

ALGORITMO SIMULATED ANNEALING APLICADO EN LA IDENTIFICACIÓN DE ERRORES DE DIFÍCIL DETECCIÓN EN ESTIMACIÓN DE ESTADO

Simulated Annealing Algorithm Applied to the Multiple Interacting Bad Data Errors Identification in State Estimation

RESUMEN

En este artículo se utiliza la técnica de optimización *Simulated Annealing* con el fin de resolver el problema de estimación de estado cuando este presenta errores de difícil detección. Es propuesto un algoritmo que combina la estimación clásica con la teoría de puntos de apalancamiento y resuelto usando un método de optimización combinatorial denominado *Simulated Annealing*. Los resultados reportados presentan un mejor desempeño al ser comparados con los obtenidos por un estimador de estado tradicional y de igual calidad a los reportados con métodos de optimización combinatorial tales como *Búsqueda Tabú*, *Algoritmo Genético de Chu-Beasley* y *Partículas Swarm*.

PALABRAS CLAVES: Estimación de estado, Optimización combinatorial, Puntos de apalancamiento, Simulated Annealing.

ABSTRACT

In this paper a Simulated Annealing optimization technique is used in order to solve the state estimation problem when a multiple interacting bad data appear in the classical formulation of the state estimation problem. The methodology used, propose an algorithm that combines the classical estimation with the leverage points theory and solved by using the Simulated Annealing optimization technique. Results obtained presents better performance when they are compared with the obtained using a classical state estimator, and the same when optimization techniques like Tabu Search, genetic algorithm of Chu-Beasley and Particle Swarm are used.

KEYWORDS: State estimation, combinatorial optimization, leverage points and Simulated Annealing.

1. INTRODUCCIÓN

La expansión e interconexión de los sistemas eléctricos de potencia ha dificultado a los centros de control de energía (CCE) la tarea de supervisar y operar de manera segura las redes eléctricas. Uno de los principales inconvenientes dentro del análisis en tiempo real de los sistemas de potencia consiste en la determinación del estado actual de la red cuando la información proporcionada por los medidores contiene inconsistencias, generalmente atribuidas a problemas en la calibración o al tiempo de funcionamiento de los mismos [1]. En consecuencia, se determina un punto de operación diferente al estado actual del sistema, siendo esta una condición crítica ya que a partir de estos resultados se derivan un conjunto de análisis en las redes eléctricas, como son estudios de flujo de carga, análisis de fallos, estabilidad, acciones de control, entre otros.

Los CCE disponen de una herramienta matemática para determinar el punto de operación de un sistema de potencia conocida como estimación de estado, capaz de procesar la información disponible en la base de datos

HUGO ANDRÉS RUIZ F.

Ingeniero Electricista.
Profesor Catedrático.
Facultad de Ingenierías.
Universidad Tecnológica de Pereira
hugor@ohm.utp.edu.co

ELIANA MIRLEDY TORO O.

Ingeniera Industrial, MS.c
Profesor Asistente.
Facultad de Ingeniería Industrial.
Universidad Tecnológica de Pereira
elianam@utp.edu.co

RAMÓN ALFONSO GALLEGO R.

Ingeniero Electricista, Ph.D
Profesor Titular.
Programa de Ingeniería Eléctrica.
Universidad Tecnológica de Pereira
ragr@utp.edu.co

relacionada con la localización e interconexión de los componentes del sistema, el conjunto redundante de datos suministrado por el sistema SCADA (*Supervisory, Control and Data acquisition*) y la información basada en registros históricos. Una vez conocidas las variables de estado del sistema y antes de aplicar alguna acción de control, la metodología clásica de estimación de estado, denominada WLS (*Weight Least Square*), a través de un procedimiento estadístico y probabilístico somete los datos de entrada a un proceso de validez con el fin de identificar las mediciones que contienen inconsistencias para su posterior eliminación del conjunto de datos, garantizando de esta forma una correcta estimación de las variables de estado [2].

Existen tipos de error caracterizados por su difícil detección por parte de la técnica WLS, tal es el caso de los errores de tipo iterativo y conformativo [3] [4]. El análisis de este tipo de error se ha desarrollado mediante la teoría de puntos de apalancamiento en sistemas eléctricos. Datos con esta característica propician la eliminación de mediciones correctas. Lo anterior implica

una pérdida de información que afecta la calidad de los resultados estimados por la disminución en la redundancia de los datos, impidiendo en algunos casos la determinación del punto de operación del sistema, debido a que el problema de estimación se transforma en un problema de observabilidad.

