



Universitat de Girona

**LOCALIZACIÓN DE FALTAS EN SISTEMAS DE  
DISTRIBUCIÓN DE ENERGÍA ELÉCTRICA  
USANDO MÉTODOS BASADOS EN EL  
MODELO Y MÉTODOS BASADOS EN EL  
CONOCIMIENTO**

**Juan José MORA FLÓREZ**

**ISBN: 978- 84-690-4513-8**

**Dipòsit legal: GI-314-2007**



-- Tesis Doctoral --

**Localización de Faltas en Sistemas de Distribución de  
Energía Eléctrica usando Métodos Basados en el Modelo y  
Métodos de Clasificación Basados en el Conocimiento**

por

**Juan José Mora Flórez**



Programa de Doctorado en Tecnologías de la Información  
UNIVERSITAT DE GIRONA  
Girona, España 2006

---

---

Localización de Faltas en Sistemas de Distribución de Energía Eléctrica usando Métodos Basados en el Modelo y Métodos de Clasificación Basados en el Conocimiento

JUAN JOSÉ MORA FLÓREZ

© JUAN JOSÉ MORA FLÓREZ, 2006

Escola Politècnica Superior  
Teléfono:+34 972 41 83 91  
Campus Montilivi, Bloque P4  
Universitat de Girona  
Girona – Spain

Campus Montilivi  
Girona, España 2006

## Resumen

La calidad de energía eléctrica incluye la calidad del suministro y la calidad de la atención al cliente. La calidad del suministro a su vez se considera que la conforman dos partes, la forma de onda y la continuidad. En esta tesis se aborda la continuidad del suministro a través de la localización de faltas.

Este problema se encuentra relativamente resuelto en los sistemas de transmisión, donde por las características homogéneas de la línea, la medición en ambos terminales y la disponibilidad de diversos equipos, permiten localizar el sitio de falta con una precisión relativamente alta. En sistemas de distribución, sin embargo, la localización de faltas es un problema complejo y aún no resuelto. La complejidad es debida principalmente a la presencia de conductores no homogéneos, cargas intermedias, derivaciones laterales y desbalances en el sistema y la carga. Además, normalmente, en estos sistemas sólo se cuenta con medidas en la subestación, y un modelo simplificado del circuito. Los principales esfuerzos en la localización han estado orientados al desarrollo de métodos que utilicen el fundamental de la tensión y de la corriente en la subestación, para estimar la reactancia hasta la falta. Como la obtención de la reactancia permite cuantificar la distancia al sitio de falta a partir del uso del modelo, el Método se considera Basado en el Modelo (MBM). Sin embargo, algunas de sus desventajas están asociadas a la necesidad de un buen modelo del sistema y a la posibilidad de tener varios sitios donde puede haber ocurrido la falta, esto es, se puede presentar múltiple estimación del sitio de falta.

Como aporte, en esta tesis se presenta un análisis y prueba comparativa entre varios de los MBM frecuentemente referenciados. Adicionalmente se complementa la solución con métodos que utilizan otro tipo de información, como la obtenida de las bases históricas de faltas con registros de tensión y corriente medidos en la subestación (no se limita solamente al fundamental). Como herramienta de extracción de información se utilizan y prueban dos técnicas de clasificación (LAMDA y SVM). Éstas relacionan las características obtenidas de la señal, con la zona bajo falta y se denominan en este documento como Métodos de Clasificación Basados en el Conocimiento (MCBC).

La información que usan los MCBC se obtiene de los registros de tensión y de corriente medidos en la subestación de distribución, antes, durante y después de la falta. Los registros tomados en la subestación se procesan para obtener los siguientes descriptores: a) la profundidad del hueco de tensión ( $dV$ ), b) la variación de la magnitud de corriente ( $dI$ ), c) la variación de la potencia ( $dS$ ), d) la reactancia de falta ( $Xf$ ), e) la frecuencia del transitorio ( $f$ ), y f) el valor propio máximo de la matriz de correlación de corrientes ( $Sv$ ), cada uno de los cuales ha sido seleccionado por facilitar la localización de la falta. A partir de estos descriptores, se proponen diferentes conjuntos de entrenamiento y validación de los MCBC, y mediante una metodología que muestra la posibilidad de hallar relaciones entre estos conjuntos y las zonas en las cuales se presenta la falta, se seleccionan los de mejor comportamiento.

Los resultados de aplicación, demuestran que con la combinación de los MCBC con los MBM, se puede reducir el problema de la múltiple estimación del sitio de falta. El MCBC determina la zona de falta, mientras que el MBM encuentra la distancia desde el punto de medida hasta la falta, integración en un esquema híbrido toma las mejores características de cada método. En este documento, lo que se conoce como híbrido es la combinación de los MBM y los MCBC de una forma complementaria.

Finalmente y para comprobar los aportes de esta tesis, se propone y prueba un esquema de integración híbrida para localización de faltas en dos sistemas de distribución diferentes. Tanto los métodos que usan los parámetros del sistema y se fundamentan en la estimación de la impedancia (MBM), como aquellos que usan como información los descriptores y se fundamentan en técnicas de clasificación (MCBC), muestran su validez para resolver el problema de localización de faltas. Ambas metodologías propuestas tienen ventajas y desventajas, pero según la teoría de integración de métodos presentada, se ha logrado una alta complementariedad, permitiendo así la formulación de híbridos que mejoran los resultados, reduciendo o evitando el problema de la múltiple estimación de la zona en falta.

# Tabla de contenido

<b>1</b>	<b><u>Introducción</u></b>	<b>1</b>
1.1	Motivación	1
1.2	Faltas en sistemas de potencia	1
1.2.1	Tipos de faltas	1
1.2.2	Modelado de faltas	2
1.2.3	Resistencia de falta	3
1.3	Localización de faltas en sistemas de potencia	3
1.3.1	Sistemas de transmisión	3
1.3.2	Sistemas de distribución	4
1.3.3	Efecto de la resistencia de falta	4
1.3.4	Efecto de las cargas intermedias	5
1.4	Definición y acotación del problema	5
1.5	Definición del sistema de pruebas prototipo	6
1.6	Aportes de la tesis	6
1.7	Estructura del documento	7
<b>2</b>	<b><u>Métodos basados en el modelo (MBM)</u></b>	<b>10</b>
2.1	Introducción	10
2.2	Definiciones y nomenclatura fundamentales	10
2.3	Modelo generalizado del sistema de distribución para localización de faltas	11
2.4	Métodos de localización de faltas basados en el modelo	12
2.4.1	Método de la componente reactiva	12
2.4.1.1	Resumen de la técnica	12
2.4.1.2	Estimación de la distancia de falta	12
2.4.2	Método de Srinivasan et. al.	13
2.4.2.1	Resumen de la técnica	13
2.4.2.2	Estimación de distancia a la falta	13
2.4.3	Método de Girgis et. al.	14
2.4.3.1	Resumen de la técnica	14
2.4.3.2	Estimación de distancia de falta	14
2.4.4	Método de Zhu et. al.	15
2.4.4.1	Resumen de la técnica	15
2.4.4.2	Estimación de distancia de falta	16
2.4.5	Método de Aggarwal et. al.	16
2.4.5.1	Resumen de la técnica	16
2.4.5.2	Estimación de distancia de falta	17
2.4.6	Método de Ratan Das	18
2.4.6.1	Resumen de la técnica	18
2.4.6.2	Estimación de la distancia de falta	18
2.4.7	Método de Novosel et. al.	21
2.4.7.1	Resumen de la técnica	21
2.4.7.2	Estimación de distancia de falta	22
2.4.8	Método de Yang	23
2.4.8.1	Resumen de la técnica	23
2.4.8.2	Estimación de distancia de falta	23
2.4.9	Método de Saha et. al.	25
2.4.9.1	Resumen de la técnica	25
2.4.9.2	Estimación de la sección bajo falta	25
2.4.9.3	Estimación de distancia a la falta	26
2.4.10	Método de Choi et al.	26
2.4.10.1	Resumen de la técnica	26
2.4.10.2	Estimación de distancia de falta	26
2.5	Comparación de los MBM para localización faltas	27

