

Atributos Relevantes para el Diagnóstico Automático de Eventos de Tensión en Redes de Distribución de Energía Eléctrica

Victor Barrera Núñez ^{a,*}, Ronald Velandia ^b, Fredy Hernández ^b, Joaquim Meléndez ^a, Hermann Vargas ^b

^a Instituto de Informática y Aplicaciones, Universitat de Girona, Campus Montilivi, 17003, Girona, España

^b Escuela de Ingeniería Eléctrica, Electrónica y Telecomunicaciones, Universidad Industrial de Santander, Carrera 27 Calle 9, Bucaramanga, Colombia

Resumen

En este trabajo se aborda el diagnóstico de eventos o perturbaciones de tensión registradas en subestaciones de distribución. La aparición de dichos eventos se debe a causas diversas que van desde faltas en la red, el arranque de motores de inducción, energización de transformadores y conmutación de bancos de capacitores. Se propone la caracterización de estos eventos de tensión a partir de atributos extraídos directamente de la forma de onda, y que se relacionan con los fenómenos físicos asociados tanto con las causas de los eventos, como con su localización relativa respecto del punto de medida. Se ha estudiado la relevancia de dichos atributos mediante un análisis estadístico de la varianza (MANOVA). Los atributos más relevantes se han utilizado para la obtención de reglas de clasificación mediante algoritmos de aprendizaje automático. Los resultados fueron obtenidos empleando datos de 484 eventos reales y 38 eventos simulados. Copyright © 2013 CEA. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Palabras Clave:

Análisis estadístico, Calidad de la potencia eléctrica, Atributos, Eventos de tensión, Sistema basado en reglas.

1. Introducción

Desde hace algunos años se ha incrementado la necesidad de monitorizar la “calidad de la energía eléctrica”. La energía suministrada por centros de generación, fluye por redes de transporte y distribución eléctrica hasta los puntos de consumo, y debe hacerlo bajo estrictos criterios de calidad. No obstante, los procesos propios del sector eléctrico (generación, transmisión y distribución), la operación de las cargas conectadas a la red eléctrica y el efecto de factores externos (aspectos atmosféricos, animales, accidentes, etc.) pueden distorsionar la forma de onda afectando su calidad. Según la forma de la distorsión de la onda, los problemas de la calidad de onda suministrada son diversos, pudiéndose destacar entre ellos: interrupciones, armónicos fluctuaciones de tensión, subidas de tensión, huecos de tensión, entre otros (Bollen, 2000).

Debido al impacto de los huecos de tensión en las cargas industriales y los costos producidos por las averías, gastos de mantenimiento y penalización, el sector industrial (tanto consumidores como distribuidores) ha prestado especial atención a su estudio, diagnóstico y mitigación (McGranaghan et al., 2002). Un ejemplo de ello es el esfuerzo realizado por las

compañías eléctricas por instalar monitores de calidad y almacenar estos eventos para su posterior estudio con el propósito de desarrollar métodos que permitan el diagnóstico (determinación de sus causas y localización del origen) automático de dichas perturbaciones (USA Energy Department, 2003, McGranaghan, 2001, McEachern, 2001).

Las causas de los eventos de tensión se pueden clasificar en dos grandes grupos, internas y externas, dependiendo si el agente que causa la perturbación pertenece o no a la red de distribución de energía. Las causas internas se deben a la operación de elementos propios de la red, tales como, energización de transformadores/capacitores, arranque de motores de inducción, cambios de carga, etcétera (Styvaktakis, 2002), mientras que las causas externas se asocian con perturbaciones generadas por agentes externos al sistema de energía eléctrica, como lo son: descargas atmosféricas, contacto de las redes aéreas con animales y ramas de árboles, ruptura de empalmes y pérdida de aislamiento por la humedad en cables subterráneos debido a la inundación de las canalizaciones (Kulkarni et al., 2010, Barrera et al., 2010b, d).

El estudio de las formas de onda de estos eventos y su clasificación es un tema de gran actualidad y se ha discutido en trabajos anteriores. Así por ejemplo, se pueden clasificar las

* Autor en correspondencia.

Correos electrónicos: vbarrera@eia.udg.edu (V. Barrera),

URL: <http://eia.udg.edu/~vbarrera> (V. Barrera)

perturbaciones por el aspecto de las formas de onda de las tensiones eficaces (Styvaktakis, 2002) como: rectangulares, no rectangulares y escalones de tensión (Tabla 1 y Figura 1). Típicamente se asocian las formas de onda no rectangulares y los escalones de tensión con causas internas. Las primeras habitualmente corresponden a arranque de motores y energización de transformadores (no rectangulares), mientras que los escalones de tensión suelen asociarse con conmutaciones de cargas y bancos de capacitores. Las formas de onda rectangular se asocian con faltas de red (por tanto causas externas) seguidas de una pérdida de carga debido a la operación de los equipos de protección.

Tabla 1. Eventos de tensión en redes de distribución

Forma	Causa	Int-erna	Ext-erna	Balan-ceada	Desba-lanceada	Tipo falta
No rectangular	ET	✓			✓	--
	M	✓		✓		3φ
Rectangular	FR		✓		✓	1φ
			✓		✓	2φ
			✓	✓		3φ
Escalones de tensión	C	✓		✓		3φ
	CC	✓		✓		3φ

Nota: ET: Energización de transformador (Figura 1e), M: Arranque de motor de inducción (Figura 1c), FR: Falta de red (Figura 1a, b, c) C: Energización de capacitor (Figura 1f), CC: Conmutación de cargas.

El análisis de eventos causados por factores externos es más complejo que el análisis de eventos causados por factores internos. Esto se debe a que el fenómeno eléctrico ocurrido durante el arranque de motores, energización de transformadores/capacitores y cambios de carga es claramente conocido (Figura 1 d, e, f). Por el contrario, las faltas en la red las producen fenómenos diversos no siempre conocidos (contacto de la red con animales/árboles, excavadoras o condiciones atmosféricas adversas, entre otras) en el momento de su ocurrencia y difícilmente cuantificables (Figura 1 a, b, c). El uso de métodos de aprendizaje automático permite establecer criterios objetivos para su clasificación de acuerdo con la causa (árbol, animal, rayo, cable, etc) tal como se ha demostrado en trabajos previos de los autores (Barrera et al., 2010a, b, c, d) y por tanto se han dejado fuera de los alcances de este artículo.

En la literatura existen distintos tipos de métodos para el diagnóstico de eventos de tensión. Algunos se centran en la identificación del número de fases involucradas en el evento: monofásica, bifásica, trifásica, con y sin contacto a tierra, etc (Bollen, 2003), un gran número en la identificación del tipo de perturbación: hueco de tensión, subida de tensión, transitorios, parpadeos, etc (Ibrahim, et al. 2002), y otros en la identificación de la causa interna (Seon et al. 2007) o externa que origina al evento (Xu et al. 2006, 2007) (Yixin et al. 2010a, b). En (Seon et al. 2007) se presenta un método para la identificación de eventos causados por energización de transformadores, arranque de motores y fallas de red, pero esta carece de un análisis estadístico de los atributos y tampoco es validada extensivamente con datos reales y sintéticos. Xu y Yixin han contribuido recientemente en la identificación de eventos causados por animales, árboles y rayos; para tal propósito en lugar de utilizar las formas de onda, utilizan únicamente información referente a las condiciones meteorológica (estación, fecha, etc), operación de los equipos de protección (número de fases, protección activada) y tipo de línea (aérea/subterránea). Un sistema experto para la identificación de la causa de eventos de tensión es propuesto en

(Styvaktakis, 2002), el sistema es validado adecuadamente con gran cantidad de medidas reales, éste es capaz de identificar a partir de la forma de onda de tensión y corriente, aquellos eventos causados por motor, transformador, fallas de red, interrupciones, interrupciones causadas por fallas, entre otras.

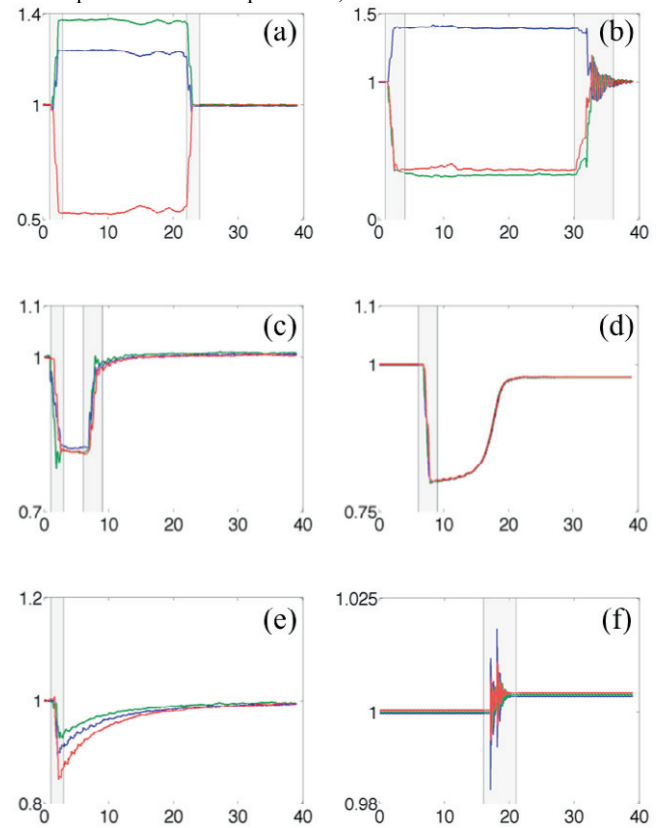


Figura 1. Tensión eficaz (p.u) Vs tiempo (ciclos). (a) Falta monofásica, (b) falta bifásica, (c) falta trifásica, (d) arranque de motor de inducción, (e) energización de transformador y (f) energización de banco de capacitores. Eventos a), b), c) y e) son reales; d) y f) son simulados, detalles sobre su captura y simulación en Sección 2.