Dado lo anterior se modela el problema de estimación de estado cuando existen múltiples mediciones de difícil detección como un problema de optimización a través de la técnica de *Simulated Annealing* usando como factor de sensibilidad metodologías de detección de puntos de apalancamiento. Se usa como sistema de prueba el sistema IEEE de 57 nodos obteniendo resultados satisfactorios.

2. ESTIMACIÓN DE ESTADO WLS

La estimación de estado a través de su metodología clásica WLS se ha utilizado para realizar el mejor ajuste de los datos. En el modelamiento matemático de dicho problema, el error estimado (e) se calcula como la diferencia entre el valor medido (z) y el calculado:

$$e = z - h(x) \quad (1)$$

donde $h(x)$ representa el conjunto de ecuaciones no lineales del sistema. Para cuantificar el error presente al realizar la estimación se emplea la ecuación de mínimos cuadrados ponderados:

$$J(x) = \sum_{i=1}^{nm} W_{ii} \cdot e_i^2 \quad (2)$$

siendo W una matriz de ponderaciones que permite darle un peso preferencial a las mediciones más exactas. Al derivar (2) con respecto a las variables de estado del sistema y siguiendo con el procedimiento presentado en [5] se obtiene la expresión empleada para la actualización de las variables de estado del sistema a través de un procedimiento iterativo:

$$\hat{x} - x = G^{-1} H^T W e \quad (3)$$

2.1 PRUEBAS ESTADÍSTICAS

El primer procedimiento se conoce con el nombre de *Prueba chi-cuadrado* (χ^2) que consiste en una prueba de hipótesis de una cola donde el resultado calculado en (2) se compara con el resultado obtenido de valores tabulados de la χ^2 con ν grados de libertad y nivel de confianza α .

Para determinar la presencia de datos erróneos en el proceso de estimación se debe cumplir que $J(\hat{x}) \geq \chi^2_{\alpha}$. En caso de identificar presencia de error, se procede a

detectar las mediciones que lo ocasionan, este procedimiento se conoce con el nombre de *Prueba t-student* (t_{α}), el cual permite la detección y eliminación de la medición que presenta mayor influencia negativa dentro del proceso de estimación.

Para identificar las mediciones con error cada uno de los errores se debe normalizar, para ello, cada uno de estos se divide por su respectiva desviación, de la siguiente forma:

$$\hat{e}_i^N = \frac{|\hat{e}_i|}{\sqrt{R_{ii}}} \quad (4)$$

Posteriormente se debe comparar el máximo valor del residuo normalizado con el valor de t_{α} . El procedimiento se repite hasta no detectarse presencia de datos erróneos.

3. IDENTIFICACIÓN DE ERRORES DE DIFÍCIL DETECCIÓN

La técnica clásica de estimación WLS en diversas implementaciones presenta resultados satisfactorios. Sin embargo, existen mediciones erradas de difícil detección como aquellas que poseen errores de tipo iterativo y conformativo, conocidas también como *puntos de apalancamiento*, en donde la metodología WLS presenta inconvenientes para su detección, debido a que durante la identificación de estas se elimina innecesariamente datos que no contienen errores, disminuyendo el nivel de redundancia, ocasionando en algunos casos problemas de observabilidad, produciendo inconvenientes en la determinación del estado del sistema.

3.1 PUNTOS DE APALANCAMIENTO

En estimación de estado, un punto de apalancamiento se define como una medición con alto grado de error, caracterizado por su difícil detección por parte de la metodología clásica WLS y su influencia negativa sobre las mediciones sin presencia de error ubicadas en regiones cercanas. Por consiguiente, los puntos de apalancamiento evitan la correcta identificación de las mediciones erróneas, debido a que sus residuos normalizados son pequeños, propiciando de esta manera la eliminación de medidas correctas.