2.6	Conclusiones	28
<b>3</b>	<b><u>Métodos de Clasificación Basados en el Conocimiento (MCBC)</u></b>	<b>29</b>
3.1	Introducción	29
3.2	Algunas aplicaciones de sistemas basados en el conocimiento a la localización de faltas en redes de distribución	29
3.3	Definición de clasificación	32
3.4	Técnicas de clasificación	32
3.5	Criterios de evaluación del método de clasificación	34
3.6	Máquinas de Soporte Vectorial	36
3.6.1	Generalidades	36
3.6.2	Metodología de SVM	36
3.6.2.1	Análisis del caso linealmente separable	37
3.6.2.2	Análisis del caso linealmente no separable	39
3.6.3	Multclasificación con SVM	41
3.6.3.1	Máquina multclasificadora	41
3.6.3.2	Máquinas biclasificadoras generalizadas	42
3.6.4	Razón para utilizar SVM en localización de faltas	44
3.7	Clasificación basada en LAMDA	44
3.7.1	Generalidades	44
3.7.2	Metodología del LAMDA	45
3.7.3	Función de pertenencia para el cálculo del DAM	46
3.7.3.1	Función de pertenencia para los espacios de descripción cuantitativos	47
3.7.3.2	Función de pertenencia para los espacios de descripción cualitativos	48
3.7.4	Aprendizaje	48
3.7.4.1	Aprendizaje supervisado de atributos cuantitativos	48
3.7.4.2	Aprendizaje supervisado de atributos cualitativos.	49
3.7.4.3	Aprendizaje no supervisado	50
3.7.5	Operadores lógicos de agregación para el cálculo del GAG	50
3.7.5.1	Conectivos mixtos de compensación	50
3.7.5.2	Sumas simétricas	51
3.7.6	Razón para utilizar LAMDA en localización de faltas	52
3.8	Conclusiones	52
<b>4</b>	<b><u>Caracterización de las señales de tensión y corriente para localización de faltas en sistemas de distribución</u></b>	<b>53</b>
4.1	Introducción	53
4.2	Definición de los descriptores	53
4.3	Caracterización del sistema mediante la obtención de descriptores	54
4.3.1	Profundidad del hueco de tensión ( $dV$ )	54
4.3.2	Variación de la magnitud de corriente ( $dI$ )	56
4.3.3	Variación de la potencia del sistema ( $dS$ )	56
4.3.4	Reactancia de falta ( $X_f$ )	58
4.3.5	Frecuencia del transitorio ocasionado por la falta ( $f$ ).	59
4.3.6	Valor propio máximo de la matriz de correlación de corrientes ( $S_v$ )	62
4.3.7	Firma del sistema	63
4.4	Variación de los descriptores propuestos como entrada a los MCBC	64
4.5	Conclusiones	67
<b>5</b>	<b><u>Aplicación de los MBM y MCBC para localización de faltas en sistemas de distribución</u></b>	<b>68</b>
5.1	Introducción	68
5.2	Localización de faltas mediante los Métodos Basados en el Modelo - MBM	68
5.2.1	Sistema de prueba y escenarios	68
5.2.2	Resultados en la estimación de la distancia de falta	69
5.2.3	Análisis comparativo de resultados	70
5.3	Localización de faltas mediante los Métodos de Clasificación basados en el	74

	Conocimiento - MCBC	
5.3.1	Metodología para localizar la zona bajo falta usando MCBC	74
5.3.1.1	Zonificación de la red	75
5.3.1.2	Definición de la estructura de los MCBC	75
5.3.1.3	Adquisición de la base de datos de falta	75
5.3.1.4	Pre-procesamiento de la señal	75
5.3.1.5	Entrenamiento del MCBC	76
5.3.1.6	Pruebas de precisión del MCBC	76
5.3.2	Aplicación de LAMDA y SVM a la localización de faltas – Pruebas comparativas	76
5.3.2.1	Zonificación de la red	76
5.3.2.2	Adquisición de la base de datos de falta	77
5.3.2.3	Pre-procesamiento de la señal	77
5.3.2.4	Entrenamiento y pruebas de precisión de los localizadores MCBC	77
5.3.2.5	Pruebas de precisión de los localizadores MCBC	77
5.3.2.6	Análisis de resultados	77
5.3.3	Perfeccionamiento de la estrategia de localización de faltas usando SVM	78
5.3.3.1	Zonificación de la red	78
5.3.3.2	Definición de las estructuras del localizador basado en SVM	79
5.3.3.3	Adquisición de datos de falta	79
5.3.3.4	Pre-procesamiento de la señal	80
5.3.3.5	Estrategia de entrenamiento de las SVM	80
5.3.3.6	Pruebas de precisión del localizador basado en SVM	82
5.3.3.7	Análisis de resultados a partir del comportamiento de los descriptores	82
5.3.4	Análisis del desempeño del localizador de faltas basado en SVM ante la variación de carga en el sistema de distribución	83
5.3.4.1	Definición de las pruebas	83
5.3.4.2	Resultados de las pruebas	83
5.3.4.3	Análisis de resultados	84
5.4	Conclusiones	86
<b>6</b>	<b><u>Estructuras híbridas para localización de faltas</u></b>	<b>88</b>
6.1	Introducción	88
6.2	Comparación entre los MCBC y MBM	88
6.3	Sistemas híbridos	89
6.3.1	Formas de integración de métodos	89
6.3.1.1	Sistemas de componente simple	89
6.3.1.2	Sistemas basados en fusión	89
6.3.1.3	Sistemas jerárquicos	90
6.3.1.4	Sistemas híbridos	90
6.4	Arquitectura híbrida para localización de faltas	91
6.4.1	Estructura básica de integración híbrida	91
6.4.2	Instanciación de la arquitecturas híbrida al problema de localización de faltas	92
6.4.3	Metodología de reducción de la múltiple estimación del sitio de falta	92
6.5	Resultados de la metodología aplicada al sistema de prueba	94
6.6	Pruebas de la metodología híbrida aplicada al sistema de 34 nodos de la IEEE	95
6.6.1	Sistema de distribución IEEE de 34 nodos	95
6.6.2	Zonificación del sistema	95
6.6.3	Configuración del MCBC	95
6.6.4	Resultados del MCBC	96
6.6.5	Resultados del MBM	101
6.6.6	Resultados del híbrido	102
6.6.7	Análisis de resultados	103
6.7	Conclusiones	103
<b>7</b>	<b><u>Conclusiones, recomendaciones y trabajo futuro</u></b>	<b>104</b>
7.1	Conclusiones generales	104
7.2	Conclusiones asociadas a los MBM	105

7.3	Conclusiones asociadas a los MCBC	105
7.4	Conclusiones asociadas a la caracterización de señales de tensión y corriente	106
7.5	Conclusiones asociadas a la aplicación de los MBM y los MCBC	107
7.6	Conclusiones asociadas a los métodos híbridos	108
7.7	Recomendaciones	108
7.8	Trabajo futuro	109
	<b><u>Bibliografía</u></b>	<b>111</b>
	<b><u>Anexo A. Sistema de pruebas prototipo</u></b>	<b>118</b>
	<b><u>Anexo B. Resumen de algunos de los algoritmos de aprendizaje</u></b>	<b>120</b>
	<b><u>Anexo C. Tablas con resultados de las pruebas de los MBM para localización de faltas</u></b>	<b>126</b>
	<b><u>Anexo D. Perfeccionamiento de las redes LAMDA para localización de faltas</u></b>	<b>130</b>
	<b><u>Anexo E. Sistema de pruebas de 34 nodos de la IEEE</u></b>	<b>136</b>