Por lo que se refiere a la localización del origen del evento en la red, se deben diferenciar dos etapas claras. Por un lado, la localización relativa (Figura 2), aguas arriba o aguas abajo del punto de medida; y por otro, la estimación exacta de la distancia.

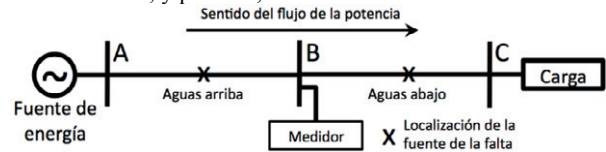


Figura 2. Diagrama unifilar simplificado de una red de distribución de energía eléctrica.

En este artículo se aborda sólo la localización relativa, dado que para la segunda etapa se requiere conocimiento adicional de la topología y los parámetros eléctricos de la red de distribución. Los autores no cuentan con esta información.

El objetivo de este trabajo es revisar los posibles atributos que pueden extraerse de las formas de onda y analizar su relevancia en la discriminación de faltas internas y externas. Para ello, se

revisan algunos atributos previamente propuestos en la literatura y otros que por su dependencia con los fenómenos físicos que intervienen en la generación y propagación del hueco de tensión se cree pueden ser relevantes.

Para las causas internas, además, se analiza la capacidad de dichos atributos para discriminar entre diferentes causas (motor, transformador, capacitor, Figura 1 d, e, f respectivamente), mientras que en el caso de causas externas, el estudio se limita a determinar las fases que intervienen en la falta (monofásica, bifásica y trifásica, Figura 1 a, b, c respectivamente). De manera similar, el estudio incluye también atributos que puede aportar información sobre la localización relativa de la fuente de la perturbación (aguas arriba y abajo) y se analizan la cantidad de estados transitorios presentes en la perturbación (evento monoestado y multiestado) teniendo en cuenta los tres tipos de eventos: rectangulares, no rectangulares y escalones de tensión (Figura 1). Se ha utilizado el análisis multivariante de la varianza (MANOVA) como herramienta para determinar la relevancia de cada atributo.

Una vez establecida la relevancia de los atributos, se propone el uso de algoritmos de extracción de reglas para obtener las reglas que permitan establecer una clasificación automática. Se han analizado seis algoritmos diferentes para tal propósito, resultando al final un conjunto de reglas capaz de distinguir entre eventos: balanceados/desbalanceados, monofásicos/bifásicos/trifásicos, multiestado/monoestado, energización de transformadores/capacitores y arranque de motores de inducción. Se han utilizado seis algoritmos con la intención de analizar la frecuencia de aparición de cada atributo en las reglas extraídas, esto ha permitido valorar cualitativamente la relevancia de los atributos, ya que cuanto mayor es su frecuencia de aparición, mayor será la relevancia del atributo.

En total se analizaron 46 atributos, de los cuales 38 describen la causa y los 8 restantes la localización relativa de la fuente de los eventos de tensión. Los atributos fueron obtenidos de eventos reales y simulados.

2. Eventos de Tensión Utilizados en el Análisis

Los 484 eventos utilizados en el análisis fueron registrados en el secundario de transformadores de 25 kV instalados en la cabecera de subestaciones de distribución suministrando energía a circuitos aéreos, subterráneos y aéreo-subterráneos. Los transformadores en estas subestaciones poseen conexión estrella-triángulo con primario solidamente aterrizado y secundario aterrizado a través de transformador Zig-Zag, ver Figura 3 (la flecha indica la ubicación del equipo de registro en la red de distribución). Los eventos tienen una frecuencia de oscilación de 50Hz, 40 ciclos de duración y 128 muestras por ciclo.

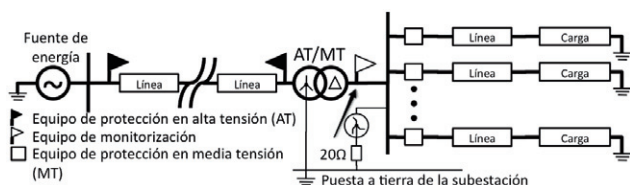


Figura 3. Diagrama unifilar de las subestaciones de distribución donde se registraron los eventos reales.

Por otro lado, se simularon 38 eventos adicionales para cubrir ejemplos no contenidos en la base de datos real empleando modelos de redes de distribución en ATP, de la siguiente manera:

- Las formas de onda debidas al arranque de motores fueron generadas a partir de la red de distribución con 4 motores de inducción que se presenta en (Yalcinkaya *et al.*, 1998). Cada motor fue energizado estando el resto en operación y fuera de operación. El torque de carga fue modificado entre el 60% y 90% del valor nominal del motor. Ningún motor se arrancó en vacío, siempre con carga previamente conectada al alimentador.
- Las formas de onda debidas a la energización de bancos de capacitores fueron generadas utilizando 2 redes de distribución presentadas en (Coury *et al.*, 1998) y (Kersting *et al.*, 2000). La primera corresponde a una red localizada en USA, y la otra a una red de 34 nodos propuesta por IEEE. En ambas redes se realizaron energizaciones con (*back-to-back commutation*) y sin (*normal commutation*) la presencia de un banco pre-energizado. Los bancos de capacitores también fueron ubicados y energizados en distintos puntos de la red con la intención de obtener diferentes frecuencias de resonancia y factor de amortiguamiento como consecuencia de la distancia (impedancia de la línea) entre el banco y el equipo de medida.

Los eventos simulados se obtuvieron para las mismas condiciones que los reales, es decir, la misma cantidad de ciclos y muestras por ciclo. Igualmente, la fuente de alimentación fue modelada con un transformador con conexión estrella-delta y aterrizado de acuerdo al diagrama unifilar de la Figura 3. A raíz de la diferencia entre el nivel de tensión de la red eléctrica real y las redes simuladas, los atributos fueron calculados a partir de las señales trifásicas en por unidad. Ambos tipos de eventos tienen estados antes del evento de tensión (prefalta), durante el evento (falta) y posterior al evento (postfalta). Tales estados son necesarios para el cálculo de los atributos. Todas estas consideraciones permiten la coherencia entre las conclusiones obtenidas a partir de ambos tipos de datos.

En la Tabla 2 se muestra la cantidad de eventos utilizados para el análisis de acuerdo tanto con su causa, como con su localización relativa.

Tabla 2. Cantidad de eventos de tensión utilizados

1 ϕ	2 ϕ	3 ϕ	C	T	M	ME	Total
203	208	25	28	11	10	37	522

Los eventos balanceados corresponden a las faltas de red trifásicas (3 ϕ), energización de capacitores (C) y arranque de motores (M), mientras que los desbalanceados a faltas monofásicas (1 ϕ), bifásicas (2 ϕ) y energización de transformadores (T), ver Tabla 1. En total son 63 eventos balanceados y 422 desbalanceados. Debido a que los eventos multiestado (ME) o evolutivos (Figura 6) involucran varias fases durante la perturbación, éstos no fueron considerados dentro de ninguna de las categorías anteriores (balanceados o desbalanceados).

La localización relativa de la fuente de cada uno de los eventos fue suministrada por la compañía de distribución, en total 257 fueron generados aguas arriba y 220 aguas abajo con respecto al punto de medida.

3. Atributos

En esta sección se definen cada uno de los atributos utilizados para el análisis.

3.1 Atributos asociados con la localización relativa de la fuente de los eventos de tensión

Los atributos propuestos para la localización relativa de los eventos de tensión se presentan y definen brevemente en la Tabla 3.

Tabla 3. Atributos asociados con la localización relativa de eventos de tensión

Descriptor	Definición
R_{ex}	Parte real de la impedancia.
R_{ey}	Parte imaginaria de la impedancia.
Re	Resistencia equivalente.
Z_{ratio}	Relación de la impedancia durante la falta y la prefalta.
$\angle(Z_{falta})$	Angulo de la impedancia durante la falta.
I_{ratio}	Relación entre la magnitud de la corriente de falta y prefalta de secuencia positiva.
$\Delta\phi$	Diferencia del ángulo de fase de la corriente de falta y prefalta de secuencia positiva.
$I\cos(\phi)$	Componente real de la corriente.

Parte real (R_{ex}) e imaginaria (R_{ey}) de la impedancia y resistencia equivalente (Re)

El algoritmo propuesto en (Tayjasanant *et al.*, 2005) utiliza las componentes de tensión y corriente de secuencia positiva para calcular los valores de R_{ex} , R_{ey} y Re . Estos atributos se estiman tomando n ciclos de las ondas de tensión y corriente incluyendo ciclos de prefalta y falta (Barrera *et al.*, 2009; Tayjasanant *et al.*, 2005). La fuente del evento se encuentra aguas abajo si los signos de R_{ex} y R_{ey} son positivos o aguas arriba si ambos son negativos. Si los signos son diferentes el resultado del algoritmo es no conclusivo. Por otro lado si Re es mayor que cero, la fuente se encuentra aguas abajo, y aguas arriba en caso contrario.

Z_{ratio} , $\angle(Z_{falta})$ – Coeficiente entre la impedancia durante la falta/prefalta y ángulo de la impedancia durante la falta

El cálculo de estos atributos se basa en el principio de la magnitud y ángulo de la impedancia vista por el medidor de calidad de energía antes (prefalta) y después (postfalta) del evento. Las impedancias se calculan tomando las señales trifásicas de tensión y corriente. La fuente del evento es aguas abajo si $Z_{ratio} < 1$ y $\angle(Z_{falta}) > 1$ y aguas arriba si $Z_{ratio} > 1$ y $\angle(Z_{falta}) < 0$ (Barrera *et al.*, 2008, 2009).

I_{ratio} – Cociente entre la magnitud de la corriente de falta y prefalta de secuencia positiva

De acuerdo con los resultados obtenidos en (Pradhan *et al.*, 2007), en un sistema radial cuando se presenta un evento de tensión aguas abajo, la magnitud de la corriente de secuencia positiva sufre un ligero incremento, motivo por el cual I_{ratio} toma valores mayores que la unidad. Caso contrario sucede cuando la ubicación de la falta se encuentra aguas arriba del punto de medida, en donde I_{ratio} toma valores menores a la unidad.