Dentro de la teoría de puntos de apalancamiento aplicada al análisis de mediciones en sistemas eléctricos, se han desarrollado metodologías que permiten conocer las medidas dentro de un sistema de potencia, que en caso de presentar errores considerables, la técnica WLS tendría inconvenientes en su detección. Las metodologías de detección de puntos de apalancamiento se derivaron del trabajo propuesto por Abur-Celik [6] y se presentan a continuación:

3.1.1 Método 1

Este procedimiento propuesto en [6] y [7] se fundamenta en el análisis de la *matriz sensitiva residual* o *matriz de varianza de los errores estimados modificada* $(I-\tilde{P})$ que resulta de aplicar las transformaciones del modelo de Abur-Celik. Esta matriz cumple con la propiedad que los elementos de la diagonal principal son iguales al cuadrado de la respectiva fila de la matriz.

$$(I-\tilde{P}) = \begin{bmatrix} (I-\tilde{P})_{11} & (I-\tilde{P})_{12} & \dots & (I-\tilde{P})_{1m} \\ (I-\tilde{P})_{21} & (I-\tilde{P})_{22} & \dots & (I-\tilde{P})_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ (I-\tilde{P})_{m1} & (I-\tilde{P})_{m2} & \dots & (I-\tilde{P})_{mm} \end{bmatrix} \quad (5)$$

Los elementos fuera de la diagonal representan la influencia que tienen los residuos de la medición i sobre los residuos de la medición j . Para detectar los puntos de apalancamiento se deben identificar cuales de las mediciones se encuentran afectadas por un conjunto de mediciones específicas denominado S_i^v $i=1, \dots, m$. El procedimiento general se presenta a continuación:

- Calcular el Jacobiano modificado $\tilde{H} = W^{1/2} H$.
- Calcular la matriz $(I-\tilde{P}) = I - [\tilde{H}(\tilde{H}'\tilde{H})^{-1}\tilde{H}']$.
- Calcular para cada medición el subíndice γ .

$$\gamma_i = \sqrt{\frac{(I-\tilde{P})_{ii} - (I-\tilde{P})_{ii}^2}{m-1}} \quad (6)$$

siendo m el número total de mediciones.

- Determinar el conjunto de mediciones S_i^v que presentan una influencia significativa sobre la medición i . Para ello, se compara el valor del subíndice γ_i con cada uno de los elementos de $(I-\tilde{P})_{ij}$ correspondientes a la i -ésima fila de esta matriz. Si $\gamma_i < |(I-\tilde{P})_{ij}|$, $j=1, \dots, m$ entonces la medición j forma parte del conjunto S_i^v .
- Por último, si para la medición i el valor de $(I-\tilde{P})_{ii}$ es menor a todos los valores de $(I-\tilde{P})_{ij}$, $\forall j \in S_i^v$, la medición i es un punto de apalancamiento.

3.1.2 Método 2

Esta metodología propuesta en [8] analiza los elementos de la diagonal de la matriz de P calculada en estimación de estado. Los elementos P_{ij} representan la influencia de

la medición j sobre el valor estimado de la medición i y los elementos P_{ii} ($0 \leq P_{ii} \leq 1$) representan la influencia de la i -ésima medición sobre la i -ésima medición estimada. Si esta influencia es alta, es decir si $P_{ii} \approx 1$, entonces la medición es llamada punto de apalancamiento, y su valor residual será pequeño haciendo difícil la tarea de detección de datos erróneos.

Las condiciones que usualmente crean puntos de apalancamiento en los sistemas de potencia son las siguientes:

- Mediciones de flujo y/o inyecciones en las terminales de líneas relativamente cortas.
- Mediciones de inyección en nodos que tienen un gran número de vecinos inmediatos.

Con base en lo anterior esta metodología considera las siguientes reglas heurísticas para determinar la existencia de puntos de apalancamiento si:

- a) Existen mediciones que cumplen con $P_{ii} > \mu_1$.
- b) La reactancia de líneas conectadas a un nodo es menor o igual que μ_2 .
- c) El número de vecinos de un nodo es mayor o igual que μ_3 .

donde $\mu_1 = 2n/m$, n es el número de variables de estado y m el número de mediciones., μ_2 es la reactancia promedio de todas las líneas de la red y μ_3 es el promedio de nodos vecinos de la red.

Una medición será considerada punto de apalancamiento si mantiene las siguientes condiciones:

- Si los incisos (a) y (b) anteriores cumplen para una medición de flujo de potencia.
- Si los incisos (a) y (b) ó (a) y (c) anteriores cumplen para una medición de inyección de potencia.