## Listado de figuras

Figura 1.1. Conexión de las redes de secuencia para faltas a) monofásica, b) fase- fase-tierra, c) fase-fase, d) trifásica a tierra y e) trifásica.	2
Figura 1.2. Efecto de la resistencia de falta para a) sistemas alimentados por un solo terminal y b) sistemas alimentados por dos terminales.	4
Figura 1.3. Diagrama unifilar del sistema de 25 kV de Saskatoon Power and Light seleccionado para pruebas.	6
Figura 2.1. Modelo generalizado de un sistema de distribución	11
Figura 2.2. Modelo generalizado de una sección de línea	12
Figura 2.3. Efecto de la resistencia de falta para sistemas alimentados por un solo	12
Figura 2.4. Circuito de las componentes superimpuestas	17
Figura 2.5. Sistema radial equivalente para una falta en F	19
Figura 2.6. Modelo simplificado del sistema de distribución	22
Figura 2.7. Circuito de las componentes superimpuestas.	22
Figura 2.8. Modelo simplificado del sistema de potencia utilizado para hallar la impedancia equivalente vista desde la subestación hasta un nodo cualquiera K	25
Figura 2.9. Modelo simplificado del sistema de potencia utilizado para hallar la impedancia equivalente vista desde la subestación hasta un nodo cualquiera K	26
Figura 3.1. Principales técnicas de clasificación, según su estructura típica de aprendizaje: a) Supervisados, b) no supervisados y c) semisupervisados	34
Figura 3.2. Hiperplanos que separan correctamente un conjunto de datos. a) Hiperplano de separación de datos. b) OSH con un mayor margen de separación entre clases.	37
Figura 3.3. Hiperplano lineal clasificador para el caso no separable.	39
Figura 3.4. La SVM no lineal establece una relación del espacio de entrada con otro de representación de dimensión alta.	40
Figura 3.5. Esquema general del método LAMDA. Adecuación del objeto $x$ a la clase $C_j$	46
Figura. 3.6. Función para el calculo de GAM para diferentes valores de $\rho$ ( $q = 1 - \rho$ )	47
Figura 4.1. Descriptores de la variación del valor eficaz de la tensión	54
Figura 4.2. Sistema de distribución seleccionado para pruebas	55
Figura 4.3. Señales de tensión para el caso de una falta monofásica A-T, en la barra 4 y con una resistencia de falta de $26\Omega$	55
Figura 4.4. Señales de tensión para el caso de una falta monofásica A-T, en la barra 11, con una resistencia de falta de $4\Omega$ .	56
Figura 4.5. Descriptores de la variación del valor eficaz de la corriente.	56
Figura 4.6. Modelo agregado de carga en la subestación	58
Figura 4.7. Modelo agregado de carga y falta en la subestación	58
Figura 4.8. Modelo agregado de impedancias en la subestación	59
Figura 4.9. Procedimiento para obtener los descriptores asociados a la señal transitoria.	60
Figura 4.10. Señal transitoria de tensión en la fase B	61
Figura 4.11. Detalles wavelet para la tensión de la fase C	61
Figura 4.12. Análisis del detalle 4 mediante la FFT	62
Figura 4.13. Curvas de coordinación de tiempo típicas del esquema de salvamento de fusibles.	64
Figura 4.14. Corrientes de línea medidas en la subestación durante una falta en el nodo 25 de la figura 4.1	64
Figura 4.15. Variación del valor eficaz de la tensión entre estado estable de falta y prefalta. Hueco de tensión	65
Figura 4.16. Variación del valor eficaz de la corriente entre los estados estables de falta y prefalta	65

Figura 4.17. Valor propio máximo de la transformada Clark-Concordia de las señales de corriente durante la falta	66
Figura 4.18. Variación de la potencia aparente entre los estados estables de falta y prefalta	66
Figura 4.19. Reactancia de la fase en falta	67
Figura 4.20. Frecuencia del transitorio medida en la fase en falta	67
Figura 5.1. Sistema prototipo de prueba. Circuito de 25kV de Saskatoon Power and Light de la ciudad de Saskatoon, Canadá.	69
Figura 5.2. Mejores resultados en la estimación de la distancia a la falta. Falta monofásica con $R_F = 0,05 \Omega$	71
Figura 5.3. Mejores resultados en la estimación de la distancia a la falta. Falta monofásica con $R_F = 25 \Omega$	71
Figura 5.4. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica con $R_F = 0,05 \Omega$	72
Figura 5.5. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica con $R_F = 25 \Omega$	72
Figura 5.6. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica a tierra con $R_F = 0,5 \Omega$	72
Figura 5.7. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica a tierra con $R_F = 25 \Omega$	73
Figura 5.8. Mejores resultados en la estimación de la distancia a la falta. Falta trifásica con $R_F = 0,5 \Omega$	73
Figura 5.9. Mejores resultados en la estimación de la distancia a la falta. Falta trifásica con $R_F = 25 \Omega$	74
Figura 5.10. Estructura general de los MCBC para localización de la falta.	75
Figura 5.11. Zonificación del sistema de potencia para localización de faltas usando MCBC – Cinco zonas	76
Figura 5.12. Zonificación del sistema de potencia para localización de faltas usando SVM	79
Figura 5.13. Estructura de localización de zona dependiente del tipo de falta - DTF	79
Figura 5.14. Estructura de localización de zona No dependiente del tipo de falta - NDTF	79
Figura 5.15. Ejemplo gráfico de la selección de $C$ y de $\sigma$ mediante búsqueda de malla	81
Figura 5.16. Comportamiento del descriptor de frecuencia en la fase en falta monofásica para diferentes condiciones de carga	85
Figura 5.17. Comportamiento del descriptor de reactancia en la fase en falta monofásica para diferentes condiciones de carga	86
Figura 5.18. Comportamiento del descriptor de reactancia una fase que no esta en falta, ante una falta monofásica para diferentes condiciones de carga	86
Figura 6.1. Esquema del sistema basado en fusión	90
Figura 6.2. Esquema del sistema jerárquico	90
Figura 6.3. Esquema del sistema híbrido	90
Figura 6.4. Arquitectura funcional de un híbrido.	91
Figura 6.5. Sistema híbrido basado en MCBC y MCBC para localización de faltas	92
Figura 6.6. Aplicación del principio de intersección para resolver el problema de la múltiple estimación de la localización de la falta. a) Posibles localizaciones ( $F_1$ , $F_2$ , $F_3$ y $F_4$ ), detectadas en diferentes secciones por un método basado en la estimación de la impedancia. b) Activación de la zona 5 como zona de falta	93
Figura 6.7. Zonificación del sistema prototipo de la IEEE de 34 nodos, de tensión nominal de 24.9 kV	95

## Listado de tablas

Tabla 2.1. Tensiones y corrientes seleccionadas de un sistema trifásico para el cálculo de la impedancia aparente	15
Tabla 2.2. Fase de referencia para los diferentes tipos de faltas	20
Tabla 2.3. Tensiones y corrientes para diferentes tipos de falta	23
Tabla 2.4. Sustituciones para faltas en múltiples fases	25
Tabla 2.5. Comparación de los MBM para la localización de faltas.	28
Tabla 3.1. Comparación de los métodos analizados basados en el conocimiento para localización de faltas.	31
Tabla 3.2. Funciones de aprendizaje ( $q = 1 - \rho$ )	49
Tabla 5.1. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta $R_f = 0,05\Omega$	69
Tabla 5.2. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta $R_f = 5\Omega$	70
Tabla 5.3. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta $R_f = 25\Omega$	70
Tabla 5.4. Cantidad de registros de falta usados para el entrenamiento y en las pruebas de precisión	77
Tabla 5.5. Resultados de precisión para cada uno de los grupos de descriptores de la prueba preliminar de los MCBC.	78
Tabla 5.6. Cantidad de datos de entrenamiento y prueba del localizador basado en SVM	80
Tabla 5.7. Valores de la constante C y el parámetro $\sigma$ del kernel RBF, obtenidos como resultado de la validación cruzada y la búsqueda en malla	81
Tabla 5.8. Resultados de la prueba de los localizadores basados en SVM para una zonificación de siete zonas del sistema de prueba.	82
Tabla 5.9. Resultados de precisión para el caso de carga diferente a la carga nominal	84
Tabla 5.10. Resultados de precisión promedio para distintos tipos de carga	85
Tabla 6.1. Comparación entre los MBM y los MCMC. Ventajas y desventajas	89
Tabla 6.2. Resultados de las pruebas con el método híbrido integrado	94
Tabla 6.3. Cantidad de registros de falta usados para el entrenamiento y en las pruebas de precisión.	96
Tabla 6.4. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FT	96
Tabla 6.5. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FF	97
Tabla 6.6. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FFT	97
Tabla 6.7. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-3F	98
Tabla 6.8. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FT	98
Tabla 6.9. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FF	99
Tabla 6.10. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FFT	99
Tabla 6.11. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-3F	100
Tabla 6.12. Resultados de precisión promedio en la prueba para todas las combinaciones de uno a cuatro descriptores	100