$\Delta\phi$ – Cambio de fase en la corriente de secuencia positiva

Este atributo corresponde a la diferencia del ángulo de fase entre las componentes de secuencia positiva de la corriente de falta y prefalta. Si la diferencia es mayor que cero la fuente del evento se encuentra aguas abajo, en caso contrario aguas arriba del equipo de medida. Una explicación detallada de este atributo es dada en (Pradhan *et al.*, 2007; Barrera *et al.*, 2009).

$I\cos(\phi)$ – Componente real de la corriente

Este atributo se define como la integral de la componente real de la corriente de falta. Para eventos aguas arriba se cumple $I\cos(\phi) < 0$ y para eventos aguas abajo $I\cos(\phi) > 0$. El valor de $I\cos(\phi)$ depende de la dirección del flujo de corriente (Hamzah *et al.*, 2005).

3.2 Atributos asociados con la causa de los eventos de tensión

Algunos de los atributos utilizados han sido calculados a partir de la secuencia de valores eficaces de la componente fundamental de tensión o corriente respectivamente. La secuencia de valores eficaces fue obtenida aplicando la Transformada Rápida de Fourier (FFT) a través de una ventana deslizante de un ciclo de longitud.

Debido a la cantidad de atributos utilizados se decidió describirlos por grupos de acuerdo con las operaciones matemáticas necesarias para calcularlos, es decir, incrementos, pendientes, desviaciones y otros.

Pendientes de tensión y corriente durante la inserción de la falta

La pendiente en cada fase de tensión y corriente se obtiene durante el primer estado de transición del evento, como se muestra en la Figura 4. Estas pendientes son utilizadas para evaluar el grado de simetría de las tres fases de tensión y corriente durante la inserción de la falta. Finalmente fueron utilizados como atributos la pendiente máxima de tensión (PV_{max}) y de corriente (PI_{max}).

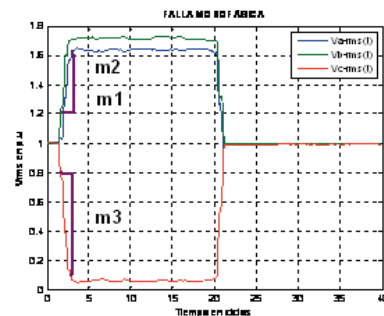


Figura 4. Estimación de pendientes de la señal de tensión eficaz durante el primer estado transitorio

Atributos que involucran la desviación estándar de magnitudes

La Tabla 4 muestra un conjunto de atributos que se obtienen a partir de la desviación estándar de pendientes de tensión y corriente al igual que la desviación de magnitudes de tensión y corriente eficaces durante el primer estado transitorio (Figura 4).

Los atributos σ_{PV} y σ_{PI} miden el grado de desbalance tomando valores altos para eventos desbalanceados y bajos para balanceados. Igualmente, σ_{maxV} y σ_{maxI} miden al grado de simetría existente en las fases de tensión y corriente durante la evolución de la perturbación, respectivamente.

Tabla 4. Atributos que involucran la desviación estándar de magnitudes

Atributo	Definición
σPV	Desviación estándar de las tres pendientes de tensión.
σPI	Desviación estándar de las tres pendientes de corriente.
$\sigma maxV$	Máxima desviación estándar de la señal de tensión a lo largo del evento.
$\sigma maxI$	Máxima desviación estándar de la señal de corriente a lo largo del evento.

Atributos que involucran el cociente de magnitudes

Los atributos que involucran cocientes de magnitudes son útiles para evaluar los cambios existentes entre los estados de prefalta y postfalta. La Tabla 5 muestra un conjunto de atributos que utilizan el cociente de magnitudes como: la potencia aparente, tensión, corriente y factor de potencia. En general, este grupo de atributos se incluyeron con la intención de caracterizar aquellos eventos que involucran cambios de carga en la red eléctrica como arranque de motores de inducción y energización de capacitores.

Tabla 5. Atributos que involucran el cociente de magnitudes

Atributo	Definición
$Sratio$	Relación entre la potencia aparente de postfalta y prefalta.
$Vratio$	Relación de tensión de postfalta y prefalta.
$Iratio$	Relación de corriente de postfalta y prefalta.
$FPratio_{Cero}$	Relación entre el factor de potencia de postfalta y prefalta. Ángulos estimados utilizando los cruces por cero de la señal de tensión y corriente.
$FPratio_{FFT}$	Relación entre el factor de potencia de postfalta y prefalta. Ángulos estimados utilizando FFT.
$FPratio_{P&S}$	Relación entre el factor de potencia de postfalta y prefalta. Ángulos estimados utilizando la potencia activa (P) y aparente (S).

El cociente de la potencia aparente $Sratio$ permite identificar los eventos de tensión provocados por el arranque de grandes motores de inducción. Los eventos de tensión causados por arranque de motores de inducción producen un incremento considerable de la corriente (Figura 5) durante y después del arranque del motor de inducción y por tanto $Iratio$ y $Sratio$ toman valores mayores a la unidad.

El cociente entre los factores de potencia de postfalta y prefalta es superior a la unidad ante la energización de bancos de capacitores (Blanco y Jagua et al., 2009). Este atributo se obtuvo de tres maneras: (1) Calculando los ángulos de desfase entre la tensión y la corriente a partir de los cruces por cero de las señales instantáneas ($FPratio_{Cero}$), (2) utilizando FFT ($FPratio_{FFT}$) y (3) a partir del cociente entre la potencia activa y la potencia aparente ($FPratio_{P&S}$) (Velandia y Hernández et al., 2010). El atributo fue calculado de varias formas con la intención de evaluar la capacidad de los tres distintos enfoques.

$Vratio$ toma valores mayores que la unidad cuando ocurre la energización de un banco de capacitores, ya que éstos al ser instalados en las redes de distribución mejoran la regulación, incrementando el nivel de tensión.

Atributos que involucran la cantidad de segmentos de transición

Los segmentos de transición hacen referencia a cada uno de los instantes donde se inicia y finaliza un estado de transición durante el evento. Por ejemplo, en la Figura 1a se presenta un hueco de tensión causado por una falta monofásica, el cual

contiene cuatro segmentos (rectas verticales) delimitando dos estados de transición.

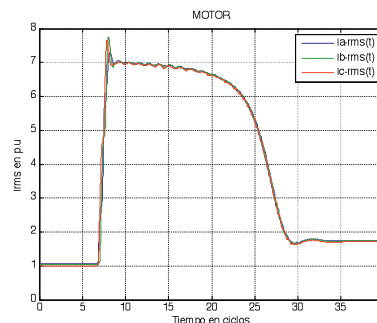


Figura 5. Magnitud de la corriente durante el arranque de un motor de inducción

Se puede evidenciar que los eventos monoestado (Figura 1a, b, c) tienen cuatro segmentos de transición mientras que los multiestado (Figura 6) tienen más de cuatro. Por tal razón, la cantidad de segmentos de transición en este artículo se utiliza para discriminar entre eventos multiestados y monoestados. La Tabla 6 muestra los atributos empleados para discriminar entre ambos tipos de eventos.

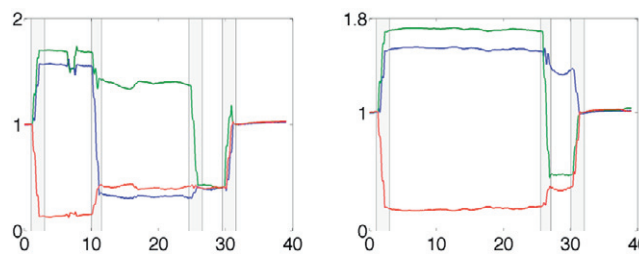


Figura 6. Tensión eficaz de evento multiestado o evolutivo.

Tabla 6. Atributos que involucran la cantidad de segmentos de transición de la señal de tensión

Atributo	Definición
$Segm(\delta=0,1; m=128)$	Número de segmentos obtenidos para una segmentación de señal de umbral $\delta=0,1\%$ y rata de submuestreo $m=128$ muestras.
$Segm(\delta=8; m=64)$	Número de segmentos obtenidos para una segmentación de señal de umbral $\delta=8\%$ y rata de submuestreo $m=64$ muestras.
σE	Desviación de las magnitudes de tensión en cada uno de los estados.

Segm(δ=8, m=64) y Segm(δ=0,1, m=128) – Numero de segmentos de transición

La identificación de los segmentos de transición se realiza aplicando la segmentación basada en la primera derivada de la señal de tensión eficaz descrito en (Bollen et al., 2007; Barrera et al., 2010b). Con la intención de distinguir entre eventos multiestado y monoestado se utilizó el atributo $Seg(\delta=8; m=64)$ con una rata de submuestreo m de 64 muestras y un valor umbral δ de 8%. δ corresponde a un valor umbral para la estimación del inicio y final de los estados de transición. Tales parámetros son adecuados para segmentar eventos de tensión (Barrera et al., 2010b). Este atributo tiende a tomar valores superiores a cuatro para eventos multiestado e igual a cuatro para monoestado. El

atributo $Seg(\delta=0,1; m=128)$ es otra alternativa planteada para segmentar eventos de tensión, permitiendo así comparar y evaluar dos atributos asociados con la discriminación de eventos monoestados y multiestados. $Seg(\delta=0,1; m=128)$ y $Seg(\delta=8; m=64)$ son útiles para detectar cambios pequeños y bruscos de magnitud en el valor eficaz de la señal (Barrera et al., 2010b), respectivamente.