4. METODOLOGÍA DE SOLUCIÓN

4.1 CODIFICACIÓN

Las variables de un problema de optimización pueden ser representadas a través de valores enteros, reales o binarios. Dado que en estimación de estado inicialmente se cuenta con un conjunto de mediciones, donde algunas de ellas en determinado momento pueden ser o no incluidas en el proceso, se utiliza una codificación de tipo binario. Por tanto, los posibles estados de una medición i se definen por una variable d_i que toma los siguientes valores:

$d_i = 0$, si la i -ésima medición es incorrecta.
 $d_i = 1$, si la i -ésima medición es correcta.

Teniendo en cuenta lo anterior, una alternativa de solución es representada de la siguiente forma:

$$d_k = \begin{array}{|c|c|c|c|c|c|} \hline & 1 & 2 & 3 & 4 & \dots & nm \\ \hline & 0 & 1 & 1 & 1 & \dots & 0 \\ \hline \end{array}$$

Figura 1. Codificación del problema

4.2 ALGORITMO SIMULATED ANNEALING

El recocido simulado (SA, "Simulated Annealing") es una técnica metaheurística introducida a comienzos de los 80 por Kirkpatrick y Vecchi, e inspirada en un procedimiento físico de templado de metales usado en la metalurgia para llevar un sólido a un estado de equilibrio térmico. El recocido consta de tres etapas: la primera consiste en una fase de calentamiento a una temperatura determinada que depende del tipo de material y de su grado de deformación; en la segunda fase se mantiene la temperatura, permitiendo de esta manera, una reorganización molecular en estados de mínima energía. Finalmente se realiza la fase de enfriamiento controlado, disminuyendo gradualmente la temperatura, hasta lograr un sólido cuyas partículas se encuentran en un estado cristalino perfecto. Al iniciar la etapa de enfriamiento, para cada valor de temperatura debe permitirse el alcance del equilibrio térmico. De lo contrario el sólido presentará una estructura amorfa en lugar de la estructura cristalina de más baja energía.

El procedimiento que dio origen a esta metodología de optimización se denomina algoritmo de Metrópolis propuesto alrededor de 1953, con el que se estudian las propiedades de equilibrio en el análisis del comportamiento microscópico de los cuerpos. El algoritmo de Metrópolis se basa en técnicas de simulación de Monte Carlo generando una secuencia de estados de un sólido, es decir, dado un sólido en un estado i y con energía E_i , se genera el siguiente estado j con energía E_j , mediante la aplicación de un mecanismo que lo conduce al estado siguiente a través de una pequeña perturbación. Si la diferencia de energía $E_j - E_i$ es menor o igual a cero, el estado j es aceptado. Si la diferencia de energía es mayor que cero, el estado j es aceptado con cierta probabilidad, la cual está dada por:

$$P_{aceptacion} = \exp\left(\frac{E_j - E_i}{k_b T}\right) \quad (7)$$

donde T representa la temperatura y k_b es la constante de Boltzmann. Esta regla de aceptación se conoce como *criterio de Metrópolis*.

De forma análoga el criterio de aceptación empleado en el algoritmo SA determina si j se acepta a partir de i al comparar los valores de la función objetivo (f) y aplicando la siguiente probabilidad de aceptación:

$$P_{aceptacion} = \begin{cases} 1 & \text{si } f(j) \leq f(i) \\ \exp\left(\frac{f(i) - f(j)}{T}\right) & \text{si } f(j) > f(i) \end{cases} \quad (8)$$

Se puede apreciar en (8) que la probabilidad de aceptar una nueva configuración de peor calidad, resulta de comparar la expresión correspondiente con un número aleatorio generado dentro de una distribución de probabilidad uniforme en el intervalo $[0,1]$.

Los parámetros que intervienen en el algoritmo son presentados a continuación:

4.2.1 Temperatura inicial

El valor inicial de temperatura T_o se obtiene de forma constructiva, simulando las transiciones hechas al inicio del proceso para la primera cadena de Markov.

El concepto físico de temperatura dentro del algoritmo SA no presenta un significado real, sino que ha de ser considerado como un parámetro por calibrar. Existen diversas formas de calcular la temperatura inicial, una de las más utilizadas se describe a continuación:

- Generar una alternativa de solución.
- Inicializar $T_o = 0$.
- Ejecutar la cadena m_o .
- Crear una nueva alternativa.
- Si $(f(i) - f(j)) \leq 0$ entonces $m_1 = m_1 + 1$ y aplique la fórmula:

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

En caso contrario, $m_2 = m_2 + 1$ y aplique la fórmula para el cálculo de T_o .