Tabla 6.13. Resultados de la pruebas con el MBM en caso de faltas a tierra	101
Tabla 6.14. Resultados de la pruebas con el MBM en caso de faltas bifásicas	101
Tabla 6.15. Resultados de la pruebas con el MBM en caso de faltas bifásicas a tierra	101
Tabla 6.16. Resultados de la pruebas con el MBM en caso de faltas trifásicas	101
Tabla 6.17. Resultados de la pruebas con el método híbrido integrado. Caso de faltas monofásicas a tierra	102
Tabla 6.18. Resultados de la pruebas con el método híbrido integrado. Caso de faltas bifásicas	102
Tabla 6.19. Resultados de la pruebas con el método híbrido integrado. Caso de faltas bifásicas a tierra	102
Tabla 6.20. Resultados de la pruebas con el método híbrido integrado. Caso de faltas trifásicas	102

# 1 Introducción

## 1.1 Motivación

La calidad de la energía eléctrica se ha convertido en un tema de estudio de gran interés para los operadores de red y para los usuarios. Por ello, la actividad investigativa en este campo ha experimentado un fuerte avance en las últimas décadas. Entre los aspectos más importantes de la calidad se tienen en cuenta, la forma de onda, la continuidad del servicio y la atención al cliente [BOLL00]. La importancia actual de estos aspectos está asociada a que la participación del capital privado en el sector eléctrico obliga a establecer pautas de remuneración e índices mínimos de calidad [CREG02] [GELL02]. La calidad del producto implica satisfacer condiciones de calidad de onda (o de la potencia) y de continuidad.

En lo relacionado con la continuidad del servicio, tema de esta tesis, existen dos indicadores ampliamente difundidos para el caso Español: a) el índice TIEPI que mide el tiempo de interrupción equivalente de la potencia instalada, durante el cual la energía no es suministrada; y b) el índice NIEPI que determina el número total de interrupciones equivalentes de la potencia instalada [LE5497][RD1955]. Los índices de naturaleza similar, pero que tiene mayor difusión a nivel mundial son el SAIDI y el SAIFI [BILL84]. La determinación de índices facilita el establecimiento de límites de parte de las entidades reguladoras, de tal forma que los operadores de red y los usuarios están obligados a satisfacerlos so pena de pagar compensaciones por su incumplimiento [BOLL02] [HEYD01].

Por otra parte, la continuidad del suministro se ve afectada por muchas causas entre las cuales la ocurrencia de faltas en la red es la más importante. Por tanto, la realización de esta tesis está motivada en su mayor parte por la presunción que: *“La eficiente y oportuna localización de las faltas, permitirá a las distribuidoras de energía eléctrica, mejorar sus índices de continuidad del servicio”*. La rápida localización de faltas permanentes, permitirá disminuir la duración de las interrupciones (TIEPI o SAIDI), mientras que la continua monitorización y localización de faltas (permanentes y transitorias), permite determinar las debilidades del sistema, y de esa forma fortalecer la red para que la ocurrencia de faltas sea cada vez menor (NIEPI o SAIFI). Adicionalmente, y como consecuencia de las acciones operativas posteriores a la localización de la falta (conmutación de seccionadores), es posible reducir el área afectada y restablecer prontamente zonas aledañas, sin que la pérdida transitoria del servicio alcance a ser contabilizada como interrupción.

Finalmente, el problema de localización de faltas en sistemas de distribución, es un problema complejo y aún no está resuelto debido a las características propias de estas redes [IEEE04]. Como visión de aporte a su solución, se plantea complementar los métodos clásicos de localización basados en el análisis circuital del modelo, mediante técnicas de clasificación, para reducir la incertidumbre asociada a la localización de la falta.

## 1.2 Faltas en sistemas de potencia

### 1.2.1 Tipos de faltas

Los tipos de faltas que experimentan los sistemas de potencia se caracterizan en dos grandes grupos: serie y paralelo [ANDE95]. Las faltas serie están asociadas generalmente a la ruptura de conductores, pero éstos no involucran contacto con tierra u otra fase. Las faltas paralelo están asociadas a cortocircuitos a tierra o con otras fases y este grupo se subdivide en cinco tipos: a) Monofásicas, b) Bifásicas, c) Bifásicas a tierra, d) Trifásicas y e) Trifásicas a tierra. Los métodos comúnmente utilizados para localización en líneas aéreas, hacen referencia a las faltas de tipo paralelo, dadas sus altas tasas de ocurrencia [CROZ99]. Las faltas tipo serie o de alta impedancia, están fuera del alcance de esta tesis.

Estudios han determinado que aproximadamente el 80% del total de faltas en el sistema de potencia, corresponden a faltas en el sistema de distribución [BOLL93]. Del total de faltas en sistemas de distribución, aproximadamente el 70% corresponde a faltas monofásicas a tierra [WEST64] [RAVI77].

### 1.2.2 Modelado de faltas

El modelado clásico de las faltas en los sistemas de potencia utiliza la teoría de las componentes simétricas [ANDE95]. Los diferentes tipos de faltas para sistemas equilibrados se modelan como se presenta en la figura 1.1 y se resume a continuación. Para obtener el modelo, se consideran las condiciones de falta, se obtienen las componentes de secuencia de las corrientes y las tensiones, y se determina la conexión de las redes de secuencia. A partir del modelo de la falta, cuando ésta ocurra, es necesario calcular las tensiones y corrientes de secuencia y finalmente obtener los valores de fase.

El modelo general de una falta monofásica a tierra de la fase A, a través de una resistencia de falta  $R_g$ , se representa como la interconexión en serie de las redes de secuencia, positiva, negativa y cero, y tres veces el valor de la resistencia de falta.

Las faltas bifásicas entre las fases B y C, a través de una resistencia  $R_f$ , donde simultáneamente existe una falta a tierra entre las mismas fases a través de una resistencia  $R_g$ , está modelada como la interconexión paralelo de las redes de secuencia positiva, negativa y cero. Las redes de secuencia positiva y negativa incluyen la resistencia  $R_f$ , mientras que la red de secuencia cero incluye la resistencia de  $R_f$  y tres veces la resistencia de falta a tierra  $R_g$ .

Para el caso de las faltas bifásicas entre las fases B y C, a través de una resistencia  $R_f$ , sin conexión a tierra, el modelo se obtiene de la interconexión paralelo de las redes de secuencia positiva y negativa. La red de secuencia cero permanece aislada.

Finalmente, para el caso de faltas trifásicas que no involucran tierra, las redes de secuencia permanecen independientes. Si se considera que el sistema es equilibrado y operaba antes de la falta bajo condiciones balanceadas, las redes de secuencia negativa y cero no tienen fuentes.