σE – Desviación de los estados estables

Este atributo se define como el incremento máximo de tensión en p.u. que se presenta entre estados estables consecutivos en la secuencia de valores eficaces de la tensión. Los instantes de inicio y final de cada uno de los estados estables se estiman mediante la primera derivada de la secuencia de tensión eficaz. Una vez identificados los estados estables, se calcula el valor medio de la tensión durante cada uno de ellos. Luego, se estiman los cambios porcentuales de tensión entre cada uno de los estados estables consecutivos, y en cada una de las tres fases (Blanco y Jagua et al., 2009). σE corresponde al máximo cambio obtenido de las tres fases. Con este atributo es posible estimar la magnitud de los cambios de tensión presentes en eventos multiestados, permitiendo discriminarlos de los monoestado.

Atributos que involucran incrementos de magnitudes

Los atributos pertenecientes a este grupo emplean incrementos o diferencias de magnitudes eléctricas (Tabla 7). Por ejemplo, mediante la diferencia entre las potencias reactivas de prefalta y postfalta es posible identificar eventos causados por energización de bancos de capacitores, ya que al ser insertados en la red eléctrica generan cambios significativos en la potencia reactiva demandada por la red.

Tabla 7. Atributos que involucran incrementos de magnitudes

Atributo	Definición
ΔQ	Diferencia de potencias reactivas de prefalta y postfalta.
θVI_{FFT}	Diferencia de los ángulos de la impedancia de postfalta y prefalta calculada por la FFT.
θVI_{cero}	Diferencia de los ángulos de la impedancia de postfalta y prefalta calculada por el cruce por cero de la tensión y la corriente.

El valor relativo de la potencia reactiva después de la energización de un banco de capacitores es menor que el valor relativo de la potencia reactiva previa a la inserción del mismo, este hecho es debido a que la potencia reactiva del banco de capacitores reduce el efecto de la potencia reactiva inductiva en el sistema. De manera similar, θVI_{cero} y θVI_{FFT} toman valores negativos ante la energización de un banco de capacitores a raíz de que la señal de corriente tiende a adelantar la señal de tensión.

Otros atributos adicionales

Los atributos descritos en esta sección permiten caracterizar los eventos de tensión respecto a su grado de desbalance, forma de la señal de tensión eficaz, contenido de componentes armónicas y otros, ver Tabla 8. A continuación se detallan algunos de estos atributos.

δL – Relación angular de las longitudes

Es posible evaluar el grado de desbalance de un evento mediante la estimación del número de fases que intervienen en el

mismo. La aplicabilidad de este atributo radica en la identificación de eventos de tensión originados por faltas de red monofásicas y bifásicas.

Tabla 8. Atributos adicionales

Atributo	Definición
RCV	Tensión compleja remanente.
PNF	Factor positivo – negativo.
GV_0	Ganancia de tensión de secuencia cero.
LV	Pérdida de tensión basada en RCV.
LV_3	Pérdida de tensión basada en las tres tensiones.
$MaxV_0$	Amplitud máxima de tensión de secuencia cero.
I_2I	Segunda componente armónica de la corriente.
Cf	Coefficiente de forma.
δL	Relación angular de las longitudes.
ΔLV	Diferencia máxima de longitudes de tensión.
ΔLI	Diferencia máxima de longitudes de corriente.
$\sigma \delta L$	Desviación de las longitudes eléctricas.
$\Delta minV$	Diferencia mínima de valores eficaces de tensión.
$\Delta minI$	Diferencia mínima valores eficaces de corriente.

Tomando como punto de referencia la primera muestra de tensión eficaz, se trazan rectas a cada uno de los puntos consecutivos a éste y se determina el valor de la proyección vertical de cada una de las rectas trazadas (Figura 7). El proceso se realiza para cada una de la fases. Con estas proyecciones se realiza una sumatoria, una para cada fase, y a cada resultado se le denomina “longitud de la fase (ℓ_A, ℓ_B, ℓ_C)” las cuales se denominan “longitudes eléctricas” (Blanco y Jagua et al., 2009; Barrera et al., 2010b).

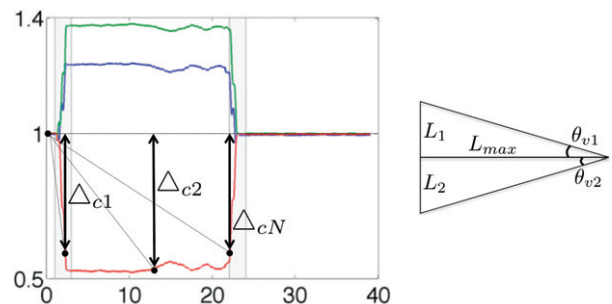


Figura 7. (a) Método de cálculo del atributo δL , (b) triángulo de longitudes eléctricas.

- ℓ_A, ℓ_B y ℓ_C son proyectadas en el triángulo de longitudes eléctricas de la Figura 7b. Donde L_{max} corresponde al máximo valor entre las tres longitudes [$L_{max} = \max(\ell_A, \ell_B, \ell_C)$], y L_1 y L_2 las dos longitudes restantes. A partir del triángulo se observa que:
 - Para faltas trifásicas los ángulos θ_1 y θ_2 serán aproximadamente 45° , $\theta_1 \sim \theta_2 \sim 45^\circ$.
 - Ante faltas bifásicas uno de los dos ángulos será aproximadamente 45° , $\theta_1 = 45^\circ$ o $\theta_2 = 45^\circ$.
 - Para faltas monofásicas $L_{max} \gg L_1$ y $L_{max} \gg L_2$ y $\theta_1 \ll 45^\circ$ y $\theta_2 \ll 45^\circ$.

Finalmente δL es calculado como $\delta L = \max(\theta_1/45^\circ, \theta_2/45^\circ)$. Ambos ángulos son normalizados en 45° porque es el valor máximo que éstos pueden tomar. Por tanto, δL toma para faltas monofásicas y bifásicas valores mucho menores y cercanos a la unidad respectivamente.

El atributo $\sigma \delta L$ se obtiene como la desviación estándar de las tres longitudes eléctricas [$\sigma \delta L = \sigma(\ell_A, \ell_B, \ell_C)$]. $\sigma \delta L$ toma valores

bajos para eventos balanceados, y valores mucho mayores que cero en presencia de eventos desbalanceados. ΔLV y ΔLI corresponden a la diferencia máxima entre las longitudes eléctricas de tensión y corriente respectivamente.

CF – Coeficiente de forma

Observando la secuencia de tensiones eficaces de los eventos de tensión, se evidencia que algunos presentan tendencia rectangular (Faltas de red, Figura 1 a, b y c) y otros triangular (Energización de transformadores, Figura 8 y 1e). CF toma valores cercanos a cero en presencia de faltas de red (Figura 1 a, b y c) (tendencia no triangular) y cercanos a la unidad ante eventos causados por la energización de transformadores (tendencia triangular) (Blanco y Jagua *et al.*, 2009; Barrera *et al.*, 2010b).

CF utiliza la muestra de inicio y final de la perturbación así como también la muestra donde la tensión eficaz es mínima para valorar la tendencia triangular de la secuencia de tensión (Figura 8).

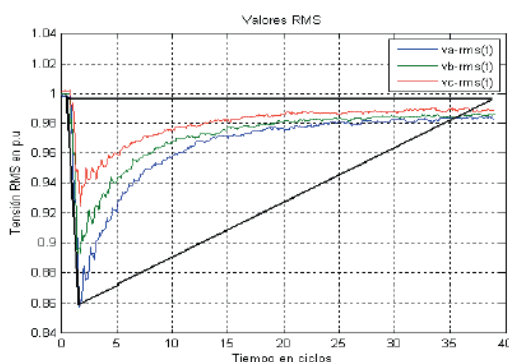


Figura 8. Secuencia de tensiones eficaces ante la energización de un transformador.

RCV y PNF – Tensión característica y Factor Positivo Negativo

Los valores mínimos de ambos atributos contienen información sobre la magnitud y grado de desbalance de los eventos de tensión. En este artículo, para calcularlos se utilizó el algoritmo de las seis fases (Bollen, 2003). El atributo RCV corresponde a una señal de tensión que representa a las tres tensiones de fase.

Diferencia mínima de valores eficaces de tensión y corriente (Δ_{minV} , Δ_{minI})

Δ_{minV} es la diferencia mínima existente entre la fase con menor tensión eficaz durante el evento, y las dos fases restantes, esto es, Δ_{minV} corresponde al valor mínimo entre ΔV_1 y ΔV_2 en la Figura 9 ($\min[\Delta V_1, \Delta V_2]$). De la misma manera, Δ_{minI} corresponde a $\min[\Delta I_1, \Delta I_2]$ (Velandia y Hernández *et al.*, 2010).

Para faltas monofásicas Δ_{minV} y Δ_{minI} toman valores altos, para bifásicas valores medios y valores bajos en presencia de faltas trifásicas. Este hecho le brinda la posibilidad a este atributo de discriminar los tres tipos de faltas existentes.

4. Análisis Estadístico de Atributos

La herramienta estadística MANOVA (*Multivariate Analysis of Variance*) realiza el análisis de la varianza para más de una variable dependiente. En este caso, las variables dependientes corresponden a los atributos y las independientes a la localización relativa y la causa de los eventos. En este trabajo MANOVA fue utilizada para determinar la relevancia de cada atributo respecto de la localización relativa y la causa de los eventos de tensión. Como resultado, es posible conocer el grado de influencia o efecto de la causa/localización de los eventos de tensión sobre cada uno de los atributos planteados. En otras palabras, MANOVA permite responder esta pregunta: **¿Cuál es la importancia/relevancia de cada atributo con respecto a la localización relativa y a la causa del evento de tensión?** Por tanto, es posible conocer el grado de incidencia (Calidad) de la localización/causa del evento de tensión sobre cada uno de los atributos utilizados en el análisis.

