberhart, R. & Yuhui S. (2001). *Particle Swarm Optimization: Developments, Applications and Resources*. Proceedings of the Congress on Evolutionary Computation, 81-86.
- Faber, R., Jockenhövel T. & Tsatsaronis, G. (2008). Dynamic optimization with simulated annealing. *Computers and Chemical Engineering* 29(2), 273–290.
- Gallego, R.A., Escobar, A.H. & Toro, E.M. (2008). *Técnicas Metaheurísticas de Optimización*. 2a. ed. Pereira: Texto Universitario.
- Gil, W.J. (2011). *Utilización de técnicas metaheurísticas en la búsqueda de parámetros óptimos para la calibración de las máquinas de soporte vectorial (MSV), para la localización de fallas en sistemas de distribución*. Tesis de pregrado, Programa de Ingeniería Eléctrica, Universidad tecnológica de Pereira, Pereira, Colombia
- Glover, F.W. & Kochenberger, G.A. (2002). *Handbook of Metaheuristics*. 1a. ed. EEUU: Kluwer Academic Publishers.
- Gómez, M. (2008). *Sistema de Generación Eléctrica con pila de Combustible de óxido sólido alimentado con Residuos forestales y su optimización mediante Algoritmos basados en nubes de partículas*. Tesis Doctoral. Departamento de Ingeniería Eléctrica, Electrónica y de Control, Universidad Nacional de Educación a Distancia Escuela Técnica Superior de Ingenieros Industriales, Madrid, España.
- Gutiérrez, J.A., Mora, J.J. & Pérez, S.M. (2010). Strategy based on genetic algorithms for an optimal adjust of a support vector machine used for locating faults in power distribution systems. *Revista de la Facultad de Ingeniería Universidad de Antioquia*, 53, 174-184.
- IEEE Std 37.114. (2004). IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines. *Power System Relaying Committee*, 1-36.
- Kennedy, J., Eberhart, R., & Shi, Y. (2001). "Swarm Intelligence". San Francisco: Morgan Kaufmann Publishers.
- Ma, J., Zhao, Y. & Ahatl, S. (2002). Osu svm classifier toolbox. Software available at <http://sourceforge.net/projects/svm/files/svm/>.
- Moguerza J.M. & Muñoz A. (2006). Support Vector Machines with Applications. *Statistical Science* 21(4), 322-336.
- Mora, J.J. (2006). *Localización de Fallas en Sistemas de Distribución de Energía Eléctrica usando Métodos basados en el Modelo y Métodos de Clasificación Basados en el Conocimiento*. Tesis Doctoral. Departamento Tecnologías de la Informática, Universidad de Girona, Girona, España.
- Mora, J.J., Meléndez, J. & Bedoya, J.C. (2006) Implementación de protecciones y simulación automática de eventos para localización de fallas en sistemas de distribución de energía. *Ingeniería y competitividad*, 8(1), 5-14.
- Mora, J.J., Meléndez, J. & Carrillo, G. (2008). Comparison of impedance based fault location methods for power distribution systems. *Electric Power Systems Research*. 78(4), 657-666.
- Morales, G.A. & Gómez, A. (2005). Estudio e Implementación de una herramienta basada en Máquinas de soporte vectorial aplicada a la Localización de fallas en sistemas de distribución. Trabajo de Grado. Facultad de Ingenierías Fisicomecánicas. Universidad Industrial de Santander, Santander, Colombia.
- Morales, G.A., Mora, J.J., & Pérez, S.M. (2008). Diseño conceptual de un localizador de fallas para sistemas de distribución usando minería de datos y análisis circuital. *Scientia et Technica*, 38, 89-94.
- Peréz, L.H., Mora, J.J., & Pérez, S.M. (2010). Diseño de una herramienta eficiente de simulación automática de fallas en sistemas eléctricos de potencia. *Dyna*, 77(164), 178-188.

Short, T: (2003). *Electric Power Distribution Handbook*. CRC press. 1, 8-30.

Rossi, O.D., Samples, M., Birattari, M., Chiarandini, M., Dorigo, M., Gambardella, L., Knowles, J., Manfrin, M., Mastrilli, M., Peachter, B., Paquete, L. & Stützle, T. (2002). Comparison of the Performance of Different Metaheuristics on

the Timetabling Problem. Springer-Verlag Berlin Heidelberg. 329-351

Thukaram, D., Khincha, H. & H. Vijaynarasimha, H. (2005). Artificial Neural Network and Support Vector Machine Approach for Locating Faults in Radial Distribution Systems. *IEEE Transactions on Power Delivery*, 20, 710-721.