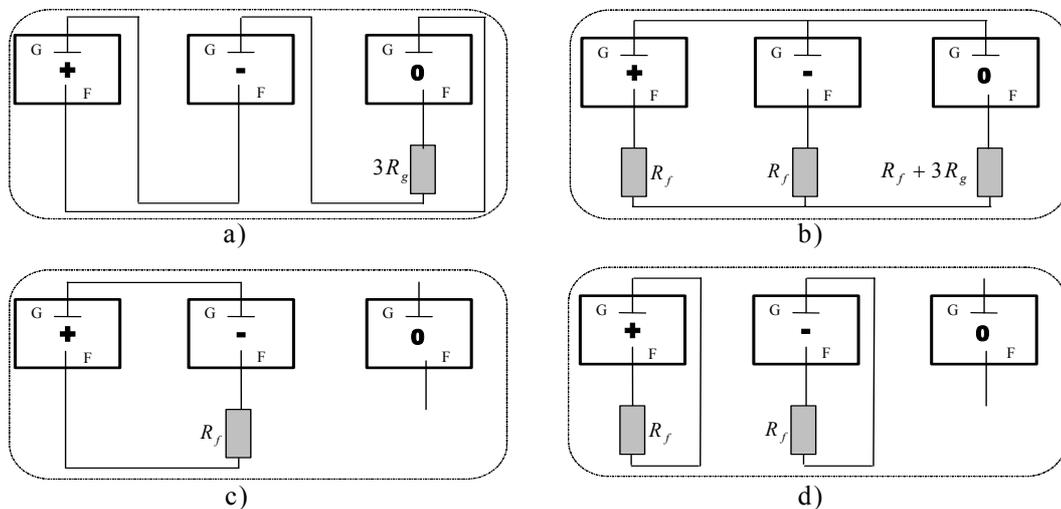


Figura 1.1. Conexión de las redes de secuencia para faltas a) monofásica, b) fase- fase-tierra, c) fase-fase, d) trifásica.

Para las faltas en las fases mencionadas, la referencia de las componentes simétricas es la fase A. Las faltas en fases diferentes a las mencionadas, se analizan a partir del cambio en la referencia de las componentes simétricas [PHAD95].

### 1.2.3 Resistencia de falta

Otro aspecto de fundamental importancia en este análisis es el relacionado con la resistencia de falta. La mayoría de las faltas es del tipo resistivo y pueden tener resistencia de arco y de tierra [WARR68]. Esta resistencia puede ser constante para toda la duración de la falta o puede variar

con la elongación y extinción del arco. En faltas fase-fase, la resistencia de falta se debe completamente al arco. Para faltas que involucran la tierra, la resistencia de falta incluye ambos tipos de resistencia. En [DAS98] se presenta un análisis adicional de la resistencia de arco y la resistencia de tierra. Sin embargo en [DAGE00] se considera que para el modelado de faltas paralelas en líneas de distribución, los valores típicos de resistencia de falta a considerar son menores o iguales a  $40\Omega$ .

### **1.3 Localización de faltas en sistemas de potencia**

Los métodos de localización de faltas más rudimentarios consistían en la inspección visual, el seccionamiento, la reparación y la posterior energización de los circuitos en falta reportados por los clientes. Estos métodos consumen mucho tiempo debido a la extensión de las redes, y además pueden ser peligrosos para las personas o causar daños en los equipos, cuando por cualquier circunstancia se energiza el circuito cuando todavía está en falta. Normalmente estos métodos se complementan con el uso de detectores de paso de corrientes de falta, ubicados estratégicamente en las estructuras de soporte a lo largo de los alimentadores del sistema. Estas alternativas hacen que el método sea económicamente más costoso y normalmente, bajo condiciones adversas de clima o por falta de mantenimiento, estos indicadores pueden deteriorarse y salir de servicio [DAS98].

#### **1.3.1 Sistemas de transmisión**

Para el caso específico de localización de faltas en líneas de transmisión, el problema se ha abordado mediante la utilización de las señales de tensión y corriente medidas en uno o en ambos terminales de la línea. Estos métodos comúnmente se dividen en tres categorías, los basados en las ondas viajeras, los basados en las componentes de alta frecuencia y los basados en la componente fundamental [DAS98].

Los métodos basados en las ondas viajeras aprovechan que éstas se inician cuando ocurre una falta y viajan por la red y desde este sitio hasta los terminales de la línea. Estos métodos tienen muy buen desempeño en la medida en que se conozcan los parámetros inductivos y capacitivos de la línea, la configuración de la red y tanto la clase y precisión como el ancho de banda del transductor utilizado en la toma de los registros. Este método tiene la desventaja del coste de implementación [ANCE94].

De otra parte, la medición de las componentes de alta frecuencia en las que se fundamenta el segundo grupo de métodos, tienen la desventaja adicional de su complejidad y alto coste debido a que requieren de filtros sintonizados para la medición de estas señales [TAKA81].

Los métodos basados en la medición del fundamental de tensión y de corriente de los terminales de la línea, estiman la impedancia equivalente desde el punto de medición. La impedancia estimada a partir de las mediciones, es comparada luego con una impedancia equivalente calculada con base en los parámetros de la línea. Para el caso de líneas de transmisión, estos métodos normalmente usan las medidas tomadas en ambos extremos de la línea, y consideran una línea homogénea y sin cargas intermedias. Estos métodos han sido preferidos por las empresas encargadas de las redes de transmisión debido a su simplicidad y bajo coste [DAS98].

#### **1.3.2 Sistemas de distribución**

Para el caso de la localización de faltas en sistemas de distribución, se prefiere el último grupo de métodos por las ventajas de simplicidad y bajo coste. Sin embargo, debido a que la explotación de estos sistemas se hace de forma radial (aunque la construcción sea mallada), la opción de medidas en ambos extremos de la línea no es posible en la práctica. Adicionalmente, los métodos de localización de faltas aplicados a distribución deben considerar la no homogeneidad de la línea, la topología radial y altamente ramificada de la red, la presencia de circuitos laterales monofásicos y la presencia de cargas intermedias.

Una desventaja de estos métodos, originada por la medición en un solo terminal, es que cuando se tiene una topología altamente ramificada se puede presentar una múltiple localización del sitio de falta, debido que la impedancia equivalente puede coincidir con varios nodos, cada uno en un ramal

diferente. Otra desventaja de estos métodos es la alta dependencia de los parámetros del modelo de la línea y de la carga, en cuanto a su exactitud.

Para solucionar en parte los problemas anteriores, algunos autores han propuesto el uso de información adicional que complemente las medidas (pre-falta y falta) de la fundamental de la tensión y corriente [LEE04][GIRG93]. Esta información está principalmente asociada a considerar las características de las señales de corriente, originadas por la activación de elementos de protección y control como reconectores, seccionadores y fusibles, ubicados a lo largo del sistema de distribución. Esta alternativa involucra generalmente la implementación de sistemas SCADA, lo cual tiene la desventaja de incrementar el coste de la solución.

### 1.3.3 Efecto de la resistencia de falta

En caso de sistemas de transmisión, que normalmente tienen alimentación por ambos terminales de la línea, se presenta un incremento en el valor de la impedancia con el aumento de la resistencia de falta. Sin embargo, el valor de la reactancia no es constante, debido al desfase de la caída de tensión en la resistencia de falta. Este fenómeno es ocasionado por la corriente que proviene de la fuente equivalente del otro extremo de la línea, que normalmente no está en fase con la corriente medida, tal como se presenta en la figura 1.2a.  $I_{mf}$  es la corriente de falta medida en M;  $I_f$  es la corriente de falta total que es igual a la suma de  $I_{mf}$  e  $I_{nf}$ . El problema radica en que no se conoce la corriente  $I_{nf}$ . La admitancia aparente  $Z_m$  estimada con los valores medidos en M es igual a la impedancia  $sZ_{mn}$  más la resistencia de falta  $R_f$  afectada por las dos corrientes de falta [PHAD95].

Para los sistemas de distribución y como consecuencia de la presencia de resistencia de falta, el valor de la impedancia aparente, es decir la estimada a partir de las medidas de tensión y de corriente en la subestación, es mayor que la impedancia de la línea desde la subestación hasta el punto de falta. Para este caso, la reactancia desde la subestación hasta el sitio de falta permanece constante, con cualquier valor de resistencia de falta y si no se considera el efecto de la carga. En la figura 1.2b se presenta el caso para una falta con una resistencia  $R_f$ , a un porcentaje  $s$  de la distancia MN de la línea [WARR68].  $Z_m$ , es la impedancia aparente y  $sZ_{mn}$  representa la impedancia de la línea hasta la falta.