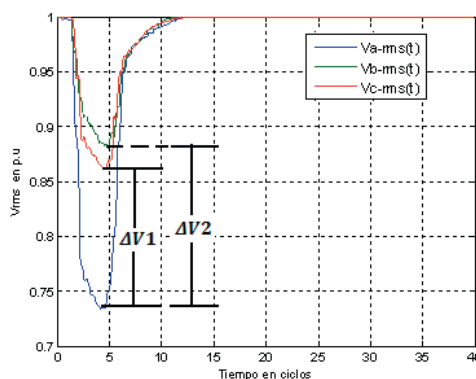


Figura 9. Diferencia mínima de valores eficaces de tensión (Δ_{minV}).

Para la identificación de los atributos relevantes se establecieron 4 escenarios asociados con la causa, y 1 escenario para atributos asociados con la localización relativa. Los resultados se presentan de la Tabla 9 a la Tabla 13.

4.1 Análisis estadístico: Causa de los eventos de tensión

Los atributos que no están relacionados directamente con la causa de los eventos fueron excluidos para evitar la introducción de ruido en el análisis estadístico de los mismos. Por ejemplo, en el escenario de eventos balanceados no tiene sentido incluir atributos que valoren el grado de desbalance de las señales de tensión y corriente, porque los eventos balanceados por su naturaleza poseen grado de desbalance casi nulo, por el contrario, los eventos balanceados deben ser discriminados a partir de atributos que relacionen los estados de prefalta y postfalta.

En la Tabla 9 es posible apreciar la relevancia o calidad de cada atributo con respecto al grado de desbalance del evento de tensión. En este escenario, la variable independiente (grado de desbalance) toma valores “Evento desbalanceado” o “Evento balanceado”. Una calidad próxima al 100% indica que la mayor parte de la variabilidad del atributo esta asociada con la causa del evento de tensión, y por tanto el atributo es altamente relevante respecto de la causa del evento. Por ejemplo, el atributo $\sigma_{max}V$ (96,6%) obtuvo la mayor relevancia seguido por $\sigma\delta L$ (95,7%) y Δ_{minV} (92,4%).

También es posible apreciar que los atributos basados en la corriente no tienen una relevancia significativa respecto al grado de desbalance del evento de tensión, excepto para $Inratio$ (71,6%)

y $\Delta minI$ (78,8%). Por lo tanto, es posible afirmar que la señal de tensión contiene mayor información sobre el grado de desbalance de los eventos que la señal de corriente.

Los atributos RCV , PNF , δL , ΔLV y $MaxV_0$ (Tabla 10) obtuvieron los índices de calidad más altos indicando que son capaces de discriminar las faltas monofásicas, bifásicas y las causadas por energización de transformadores. En este escenario la variable independiente (causada del evento) toma valores “Falta monofásica”, “Falta bifásica” y “Energización de transformador”.

Tabla 9. Resultados MANOVA para el grado de desbalance de eventos de tensión

Atributo	Definición	Calidad
$\Delta minV$	Diferencia mínima de valores eficaces de tensión.	92,40%
$\Delta minI$	Diferencia mínima de valores eficaces de corriente.	78,80%
σPV	Desviación estándar de las 3 pendientes de tensión.	91,40%
σPI	Desviación estándar de las 3 pendientes de corriente.	4,00%
RCV	Tensión compleja remanente.	90,40%
PNF	Factor positivo negativo.	90,00%
$\sigma maxV$	Máxima desviación de las 3 señales de tensión.	96,6%
$\sigma maxI$	Máxima desviación de las 3 señales de corriente.	5,00%
$\sigma \delta L$	Desviación de las longitudes eléctricas.	95,70%
$Vnratio$	Máxima tensión de neutro.	88,00%
$Inratio$	Máxima corriente de neutro.	71,60%
GV_0	Ganancia de tensión de secuencia cero.	20,00%
LV	Pérdida de tensión basada en RCV.	11,00%
LV_3	Pérdida de tensión basada en las tres tensiones.	2,00%
$MaxV_0$	Máxima de tensión de secuencia cero.	30,00%

Nota: Evento desbalanceado: Faltas monofásicas, bifásicas y energización de transformador. Evento balanceado: Arranque de motores, bancos de capacitores y faltas trifásicas.

Tabla 10. Resultados MANOVA para eventos desbalanceados

Atributo	Definición	Calidad
$\Delta minV$	Diferencia mínima de valores eficaces de tensión.	60,00%
$\Delta minI$	Diferencia mínima de valores eficaces de corriente.	84,60%
σPV	Desviación estándar de las 3 pendientes de tensión.	63,00%
σPI	Desviación estándar de las 3 pendientes de corriente.	14,00%
RCV	Tensión compleja remanente.	98,00%
PNF	Factor positivo negativo.	98,00%
$\sigma maxV$	Máxima desviación de las 3 señales de tensión eficaz	71,00%
$\sigma maxI$	Máxima desviación de las 3 señales corriente eficaz.	14,00%
δL	Relación angular de las longitudes.	89,00%
ΔLV	Diferencia máxima de longitudes de tensión.	85,00%
ΔLI	Diferencia máxima longitudes de corriente.	82,60%
CF	Coefficiente de forma.	85,40%
$Vnratio$	Máxima tensión de neutro en p.u.	72,00%
$Inratio$	Máxima corriente de neutro en p.u.	68,70%
GV_0	Ganancia de tensión de secuencia cero.	58,00%
LV	Pérdida de tensión basada en RCV.	8,00%
LV_3	Pérdida de tensión basada en las tres tensiones.	15,00%
$MaxV_0$	Máxima de tensión de secuencia cero.	88,50%

Nota: Se incluyeron sólo los eventos etiquetados como: Faltas monofásicas, bifásicas y energización de transformadores.

En Tabla 11 se muestra que los atributos basados en la relación de potencias, corrientes y factores de potencia poseen mayor relevancia con respecto a los eventos balanceados. El atributo $FPratio_{Cero}$ es el más relevante (85%) de los 3 atributos relacionados con el cociente entre los factores de potencia de postalta y prefalta.

Como cuarto escenario se establecieron los atributos relacionados con eventos multiestado (evolutivo, Figura 6) y monoestado (faltas de red, Figura 1 a, b y c). Según MANOVA, se observa en la Tabla 12 que para distinguir eventos monoestado

de multiestados, es mejor utilizar el atributo basado en segmentación $Segm(\delta=8, m=64)$ (90,7%) que el atributo σE (71,7%).

Tabla 11. Resultados MANOVA para eventos balanceados

Atributo	Definición	Calidad
$Sratio$	Relación entre la potencia aparente de postalta y prefalta.	98,00%
$Vratio$	Relación de postalta y prefalta.	76,30%
$Iratio$	Relación de corriente de postalta y prefalta.	98,00%
$FPratio_{FFT}$	Relación entre el factor de potencia de prefalta y postalta. Ángulos estimados utilizando FFT.	42,2%
$FPratio_{P\&S}$	Relación entre el factor de potencia de prefalta y postalta. Ángulos estimados utilizando la potencia activa (P) y aparente (S).	4,00%
$FPratio_{Cero}$	Relación entre el factor de potencia de prefalta y postalta. Ángulos estimados utilizando los cruces por cero de la señal de tensión y corriente.	85,00%
θVI_{FFT}	Diferencia de los ángulos de la impedancia de postalta y prefalta. Ángulos estimados utilizando FFT.	5,00%
θVI_{Cero}	Diferencia de los ángulos de la impedancia de postalta y prefalta. Ángulos estimados utilizando los cruces por cero de la señal de tensión y corriente.	21,00%
ΔQ	Diferencia de potencias reactivas de prefalta y postalta.	70,00%

Nota: Se incluyeron sólo los eventos etiquetados como: Faltas trifásicas, energización de banco de capacitores y arranque de motores.

Tabla 12. Resultados MANOVA para eventos multiestado y monoestado

Atributo	Definición	Calidad
$Segm(\delta=0,1, m=128)$	Número de segmentos obtenidos para una segmentación de señal de umbral $\delta=0.1\%$ y rata de submuestreo $m=128$ muestras.	77,60%
$Segm(\delta=8, m=64)$	Número de segmentos obtenidos para una segmentación de señal de umbral $\delta=8\%$ y rata de submuestreo $m=64$ muestras.	90,70%
σE	Desviación de las magnitudes de tensión en cada uno de los estados.	71,70%

Nota. Evento monoestado: Faltas monofásicas, bifásicas y trifásicas. Evento multiestado: Faltas de red multiestado.

Tabla 13. Resultados MANOVA para la localización relativa de los eventos de tensión

Atributo	Definición	Calidad
Rex	Parte real de la impedancia.	72,30%
Rey	Parte imaginaria de la impedancia.	76,30%
Re	Resistencia equivalente.	46,20%
$Zratio$	Relación de la impedancia durante la falta y la prefalta.	87,00%
$\acute{a}ngulo(Zfalta)$	Ángulo de la impedancia durante la falta.	60,70%
$Irratio$	Relación entre la magnitud de la corriente de falta y prefalta de secuencia positiva.	98,00%
$\Delta \phi$	Diferencia del ángulo de fase de la corriente de falta y prefalta de secuencia positiva.	70,90%
$I\cos(\phi)$	Componente real de la corriente.	79,00%

Nota: Los eventos incluidos en el análisis fueron validados y etiquetados previamente como “Aguas arriba” y “Aguas abajo” por parte de la compañía de distribución.

De igual manera se nota que el atributo $Segm(\delta=0,1, m=128)$ tiene menor relevancia que $Segm(\delta=8, m=64)$, ya que los parámetros de segmentación del primero son adecuados para discriminar eventos de un solo estado transitorio (energización de transformadores y energización de capacitores) mientras que el segundo para eventos de varios estados transitorios (Barrera et al., 2010b).

4.2 Análisis estadístico relacionado con la localización relativa de la fuente de eventos de tensión

En la Tabla 13 se listan los atributos asociados con la localización relativa de los eventos y su respectiva calidad al igual como se mostraron para los 4 escenarios del análisis de causas de eventos de tensión.

Los resultados obtenidos en la Tabla 13 resaltan los atributos *Iratio* y *Zratio*. Este último es un atributo complementario con *Ángulo(Zfalta)*, ambos incrementan la precisión en la estimación del origen del evento de tensión. En su combinación se basa la robustez del algoritmo *Distance Relay* (Barrera et al., 2009).