- Si, $m_0 = m_1 + m_2$ terminó la cadena y el valor de T_o (calculado en la última iteración) se asume como el valor inicial de temperatura.

donde m_1 es el número de transiciones propuestas de i a j para las cuales $f(i) \leq f(j)$, m_2 es el número de transiciones propuestas de i a j para las cuales

$f(i) > f(j)$, $\Delta F +$ es el incremento medio en el costo de las m_2 transiciones, calculado como:

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

4.2.2 Longitud de la cadena

A medida que la temperatura baja, también lo hace la probabilidad de aceptación y es necesario aumentar el número de posibilidades a evaluar. Este número de posibilidades en cada paso del proceso iterativo se conoce con el nombre de cadena de Markov o número de tentativas N_k . La longitud de la cadena N_k define el número de vecinos por explorar en el nivel k de temperatura T_k . Como regla general se maneja una longitud creciente a medida que la temperatura disminuye, con el fin de alcanzar el equilibrio de energía y minimizarla para cada estado de temperatura.

La expresión utilizada para actualizar el valor de la longitud de la cadena se presenta a continuación:

$$N_{k+1} = \rho \cdot N_k \quad (11)$$

donde ρ es un valor fijo para el incremento de la cadena.

4.2.3 Tasa de enfriamiento

Corresponde al porcentaje de disminución de la temperatura, afectando la probabilidad de aceptación de soluciones de peor calidad. El cálculo de la temperatura en el nuevo ciclo iterativo puede efectuarse de la siguiente manera:

$$T_{k+1} = \eta \cdot T_k \quad (12)$$

El valor de η es recomendado en la literatura especializada en un rango de [0.8, 0.99].

4.2.4 Definición de vecindad

La estructura de vecindad es determinante para el buen desempeño del algoritmo. Para esta aplicación el esquema de vecindad es el siguiente:

- Integrar una medida en el vector solución.
- Retirar una medida en el vector solución.
- Intercambiar dos medidas en el vector solución. Ingresar y retirar una medida.

Todas las alternativas cuentan con igual probabilidad de ocurrencia. El algoritmo inicialmente intenta el ingreso; de ser aceptada regresa a otra propuesta de vecindad, en caso contrario, intenta retirar y así sucesivamente.

4.2.5 Criterio de parada

Existen diferentes procedimientos para definir el criterio de parada:

- Fijar un número determinado de niveles de temperatura a analizar.
- Detener la búsqueda al finalizar el nivel de temperatura T_k si en este no se obtiene una solución de mejor calidad que la encontrada en los niveles anteriores.
- Detener la búsqueda cuando en el último nivel de temperatura no se haya encontrado un número mínimo de aceptaciones.
- Ejecutar el proceso un número especificado de iteraciones.

4.2.6 Descripción del algoritmo SA

A continuación se presenta el diagrama de flujo correspondiente al algoritmo SA general empleado para un problema de minimización.

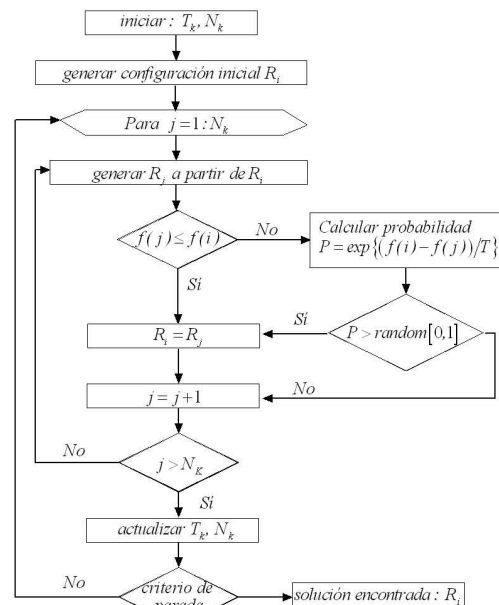


Figura 2. Descripción del algoritmo

5. RESULTADOS

Para la validación de la metodología propuesta se emplea como sistema de prueba la red IEEE de 57 nodos, 65 líneas y un total de 390 mediciones. Realizado un análisis de la información se encontraron 62 mediciones clasificadas como posibles puntos de apalancamiento. De estas se alteraron 20 para conformar una base de datos. Para el estudio se corrieron múltiples casos, en los cuales participan 20 o menos de las mediciones que fueron alteradas. Los valores de los parámetros del SA empleados en la simulación se presentan en la tabla 1.