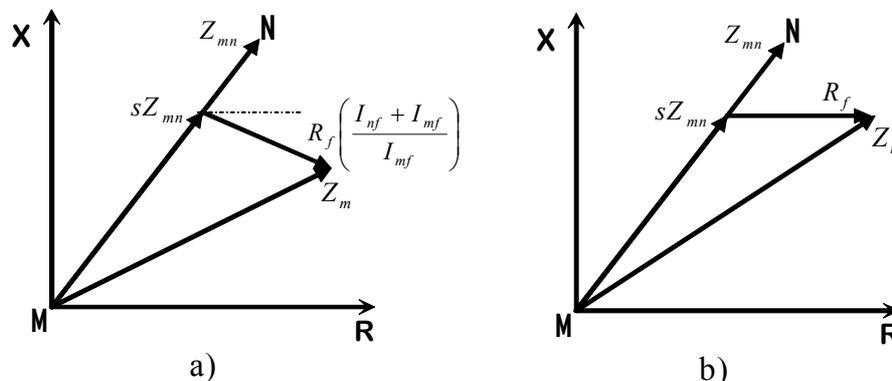


Figura 1.2. Efecto de la resistencia de falta para a) sistemas alimentados por dos terminales y b) sistemas alimentados por un solo terminal, sin carga

### 1.3.4 Efecto de las cargas intermedias

Las cargas intermedias causan problemas en la exactitud de la estimación de la distancia, debido que la corriente medida en la subestación, no corresponde a la corriente de falta sino a la suma de ésta con la de carga a la tensión de falta [IEEE04]. Estas corrientes no siempre son fáciles de estimar, si no se cuenta con un modelo adecuado de la carga. Por lo tanto la impedancia medida no corresponde a la de la línea entre la subestación y el sitio de falta.

## 1.4 Definición y acotación del problema

El problema de la localización de faltas en sistemas eléctricos no es nuevo. Las primeras investigaciones que trataron este problema se enfocaron en la localización de faltas en líneas de transmisión, mediante la utilización de las medidas tomadas en ambos extremos de una línea normalmente homogénea [IEEE04]. En los últimos años y motivados por la nueva estructura competitiva del mercado eléctrico, se han incentivado en gran medida los estudios relacionados con la localización del sitio de falta en sistemas de distribución, donde el problema es más complejo

Para el caso de los sistemas de transmisión, el problema está parcialmente resuelto debido a la disponibilidad de sofisticados equipos de medición en las subestaciones de inicio y finalización de la línea, las características como la homogeneidad y la ausencia de cargas intermedias entre las subestaciones de transmisión. Sin embargo, el problema no está totalmente resuelto para faltas de alta impedancia o para faltas evolutivas. Para este último caso, debido al cambio en las condiciones de la falta, los algoritmos de localización disponibles pueden tener errores considerables [CHOI04].

La localización de faltas en sistemas de distribución es un problema de mayor complejidad debido a que estos sistemas pueden estar desbalanceados, tener una topología radial con muchas ramificaciones o laterales monofásicos, cargas intermedias, y tener conductores de diferentes calibres. Adicionalmente, debido a la gran cantidad de subestaciones de distribución y la comparativamente baja potencia que alimentan, estos sistemas cuentan con muy poca instrumentación. En el mejor de los casos, se limita a medidores de tensión y corriente, algunas protecciones como relés 50/51 en la subestación, y elementos propios de automatización como reconectores, seccionadores y fusibles ubicados a lo largo del sistema [PHAD95].

De otra parte, los primeros métodos utilizados para localización de faltas consisten en una inspección visual de los circuitos en falta, reportados por los clientes. Otros métodos implementados posteriormente, dada la poca practicidad de los anteriores, utilizan las medidas de tensión y corriente en ambos o en un extremo de la línea. Estos últimos métodos se dividen en tres categorías: Métodos basados en ondas viajeras, métodos que usan componentes de alta frecuencia y métodos que usan el valor de la componente fundamental de tensión y corriente o mejor conocidos como métodos basados en la impedancia. Los de esta última categoría se pueden clasificar también como métodos que usan medidas en un terminal de la línea y los que usan medidas en ambos terminales [DAS98].

Adicionalmente, los métodos basados en la estimación de la impedancia utilizan los parámetros de modelo del sistema y los valores eficaces del fundamental de tensión y corriente de pre-falta y falta, para calcular la distancia entre el punto de medida, que normalmente es la subestación, y el sitio de falta. Debido a este aspecto y a la topología radial altamente ramificada, esta distancia puede coincidir en varios sitios del sistema y por tanto estos métodos tienen el problema de la múltiple estimación del sitio de falta. Adicionalmente, la exactitud de la estimación es altamente dependiente de un buen modelo del sistema [ZHU97].

Considerando lo expuesto y la revisión bibliográfica previa, el problema está relacionado con la localización de faltas paralelas en sistemas de distribución, el cual no ha sido completamente resuelto. Particularmente, esta tesis está orientada al análisis y la propuesta de una solución utilizando medidas en un solo terminal de la línea, que en este caso corresponde a la subestación de distribución.

Para la solución del problema se propone el uso de la información proveniente de las medidas de tensión y de corriente mediante métodos de clasificación. Esto permitirá reducir el problema de la múltiple estimación del sitio de falta, así como la alta dependencia de un buen modelo del sistema de distribución, propio de los métodos basados en la estimación de la impedancia.

## 1.5 Definición del sistema de pruebas prototipo

El sistema de distribución seleccionado para la realización de las pruebas de los diferentes métodos es un circuito de 25kV de Saskatoon Power and Light de la ciudad de Saskatoon, Canadá presentado en [DAS98] y frecuentemente citado y utilizado por otros investigadores del problema

de la localización de faltas. En la figura 1.3 se presenta el diagrama unifilar del sistema prototipo. Los parámetros de las líneas y carga se presentan en el anexo A.

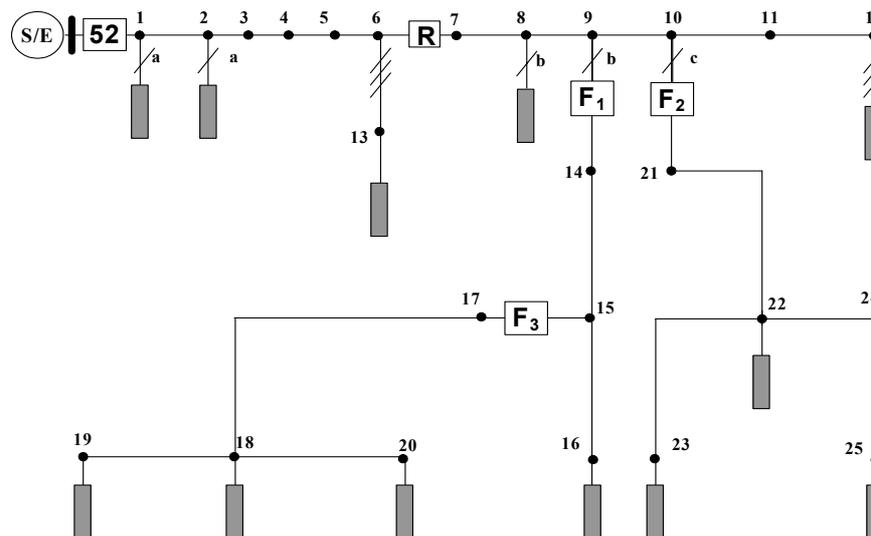


Figura 1.3. Diagrama unifilar del sistema de 25 kV de Saskatoon Power and Light seleccionado para pruebas. Adicionalmente, para la implementación y simulación de faltas en el circuito de pruebas se utiliza el “Alternative Transients Program”(ATP)[LEUV87]. Se realizó una simulación extensiva de faltas con las que se configuró una base de datos conformada por 2982 faltas, en donde se incluyen trifásicas, bifásicas, bifásicas a tierra y monofásicas para 21 valores diferentes de resistencia de falta (desde 0,05 hasta 40  $\Omega$ ).