5. Extracción de Reglas de Decisión

A partir de los resultados MANOVA se seleccionaron los atributos con mayor relevancia (Calidad mayor a 50%) ver Tabla 14. Posteriormente estos atributos fueron utilizados como entrada de seis algoritmos de extracción de reglas con la intención de extraer reglas de decisión que permitan diagnosticar de manera automática los eventos de tensión. La selección de estos atributos permitió reducir la cantidad de reglas extraídas por los algoritmos, ya que en ellos cuanto mayor es la cantidad de atributos, mayor la cantidad de reglas extraídas. La anterior es la ventaja de identificar los atributos relevantes mediante MANOVA.

Tabla 14. Atributos seleccionados para cada escenario

Escenario 1: Eventos balanceados/desbalanceados	$\Delta_{min}V, \Delta_{min}I, \sigma_{PV}, RCV, PNF, \sigma_{max}V, \sigma_{\delta}L, Vnratio, Inratio$
Escenario 2: Eventos Desbalanceados	$\Delta_{min}V, \Delta_{min}I, \sigma_{PV}, RCV, PNF, \sigma_{max}V, \delta L, \Delta LV, \Delta LI, CF, Vnratio, Inratio, GV_0, MaxV_0$
Escenario 3: Eventos balanceados	$Sratio, Vratio, Iratio, FPratio_{Cerro}, \Delta Q$
Escenario 4: Eventos multiestados	$\sigma E, Segm(\delta=0,1 m=128), Segm(\delta=8, m=64)$
Escenario 5: Localización relativa de la fuente de los eventos	$Rex, Rey, Zratio, \text{ángulo}(Zfalta), Iratio, \Delta\phi, I\cos(\phi)$

5.1 Algoritmos utilizados para la extracción de reglas de decisión

Los algoritmos utilizados fueron *Conjunctive Rule*, *JRip*, (*RIPPER*), *OneR*, *PART*, *Ridor* (Hernández et al., 2004) y *CN2* (Clark et al., 1981, 1991). Los algoritmos se aplicaron empleando las plataformas de minería de datos *Weka* y *Orange-Canvas*. El tratamiento de los eventos y cálculo de atributos fue realizado utilizando Matlab®. A continuación se da una breve descripción de los algoritmos de extracción seleccionados:

- *Conjunctive Rule*: Implementa una sola regla de aprendizaje conjuntiva a partir de la comparación de un conjunto de datos validados.
- *JRip*: Basado en el algoritmo *RIPPER* (Poda incremental Repetida para la Reducción del Error). Utiliza varias comparaciones al mismo tiempo, construye un conjunto de reglas por separado y luego establece comparaciones entre ellas.
- *OneR*: Construye reglas usando el clasificador (algoritmo) 1R. Utiliza el error mínimo para la predicción de nuevas reglas discretizando atributos numéricos.
- *PART*: Genera una lista de reglas de decisión en orden de jerarquía. En esencia construye una regla, elimina las instancias que cubre y continúa creando reglas recursivamente para las instancias finales hasta que no queda ninguna instancia.
- *Ridor*: Basado en el algoritmo *Ripple Down Rule*. Genera una regla por defecto (predeterminada) y luego toma un conjunto de

reglas que predicen clases para la regla predeterminada con el mínimo de error. Entonces genera el mejor conjunto de reglas hasta lograr disminuir el error.

- *CN2*: Elimina ruido en los datos y evita el sobre ajuste en los mismos presentado en otros algoritmos. Selecciona la regla que cubre la mayor cantidad de ejemplos de una clase y pocas de las demás clases para compararlas (Clark, et al., 1981, 1991).

5.2 Reglas extraídas para eventos balanceados y desbalanceados

En la Tabla 15 se muestran las reglas extraídas y la proporción de eventos cubiertos por cada una de ellas en el primer escenario (Tabla 9) de la causa de los eventos de tensión, es decir, evento “Balanceado” o “Desbalanceado”.

Tabla 15. Reglas extraídas para eventos desbalanceados y balanceados

Clase	Algt.	Regla extraída	Cobertura
Evento desbalanceado	<i>CR</i>	$\sigma_{\delta}L > 0.131022 \rightarrow$ Desbalanceado	100%
	<i>OneR</i>	$\sigma_{\delta}L \geq 0.13102 \rightarrow$ Desbalanceado	100%
	<i>PART</i>	$\sigma_{\delta}L > 0.116368 \rightarrow$ Desbalanceado	100%
	<i>CN2</i>	$\sigma_{\delta}L > 0.116368 \rightarrow$ Desbalanceado	100%
Evento balanceado	<i>JRip</i>	$\sigma_{\delta}L \leq 0.116368 \rightarrow$ Balanceado	98%
	<i>Ridor</i>	$\sigma_{\delta}L \leq 0.11769 \rightarrow$ Balanceado	100%

Es importante notar en la Tabla 15 que todos los algoritmos han extraído la misma regla basada en el atributo “Desviación de las longitudes eléctricas ($\sigma_{\delta}L$)”, el cual obtuvo una relevancia de 95,7% en el análisis MANOVA (Tabla 9). Aunque las seis reglas poseen distintos puntos de corte para el atributo $\sigma_{\delta}L$, los puntos de corte son aproximadamente el mismo (alrededor de 0,12 a 0,13). Las reglas indican que el descriptor $\sigma_{\delta}L$ es capaz de discriminar entre eventos balanceados y desbalanceados con un adecuado grado de precisión.

5.3 Reglas extraídas para eventos desbalanceados

En la Tabla 16 se muestran las reglas extraídas para los eventos desbalanceados. Los resultados sugieren que los eventos monofásicos los discrimina adecuadamente el atributo *RCV*, ya que este atributo fue incluido en la mayoría de las reglas arrojadas por los seis algoritmos para los eventos monofásicos. Adicionalmente, y de manera similar al caso anterior, el punto de corte es aproximadamente el mismo (alrededor de 0,8), excepto para la regla extraída por *JRip* la cual no ha sido tomada en cuenta por no coincidir con las demás y por su baja cobertura (67%). El algoritmo *PART* extrajo un par de reglas con buena cobertura y basadas en el atributo *PNF*, éstas se descartaron por su baja repetitividad. Lo mismo ocurrió con las obtenidas por *CN2* (Clark, et al., 1981, 1991) debido a la baja repetitividad y cobertura.

Tabla 16. Reglas extraídas para eventos desbalanceados

Clase	Algt.	Regla extraída	Cobertura
Evento monofásico (1-fase)	<i>CR</i>	Si $RCV > 0,8646 \rightarrow$ 1-fase	100%
	<i>JRip</i>	Si $RCV \geq 0,9215 \rightarrow$ 1-fase	67%
	<i>OneR</i>	Si $RCV \geq 0,8646 \rightarrow$ 1-fase	100%
	<i>PART</i>	Si $PNF > 0,8596 \rightarrow$ 1-fase	100%
	<i>CN2</i>	$\delta L \leq 0,2524 \ \& \ V_0 > 0,0277 \rightarrow$ 1-fase	64%
Evento bifásico (2-fase)	<i>PART</i>	Si $0,222 < PNF \leq 0,734 \rightarrow$ 2-fase	100%
	<i>Ridor</i>	Si $RCV \leq 0,864592 \rightarrow$ 2-fase	100%
	<i>CN2</i>	$\delta L \leq 0,847 \ \& \ \Delta LV \leq 0,3419 \rightarrow$ 2-fase	78%
Energización Transformador (ET)	<i>PART</i>	$CF > 0,613312 \rightarrow$ ET	91%
	<i>Ridor</i>	$CF > 0,605565 \rightarrow$ ET	91%
	<i>CN2</i>	$CF > 0,61331242 \rightarrow$ ET	91%

Por otro lado, las reglas extraídas por los algoritmos y que describen los eventos causados por la energización de transformadores sugieren que este tipo de eventos toman valores para el atributo CF superiores a 0,6 aproximadamente.

Las reglas extraídas confirman la relevancia de los atributos RCV y CF, ya que éstos obtuvieron altos niveles de relevancia en el análisis MANOVA (Tabla 10); 98,0% y 85,4% respectivamente.

5.4 Reglas extraídas para eventos balanceados

Las reglas extraídas (Tabla 17) para eventos debidos a la energización de bancos de capacitores muestran que la relación de las tensiones (*Vratio*) y corrientes (*Iratio*) después y antes del evento son atributos que discriminan adecuadamente este tipo de eventos. La regla obtenida por el algoritmo *PART* cubre el 93% de los eventos causados por bancos de capacitores. Ésta indica que luego de la energización de un banco de capacitores la tensión se incrementa y la corriente se reduce. Lo anterior como consecuencia de la inyección de reactivos a la red por parte del banco capacitivo, efecto también conocido como mejoramiento de la regulación de tensión. Esta regla confirma la hipótesis inicial discutida anteriormente cuando se definieron ambos atributos.

Tabla 17. Reglas extraídas para eventos balanceados

Clase	Algt.	Regla extraída	Cobertura
Energización de bancos de capacitores (C)	CR	$FP_{ratio} > 1,014881 \rightarrow C$	64%
	JRip	$FP_{ratio} \geq 1,01959 \rightarrow C$	89%
	PART	$V_{ratio} > 0,9995 \ \& \ I_{ratio} \leq 0,9965 \rightarrow C$	93%
	Ridor	$FP_{ratio} > 1,0246 \ \& \ I_{ratio} \leq 0,9967 \rightarrow C$	89%
	CN2	$FP_{ratio} > 1,02065 \ \& \ \Delta Q \leq 0,9967 \rightarrow C$	82%
Arranque de motor de inducción (M)	JRip	$I_{ratio} \geq 1,575331 \rightarrow M$	100%
	PART	$V_{ratio} \leq 0,99954 \rightarrow M$	100%
	Ridor	$I_{ratio} \geq 0,99774 \rightarrow M$	100%
	CN2	$I_{ratio} > 1,159035 \rightarrow M$	100%
Falta trifásica (3-fase)	JRip	$FP_{ratio} > 1,000112 \rightarrow 3\text{-fase}$	80%
	PART	$FP_{ratio} = 1,000211 \rightarrow 3\text{-fase}$	76%
	Ridor	$I_{ratio} \leq 1,36718 \rightarrow 3\text{-fase}$	72%
	CN2	$S_{ratio} = 0,968 \rightarrow 3\text{-fase}$	75%

El atributo $FP_{ratio_{Cero}}$ de ahora en adelante FP_{ratio} , aunque aparece en varias reglas ninguna de ellas cubre más eventos que la regla *PART*. Esto no implica que FP_{ratio} sea un mal atributo, implica simplemente que las reglas que contienen a FP_{ratio} en su condición cubren menos eventos que la regla *PART*.