T_o	75
N_o	80
η	0.8
ρ	1.02
T_f	5

Tabla 1. Parámetros usados en la simulación

La primera prueba se realiza utilizando la metodología clásica WLS, en esta se incurre en problemas de inadecuada identificación y observabilidad. En la segunda prueba se utiliza el algoritmo SA sin hacer uso de las metodologías de puntos de apalancamiento como factores de sensibilidad. En esta se logra identificar adecuadamente las mediciones de difícil detección, sin embargo, el número de iteraciones requerido es demasiado alto.

En las pruebas 3 y 4 se emplea el algoritmo SA junto con las metodologías de detección de puntos de apalancamiento como factores de sensibilidad, logrando la identificación de las mediciones de difícil detección y disminuyendo en menos del 50% el número de iteraciones empleadas.

Dado que el algoritmo SA aplicado con los factores de sensibilidad por separado proporciona una mejora en el algoritmo, se realiza una cuarta prueba en donde se realiza una implementación simultánea obteniendo el óptimo global en 158 iteraciones promedio. Los resultados comparativos se aprecian en la tabla 2.

Procedimiento	Iteraciones Promedio	Iteraciones mejor solución encontrada
SA sin sensibilidad	817	613
SA con método 1	342	211
SA con método 2	397	229
SA con método 1 y 2	158	95

Tabla 2. Comparación entre metodologías

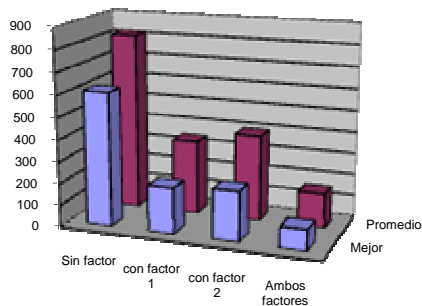


Figura 3. Número de iteraciones por método

6. CONCLUSIONES Y RECOMENDACIONES

La metodología clásica de estimación WLS no logra identificar correctamente las mediciones que presentan error cuando existen errores de tipo iterativo y conformativo. En todos los casos de prueba se presenta una inadecuada identificación de medidas, al retirar del proceso medidas correctas.

A pesar que la metodología de optimización combinatorial trabajando sin hacer uso de los factores de sensibilidad identifica correctamente las medidas de difícil detección, requiere de un número relativamente grande de iteraciones.

Cuando la metodología hace uso de los factores de sensibilidad trabajando independientemente, identifica correctamente las medidas de difícil detección y requiere de un número de iteraciones similares. Cuando existe un trabajo cooperativo en el cual se emplea de forma simultánea los factores de sensibilidad, el número de iteraciones disminuye ostensiblemente.

El método de optimización combinatorial usado en combinación con factores de sensibilidad, muestra ser promisorio para resolver problemas de estimación de estado de sistemas eléctricos de gran complejidad matemática.

7. BIBLIOGRAFÍA

- [1] Grainger, J.; Stevenson, W. D. Análisis de Sistemas de Potencia, Mc Graw Hill, México 1998.
- [2] A. Monticelli. State estimation in electric power systems – A generalized Approach. Kluwer Academic Publishers. pp. 227-265 1999.
- [3] A. Monticelli; F. Wu; M. Yen. Multiple bad data identification for state estimation by combinatorial optimization. IEEE Transactions on Power Apparatus and Systems, vol. 1, pp 361-369, 1986.
- [4] E. Asada; A. García; R. Romero. Identifying multiple interacting bad data in power system state estimation. IEEE Power Engineering Society General Meeting, vol. 1, pp. 571-577, 2005.
- [5] H. Ruiz; E. Toro; H. Salazar. Algoritmo genético modificado Chu-Beasley aplicado a la identificación de errores en la estimación de estado de sistemas eléctricos. Revista Scientia et Técnica, Ago 2007, No 37. pp. 25-30.
- [6] A. Abur; M. Celik. A robust WLAV state estimator transformations. IEEE Transaction on Power Systems, 1992, vol. pp. 106-113.
- [7] J. London; L. Alberto; L. Fabio; N. Bretas. Topologic análisis of undetectable errors in power system state estimation. X symposium of specialists in electric operational and expansion planning.
- [8] D. Romero; J. Ruiz; J. Robles. Estimación robusta para sistemas de potencia. IEEE RVP-91. Tomo 1, Jul 1991, pp 241-247.