## 1.6 Aportes de la tesis

La propuesta aquí presentada permite el aprovechamiento de información adicional a los valores eficaces de la tensión y la corriente medidos en la subestación. Con esto se logran reducir los problemas de la múltiple estimación de la localización de la falta y la alta dependencia del modelo.

La metodología propuesta para localización de faltas consiste en una estructura híbrida que permite unir los métodos que estiman la impedancia a partir del fundamental de tensión y corriente y los parámetros de la red, denominados en este documento como Métodos Basados en el Modelo (MBM), y métodos que aprovechan otro tipo de información como características de carga, activación de protecciones, comportamiento de los huecos de tensión, entre otros, denominados en este documento como Métodos de Clasificación Basados en el Conocimiento (MCBC).

Entre los aportes de la tesis se tienen:

- Análisis comparativo de varios de los más comúnmente referenciados métodos basados en el modelo para la localización de faltas.
- Análisis de los métodos basados en el conocimiento y su posible aplicación a la solución del problema de la localización de faltas.
- Análisis de las señales de tensión y corriente medidas durante faltas en sistemas de distribución, para obtener características, que en esta tesis se denominarán en adelante como “descriptores”. Éstos son la entrada de los MCBC y le permiten establecer una relación entre las medidas y el sitio de falta.
- Aplicación de técnicas de clasificación como núcleo de los MCBC para resolver el problema de la localización de faltas, enfatizando en las Máquinas de Soporte Vectorial, dado su mejor desempeño.
- Aplicación del método de validación cruzada en la selección de los parámetros del clasificador, logrando que con diferentes conjuntos de descriptores se obtenga un buen resultado y se pueda cambiar el paradigma de usar más datos para el entrenamiento que para la prueba.

- Propuesta de una arquitectura genérica e instanciación de la misma, para desarrollar métodos híbridos para localización de faltas.
- Prueba de la estructura híbrida desarrollada, mostrando su alto desempeño en el problema de localización de faltas en sistema de distribución.

## 1.7 Estructura del documento

Este documento de tesis está organizado en seis capítulos. El capítulo inicial corresponde a la introducción y contiene la presentación formal del documento de tesis.

En el capítulo dos, inicialmente se presentan la nomenclatura y las definiciones fundamentales; luego se presentan los métodos que estiman la impedancia desde el punto de medida, a partir de los valores eficaces de pre-falta y falta del fundamental de tensión y corriente y los parámetros de la red; denominados como Métodos Basados en el Modelo (MBM). Posteriormente, en este capítulo se presenta el estado del arte de los métodos más relevantes por su exactitud y actualidad, para luego realizar un análisis comparativo de los mismos. Finalmente se presentan las conclusiones del capítulo.

En el capítulo tres, se hace una presentación generalizada de los métodos que aprovechan otro tipo de información, que en este documento se denomina genéricamente como “descriptores”. Se presenta por tanto un estado del arte de los MCBC que se pueden adaptar al problema de la localización de faltas. Adicionalmente, en este capítulo se analiza el tipo de entradas o atributos usados como entradas del método, y su desempeño en términos de exactitud y condiciones de validez de las aplicaciones. También se presentan los fundamentos y la justificación de su uso en el problema de localización de faltas de dos técnicas de clasificación: a) el Algoritmo de Aprendizaje para el Análisis de Datos Multivariantes (del inglés LAMDA), y b) las Máquinas de Soporte Vectorial (del inglés SVM). Las conclusiones se presentan al final del capítulo.

La caracterización del sistema de potencia, para extraer información útil para resolver el problema de la localización de faltas en sistemas de distribución se presenta en el capítulo cuatro. Se presentan los descriptores y las técnicas de obtención de los mismos, los cuales servirán como entrada de los MCBC. Algunos de los más importantes están asociados a la variación del fundamental de tensión y corriente, la potencia aparente entre pre-falta y falta, la frecuencia de la señal transitoria durante la falta, la reactancia de falta y la forma de la corriente a consecuencia de la activación de las protecciones. En la parte final del capítulo se presentan las conclusiones.

En el capítulo cinco se presentan los resultados de las pruebas de localización y el análisis de los resultados usando tanto los MBM descritos en el capítulo dos, como los dos MCBC presentados en el capítulo tres, junto con los descriptores obtenidos en el capítulo cuatro. Se presentan, asimismo, los aportes de los MCBC sobre los MBM en el problema de la localización de faltas. Adicionalmente a la comparación entre LAMDA y las SVM, y debido a los mejores resultados, se presenta una metodología de utilización de las SVM, para las cuales se hace un estudio detallado para la localización de faltas en el sistema prototipo de prueba. Finalmente se presentan las conclusiones del capítulo.

En el capítulo seis inicialmente se presenta una comparación entre los MBM y los MCBC. Posteriormente se presentan los fundamentos conceptuales básicos de la conformación de estructuras híbridas y su instanciación para el problema de localización de faltas. A partir de estas estructuras instanciadas, se presentan las pruebas y el análisis de resultados para el sistema prototipo. Adicionalmente, se presenta un ejemplo de aplicación con un sistema de distribución “IEEE 34 Node Test Feeder” propuesto por el “Distribution System Analysis Subcommittee” donde se demuestra la validez de la propuesta de esta tesis. Las conclusiones del capítulo se presentan en la parte final y, enfatizan en los aportes de las estructuras híbridas sobre el funcionamiento individual de los MBM y MCBC.

Finalmente, en el capítulo siete se presentan las conclusiones, las recomendaciones y el trabajo futuro como resultado de esta investigación, haciendo énfasis en los aportes de la tesis.

## 1.8 Contribuciones derivadas de esta tesis

Esta disertación está parcialmente fundamentada en los aportes presentados en las siguientes publicaciones (una tesis de maestría, diez publicaciones en conferencias internacionales y seis publicaciones aceptadas en revistas, dos publicaciones con la primera revisión en revista y dos más enviadas a revistas y aún sin comentarios):

- [1] J. Mora, D. Llanos, J. Meléndez. “Classification of sags measured in a distribution substation using a fuzzy tool” ICREPQ'2003, International Conference on Renewable Energy and Power Quality. Vigo-Spain 9-11, April 2003.
- [2] J. Mora, D. Llanos, J. Meléndez. “Classification of Sags Measured in a Distribution Substation Based on Qualitative and Temporal Attributes” CIRED 2003, 17th International Conference on Electricity Distribution. Barcelona-Spain 12-15, May 2003.
- [3] J. Mora, D. Llanos, J. Meléndez. “Classification of short duration faults in transmission and distribution power systems” ECC'03, European Control Conference. University of Cambridge (UK). 1-4 September, 2003.
- [4] J. Mora. “Voltage Sag Characterization and Classification for Diagnosis in Electric Power Quality Domain”, Master dissertation. University of Girona, España. 2003.
- [5] J. Mora, J. Meléndez, J. Colomer. “Fault location in distribution systems. A voltage sag approach” ICREPQ'2004, International Conference on Renewable Energy and Power Quality. Barcelona (Spain) March 2004.
- [6] J. Mora, J. Meléndez. “Caracterización de huecos de tensión para localización de faltas en sistemas de distribución - un primer paso hacia un método híbrido”. Revista Scientia et Technica No 25. (ISSN 0122-1701). Noviembre de 2004. Pag. 45-52.
- [7] J. Mora, C. Cortés, J. Gutiérrez “Técnica de localización de faltas para un sistema de potencia radial, con cargas laterales desequilibradas y circuitos no homogéneos” Revista Scientia et Technica. N.28. (ISSN 0122-1701). Septiembre 2005, Pag. 56-62.
- [8] J. Mora, S. Muñoz, G. Carrillo “Técnicas algorítmicas de localización de faltas como alternativa para reducir el efecto de las salidas en sistemas de potencia – Una revisión”. III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov. 2005.
- [9] J. Mora, H. Serrano, G. Ordóñez “Mejoramiento de los Índices de Calidad del Servicio mediante la Caracterización de Señales para Localización de Faltas en Distribución”. III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov 2005.
- [10] J. Mora, J. Meléndez, G. Carrillo “Una Arquitectura Genérica para el Desarrollo de Sistemas Híbridos. Propuesta de localizador de faltas para mejorar los índices de calidad”. III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov. 2005.
- [11] J. Mora, M. Granada, L. Marín. “Métodos de representación del conocimiento en inteligencia artificial y su integración en sistemas híbridos de localización de faltas”. Revista Tecnura (ISSN 012-3921X), año 9, No. 17, II Semestre de 2005, Pag. 98 – 109.
- [12] J. Mora, J. Rodríguez, S. Pérez. “Método híbrido basado en la estructura de agentes para localización de faltas en sistemas de distribución de energía eléctrica”. Revista Scientia et Técnica. No 29 de 2006. (ISSN 0122-1701). Mayo de 2006. Pag. 59-65.
- [13] J. Mora, G. Carrillo, L. Pérez. “Fault location in power distribution systems using ANFIS nets and current patterns”. In proceedings of 2006 IEEE PES Transmission and Distribution Conference and Exposition, Caracas, August, 2006.
- [14] J. Mora, J Melendez, J. Bedoya. “Extensive Events Database Development using ATP and Matlab to Fault Location in Power Distribution Systems” In proceedings of 2006 IEEE PES Transmission and Distribution Conference and Exposition, Caracas, August 2006.
- [15] J. Mora, J Meléndez, G. Carrillo. “Fault Location in Power Distribution Systems Based on Signal Descriptors” In proceedings of 2006 IEEE 12th International Conference on Harmonics and Quality of Power, Porto, October, 2006.