Los eventos causados por el arranque de motores de inducción son descritos por los atributos *Vratio* e *Iratio*. Las reglas confirman que después del arranque de un motor de inducción la corriente de operación sufre un incremento y que la tensión se reduce.

Cuando ocurre una falta de red trifásica es posible que parte de la carga (tramos de red, secciones) sea o no excluida. Este hecho afecta a los atributos FP_{ratio} , *Vratio*, *Iratio* y *Sratio*. Por ejemplo, FP_{ratio} tomará valores mayores a la unidad para faltas de red trifásicas en las cuales se haya aislado parte de la red a causa de fusibles fundidos u operación de protecciones, tal fenómeno ocasiona que una falta trifásica sea confundida con un evento causado por la energización de capacitores, ya que FP_{ratio} toma valores superiores a la unidad para estos últimos eventos. FP_{ratio} tomará valores muy cercanos a la unidad en caso de no existir exclusión de carga, esto último se ve reflejado en las reglas obtenidas por el algoritmo JRip, *PART* y CN2 (Clark, et al., 1981, 1991) en el apartado correspondiente a faltas trifásicas de la

Tabla 17, nótese que los puntos de corte de ambos atributos están muy cercanos a la unidad.

Los atributos *Iratio*, FP_{ratio} , y *Vratio* obtuvieron 98%, 85% y 76,34% de relevancia en el análisis MANOVA (Tabla 11).

5.5 Reglas extraídas para eventos monoestado y multiestado

En la Tabla 18 se presentan las reglas extraídas para los eventos clasificados como monoestados y multiestados. Las reglas obtenidas sugieren que el atributo más relevante para la identificación de la cantidad de estados es σE con punto de corte en 0,24 aproximadamente.

Tabla 18. Reglas extraídas para eventos monoestados y multiestados

Clase	Algt.	Regla extraída	Cobertura
Evento monoestado (1-E)	CR	$\sigma E \leq 0,239749 \rightarrow 1\text{-E}$	98%
	OneR	$\sigma E < 0,57209 \rightarrow 1\text{-E}$	98%
	PART	$\sigma E \leq 0,238845 \rightarrow 1\text{-E}$	98%
	CN2	$\sigma E \leq 0,2388 \ \& \ Segm(\delta=8, m=64) < 5 \rightarrow 1\text{-E}$	65%
Evento multiestado (∞ -E)	JRip	$\sigma E \geq 0,240654 \rightarrow \infty\text{-E}$	81%
	PART	$\sigma E > 0,557234 \rightarrow \infty\text{-E}$	73%
	Ridor	$\sigma E > 0,5721 \rightarrow \infty\text{-E}$	77%
	CN2	$Segm(\delta=8, m=64) > 5 \ \& \ \sigma E > 0,5572 \rightarrow \infty\text{-E}$	71%

En la misma tabla se nota que los eventos monoestados toman valores menores a 0,24 y los multiestados lo contrario. Un punto de corte igual a 0,24 quiere decir que en los eventos multiestados al menos una de sus tensiones de fase experimenta un cambio de tensión entre estados consecutivos superior al 24% de la tensión de prefalta.

El conjunto de reglas tiene buena cobertura excepto para aquellas que contienen el atributo $Segm(\delta=8, m=64)$. Nótese como este atributo tiene punto de corte igual a cinco segmentos, lo cual es correcto pues los eventos monoestados tienen cuatro segmentos, mientras que los multiestados más de cinco.

5.6 Reglas extraídas para eventos aguas arriba/abajo

A partir del conjunto de reglas extraídas en este escenario (Tabla 19), se puede inferir que los atributos relevantes para la identificación de la localización relativa de la fuente de eventos de tensión, son en su orden de relevancia: I_{ratio} , $I \cdot \cos(\phi)$ y $\Delta\phi$. Este último tiene relevancia especialmente en eventos cuyo origen se encuentra aguas abajo del punto de medida.

Tabla 19. Reglas extraídas de acuerdo a la localización relativa de la fuente del evento

Clase	Algt.	Regla extraída	Cobert
Aguas Arriba	CR	$I_{ratio} \leq 0,988 \rightarrow A\text{-Arriba}$	96%
	OneR	$I_{ratio} < 0,989 \rightarrow A\text{-Arriba}$	95%
	PART	$I_{ratio} \leq 0,981 \ \& \ I \cdot \cos(\phi) \leq 34,15 \ \& \ Rex > 0 \rightarrow A\text{-Arriba}$	95%
	CN2	$0,324 < I_{ratio} \leq 0,981 \ \& \ I \cdot \cos(\phi) \leq 14,23 \rightarrow A\text{-Arriba}$	88%
Aguas Abajo	JRip	$I_{ratio} \geq 0,804 \ \& \ I \cdot \cos(\phi) \geq 14,68 \rightarrow A\text{-Abajo}$	95%
	JRip	$I_{ratio} \geq 0,996 \ \& \ I \cdot \cos(\phi) \geq 0 \rightarrow A\text{-Abajo}$	96%
	OneR	Ninguna regla extraída	--
	PART	$\Delta\phi \leq 0,07 \ \& \ I \cdot \cos(\phi) > 6,66 \rightarrow A\text{-Abajo}$	96%
	Ridor	$0,993 < I_{ratio} \leq 2,24 \ \& \ \Delta\phi \leq -0,036 \rightarrow A\text{-Abajo}$	96%
CN2	$0,542 < I_{ratio} \leq 2,22 \ \& \ \Delta\phi \leq -0,037 \rightarrow A\text{-Abajo}$	84%	

Los puntos de corte obtenidos por los algoritmos de extracción de reglas no distan de los puntos de corte teóricos propuestos por los autores de cada uno de los algoritmos de localización relativa

de donde fueron extraídos cada uno de los atributos utilizados. Por ejemplo, el punto de corte teórico para I_{ratio} es la unidad, y los puntos sugeridos por los algoritmos oscilan alrededor de este valor. Similarmente esto ocurre con el atributo $\Delta\phi$ cuyo punto de corte teórico es cero. Con el atributo $I \cdot \cos(\phi)$ no ocurre lo mismo porque el valor del punto de corte depende de cómo se calcule el atributo. En este trabajo $I \cdot \cos(\phi)$ fue calculado a través de integración numérica.

De la Tabla 19 se concluye que para eventos aguas abajo I_{ratio} y $I \cdot \cos(\phi)$ toman valores mayores que la unidad y mucho mayores que cero, respectivamente. Para eventos aguas arriba I_{ratio} y $\Delta\phi$ toman valores menores que la unidad y menores que cero, respectivamente.

6. Diagnóstico de Eventos Tensión

En esta sección se describe la utilización de las reglas extraídas como unidades de decisión dentro de metodologías para el diagnóstico de eventos de tensión.

A partir del análisis de las reglas extraídas, se generó una metodología para el diagnóstico de eventos de tensión (Figura 10). La reglas incluidas en la metodología corresponden a aquellas con mayor cobertura y frecuencia de aparición. Por ejemplo, la regla $\sigma\delta L > 0,12$ fue seleccionada a partir de los resultados de los algoritmos *PART*, *CN2*, *JRip* y *Ridor* (Tabla 15). Debido a los distintos puntos de corte, éste fue redondeado a 2 cifras decimales. Inicialmente la metodología discrimina si el evento es evolutivo y lo clasifica como multiestado. En caso contrario (evento monoestado), se continua el proceso de clasificación. Los eventos evolutivos se identifican previamente porque éstos no se pueden clasificar como balanceados o desbalanceados, debido a que algunos de ellos presentan varios tipos de falta durante la evolución del evento de tensión.

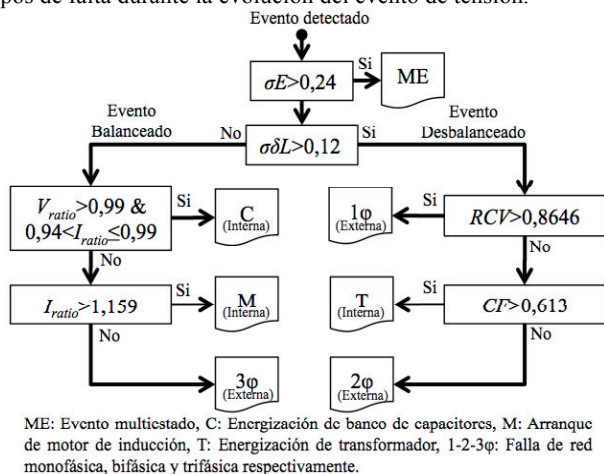


Figura 10. Metodología para el diagnóstico de eventos de tensión en sistemas de distribución radiales y sin presencia de generación distribuida.

Una vez determinado que el evento sólo contiene un estado estable (monoestado), se identifica su grado de desbalance, es decir, si éste es balanceado o desbalanceado. Lo anterior es posible gracias a que el atributo $\sigma\delta L$ es capaz de discriminar correctamente el 100% de los eventos balanceados y desbalanceados. Luego, se determina la clase de cada evento tal y como se muestra en la Figura 10.

En la Tabla 20 se presentan los resultados de clasificación de la metodología. Las filas indican la clase real de los eventos, mientras que las columnas la clase en la cual fue clasificado cada evento por la metodología. Por lo tanto, los elementos de la diagonal principal corresponden a los eventos clasificados correctamente, mientras que aquellos fuera de la diagonal a los mal clasificados.

La metodología clasificó correctamente 495 eventos de un total de 522 (94,83%). Un total de 27 eventos fueron clasificados en clases erróneas, 6 monofásicos, 1 bifásico, 8 trifásicos, 5 multiestados, 1 energización de transformador y 6 energización de capacitores.

Tabla 20. Resultados de clasificación de la metodología

Clase real	Clase estimada								Total
	1φ	2φ	3φ	ME	ET	C	M		
1φ	197	0	0	6	0	0	0	203	
2φ	0	207	0	0	1	0	0	208	
3φ	0	0	17	1	0	6	1	25	
ME	0	5	0	32	0	0	0	37	
ET	0	1	0	0	10	0	0	11	
C	0	0	6	0	0	22	0	28	
M	0	0	0	0	0	0	10	10	
Total	197	213	23	39	11	28	11	522	

Las 6 de las 25 faltas trifásicas fueron mal clasificadas porque una pequeña porción de la red fue aislada por las protecciones durante cada una de estas perturbaciones, y por tanto, el nivel de tensión se incrementó ($V_{ratio} > 1$) y el de corriente se redujo ($I_{ratio} < 1$). Este efecto ocasionó que estos 6 eventos trifásicos fueran confundidos con eventos debidos a la energización de bancos de capacitores, tal y como se había mencionado en la sección anterior.

7. Conclusión

Se definieron y analizaron eléctrica y estadísticamente atributos relacionados con la causa de los eventos, su localización relativa, su grado de desbalance y su cantidad de estados estacionarios a lo largo del evento.

Se identificaron los atributos relevantes en cada uno de los escenarios mediante un análisis multivariante de la varianza (MANOVA).

Las reglas extraídas en cada escenario de clasificación fueron discutidas, y se verificó que los atributos más frecuentes en las reglas corresponden a los identificados como relevantes en el análisis MANOVA.

Se propuso una metodología para el diagnóstico de eventos de tensión y se validó con 522 observaciones. La metodología identifica las faltas internas y externas. De las faltas internas identifica su causa y de las externas la cantidad de fases involucradas en el evento.

Las reglas extraídas y la metodología propuesta pueden ser aplicadas sólo a redes de distribución radial sin presencia de generación distribuida. Los atributos relacionados con potencias pueden verse afectados por la capacidad de cortocircuito de punto de conexión.

English Summary

Relevant Attributes for Voltage Event Diagnosis in Power Distribution Networks

Abstract

This paper focuses on diagnosis of voltage events collected in power distribution networks. Fault networks, induction motor starting, transformer energization and capacitor bank switching cause voltage events. A characterization of voltage events using attributes directly extracted from the voltage and current waveforms is done in this paper. The used attributes are highly related with the event root-cause as well as the relative location of the event source with respect to the measurement point. The relevance of each attribute has been assessed applying a statistical analysis of variance (MANOVA). The most relevant attributes have been used as input to rule-extraction algorithms in order to extract classification rules. The results were obtained using 484 real-world and 38 synthetic voltage events.

Keywords:

Statistical analysis, power quality, attributes, voltage sag event, rule based system.

Agradecimientos

Este trabajo fue financiado por el gobierno de España y Colombia a través de las siguientes entidades, proyectos y becas: MEC (BES-2007-14942, DPI 2006- 09370), Universitat de Girona (UdG), COLCIENCIAS (1110-467-21786), Universidad Industrial de Santander (UIS). Los autores extienden sus agradecimientos a ENDESA DISTRIBUCIÓN por facilitar los eventos registrados en sus circuitos de distribución.

Referencias

- Barrera V., Berjaga X., Melendez J., Herraiz S., (2008). Two new methods for voltage sag source location. *13th International Conference on Harmonics & Quality of Power*, 28 de Septiembre a 1 de Octubre, Australia.
- Barrera V., Meléndez J., Herraiz S. (2009). Evaluation of Fault relative location algorithms using voltage sag data collected at 25-kv substations. *European Transactions on Electrical Power*, **20**, 34-51.
- Barrera V., Meléndez J., Kulkarni S., Santoso S. (2010) Feature analysis and automatic classification of short-circuit faults resulting from external causes. *European Transactions on Electrical Power*, DOI: 10.1002/etep.674, January 2012.
- Barrera V., Bollen, M., Yu-Hua Gu I., Meléndez, J. (2010). Feature characterization of power quality events according to their underlying causes. *14th International Conference on Harmonics & Quality of Power*, 26-29 de Septiembre, Italy.
- Barrera V., Kulkarni S., Santoso S., Meléndez J. (2010) SVM-Based classification methodology for overhead distribution fault events. *14th International Conference on Harmonics & Quality of Power*, 26-29 de Septiembre, Italy.
- Barrera V., Kulkarni S., Santoso S., Melendez J. (2010) Feature analysis and classification methodology for overhead distribution fault events. *IEEE Power & Energy Society, 2010 General Meeting*. 25-29 de Julio, USA.
- Blanco J., Jagua J., Barrera, V., Jaimes, L. (2009) *Metodología para el diagnóstico de la causa de huecos de tensión: análisis de fallas*, Tesis de grado, Publicaciones UIS, <http://tangara.uis.edu.co/biblioweb/tesis/2009/132235.PDF> Colombia.
- Bollen M. (2000) *Understanding Power Quality Problems: Voltages Sags and Interruptions*, IEEE PRESS, New York.
- Bollen M. (2003) Algorithms for characterizing measured three-phase unbalanced voltage dips. *IEEE Transactions on Power Delivery*, **18**, 937 - 944.
- Bollen, M., Yu-Hua, I. Axelberg, P., Styvaktakis, E. (2007) Classification of underlying causes of power quality disturbances: Deterministic versus Statistical methods. *EURASIP Journal on Advances in Signal Processing*, **2007**, 1-17.
- Clark P, Niblett T. (1989) The CN2 induction algorithm. *Machine Learning Journal*, **3/4**: 261-283.
- Clark P and R. Boswell. (1991) Rule induction with CN2: Some recent improvements. In Y. Kodratoff, editor, *Machine Learning - EWSL-91*, Springer-Verlag: Berlin, 151-163.
- Coury, D.V., Tavarez C.J. (1998) Transient analysis resulting from shunt capacitor switching in an actual electrical distribution system, *8th International Conference on Harmonics & Quality of Power*, Athens.
- Grid 2030 - A National Vision for Electricity's Second 100 Years (2003), United States Department of Energy.
- Hamzah N., Mohamed A., Hussain A. (2005) Locating voltage sag source at the point of common coupling in industrial distribution systems. *IEEE International Conference on Power Electronics and Drive Systems*. 28-1 de Noviembre, Malasia.
- Hernández, J, Quintana, M.J, Ramírez, C. (2004) *Introducción a la minería de datos*. Pearson Prentice Hall, Addison-Wesley.
- Ibrahim, W., Morcos M. (2002) "Artificial Intelligence and Advanced Mathematical Tools for Power Quality Applications: A Survey", *IEEE Transactions on Power Delivery*, **17**, 668.
- Kersting, W. H. (2000) *Radial distribution test feeders*. Power engineering Society, USA.
- Kulkarni S., Lee D., Allen A., Santoso S., Short T., (2010) *Waveform Characterization of Animal Contact, Tree Contact, and Lightning Induced Faults*, IEEE Power & Energy Society, 2010 General Meeting. 25-29 de Julio, USA.
- McGranaghan M, B. Roettger. (2002) "Economic evaluation of power quality". *IEEE Power Engineering Review*, **22** (2), 8-12.
- McGranahan M. (2001) "Trends in Power Quality Monitoring", *IEEE Power Engineering Review*.
- Pradhan A.K, Routray A., Madhan S. (2007). Fault direction estimation in radial distribution system using phase change in sequence current. *IEEE Transactions on Power Delivery*, **22**, 2065-2071.
- Seon A., Dong W., Li C., Seung M. (2004) "Determination of the Relative Location of Voltage sag source According to event Cause", *Power Engineering Society General Meeting*, **1**, 620 - 625.
- Styvaktakis E., (2002) *Automating power quality analysis*, Ph.D. thesis, Chalmers University of Technology, Sweden, 2002.
- Tayjasantan T., C. Li, and W. Xu. (2005) A resistance sign-based method for voltage sag source detection. *IEEE Transactions on Power Delivery*, **20**, 2544-51, 2005.
- Velandia R., Hernández F., Barrera, V., Vargas, H. (2010) *Evaluación de algoritmos de extracción de reglas de decisión para el diagnóstico de huecos de tensión*, Tesis de grado, Publicaciones Universidad Industrial de Santander, Colombia.
- Xu L, Mo-Yuen C. (2006) A classification approach for power distribution systems fault cause identification, *IEEE Transaction on Power System*; **21**, 53-60.
- Xu L, Chow M, Taylor L. (2007) Power Distribution Fault Cause Identification With Imbalanced Data Using the Data Mining-Based Fuzzy Classification E-Algorithm. *IEEE Transaction on Power Systems*, **22**, 164-171.
- Yalcinkaya G., Bollen MHJ., Crossley P.A. (1998) Characterization of voltage sags in industrial distribution Systems. *IEEE Transactions on Industry Applications*, **34**, 682-688.
- Yixin C, Mo-Yuen C, Wenbin L, Lexin L. (2010) Evaluation of distribution fault diagnosis algorithms using ROC curves, 25 a 29 de Julio, USA.
- Yixin C, Mo-Yuen C, Wenbin L, Lexin L. Statistical Feature Selection From Massive Data in Distribution Fault Diagnosis. *IEEE Transaction on Power Systems* 2010; **25** (2), 642-648.