- [16] J. Mora, J. Meléndez, J. Bedoya. “Implementación de Protecciones y Simulación Automática de Eventos para Localización de Faltas en Sistemas de Distribución de Energía”. Revista Ingeniería y Competitividad. Número 1, Agosto de 2006 (ISSN 0123-3033). Pag. 43-51
- [17] J. Mora, V. Barrera, G. Carrillo. “Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis”. IEEE Transactions on Power Delivery. 2006 (*Accepted paper*)
- [18] J. Mora, G. Olguin, H. Serrano. “Fault Location in Power Distribution Systems Using Characterization of Voltages and Currents”. IEEE Transactions on Power Delivery. 2006 (*RI. paper*)
- [19] J. Mora, J. Cormane, G. Ordóñez. “Improving of Power Continuity Indexes in Distribution Systems by using a Statistical Model for Fault Location”. ETEP European Transactions on Power Systems. 2006 (*RI. paper*)
- [20] J. Mora, J. Meléndez, G. Carrillo. “Comparison between power distribution system fault location methods”. Electric Power System Research. 2006 (*Submitted paper*)
- [21] J. Mora, J. Meléndez, G. Carrillo. “SVM based Fault Locator applied in Power Distribution Systems”. IEEE Transaction on Power Delivery. 2006 (*Submitted paper*)

El autor ha participado en las siguientes publicaciones, que no están completamente direccionados en esa disertación, pero son consideradas por su relación con el trabajo aquí propuesto:

- [22] J. Mora, P. Warren, J. Colomer “Hybrid Approach Based On Temporal Representation And Classification Techniques Used To Determine Unstable Conditions In A Blast Furnace”, Current Topics in Artificial Intelligence (ISBN: 3-540-22218-9), May 2004. pag 125-136.
- [23] J. Mora, J. Colomer “Statistical analysis and a conceptual clustering method to rank clients of a financial bank”. Revista Scientia et Technica, N. 25 de 2004. (ISSN 0122-1701). Pag 25-75
- [24] J. Mora, J. Jaimes, G. Carrillo “Reducción de la indisponibilidad durante faltas en subestaciones de transmisión de energía eléctrica” Revista Tecnura N 15 de 2005. Pag. 77-83. (ISSN 0132-1851)
- [25] J. Mora, J. Manrique, G. Ordóñez “Efecto del deslizamiento de frecuencia en la estimación de magnitudes eléctricas mediante la DFT”. Revista Scientia et Technica N.27. (ISSN 0122-1701). Abril 2005. Pag 27-34

## 2 Métodos basados en el modelo (MBM)

### 2.1 Introducción

Los sistemas eléctricos de potencia han venido creciendo aceleradamente en las últimas tres décadas, y esto ha traído como consecuencia una ampliación de cobertura y el aumento de líneas en operación. Estas líneas experimentan faltas, que en su mayoría corresponden al grupo de faltas paralelas, las cuales son causadas por descargas atmosféricas, ruptura del aislamiento, tormentas, nieve y cortocircuitos causados por animales u otros objetos extraños.

En el caso de faltas permanentes, para la restauración del servicio a los clientes, es necesario reparar la línea y los daños mecánicos de las estructuras. El proceso de reparación es mucho más expedito si se detecta con rapidez y exactitud la localización de la falta. La rápida localización de faltas consideradas como permanentes, permitirá disminuir la duración de las interrupciones (TIEPI). De otra parte, la continua monitorización y localización de las faltas permanentes y transitorias permitirá a las empresas determinar las debilidades de su sistema y de esa forma podrán fortalecer la red, de modo que la ocurrencia de faltas sea cada vez menor (NIEPI). Adicionalmente, mediante la localización de fallas, se pueden desarrollar estrategias operativas para aislar el área fallada y restaurar prontamente las áreas aledañas, mediante estrategias de conmutación de seccionadores, evitando que muchos eventos de falla sean considerados como una interrupción [MORA05-a].

En los últimos años, el interés en el tema de localización de faltas ha aumentado debido que muchas de las empresas suministradoras de electricidad están involucradas en el ambiente competitivo y desregulado del nuevo mercado eléctrico. Uno de los aspectos diferenciadores en esta competencia está asociado al aseguramiento de la continuidad del servicio a los clientes, el cual está estrechamente relacionado con el problema de la localización de faltas [IEEE04].

Para el caso de sistemas de distribución rígidamente puestos a tierra, el problema de localización de faltas es mucho más complejo que en los sistemas de transmisión. Algunos de los factores más importantes que afectan la exactitud de los métodos de estimación de faltas son: a) inexistencia de medición de tensiones y corrientes en las cargas intermedias, b) Variabilidad de la impedancia de las cargas intermedias, c) Indiferenciabilidad entre una falta en un lateral o en el alimentador principal, d) posible desbalance del sistema y de la carga, e) variabilidad de la topología del sistema y f) posible no homogeneidad de la línea debido a cambios de conductor a lo largo de los alimentadores [ZHU97].

Como ya se explicó en el capítulo introductorio, los métodos más utilizados tanto por su facilidad de implementación como por su exactitud, son los métodos basados en la estimación de la impedancia a partir de medidas de tensión y corriente, antes y durante la falta, medidos en un solo extremo de la línea, que en este caso corresponde a la subestación de distribución [DAS98].

Para el caso de redes de distribución que no están rígidamente puestas a tierra, se utilizan configuraciones especiales para detectar contactos a tierra de los conductores. Estas redes no son de uso generalizado debido al riesgo que implican para las personas. Sin embargo, son utilizadas especialmente en industrias que requieren altos índices de continuidad en el suministro de energía [BLAC98].

Este capítulo presenta algunos métodos de localización de faltas en sistemas de distribución que utilizan los valores fundamentales de tensión y corriente de pre-falta y falta, medidas en un extremo de la línea, así como los parámetros de la red. Para propósitos de este documento, estos métodos se denominan Métodos Basados en el Modelo – MBM

### 2.2 Definiciones y nomenclatura fundamentales

M	Nodo inicial que corresponde a la subestación.
X	Nodo anterior a la falta.
F	Nodo en falta.
X+1	Nodo después de la falta.

N	Nodo final de la línea.
$s$	Distancia en p.u. desde el nodo X hasta el nodo F.
$s_T$	Distancia en p.u. desde el nodo M hasta el nodo F.
$L$	Distancia entre los nodos de la sección en falta, X y X+1.
$L_T$	Distancia entre el nodo X y el último nodo del alimentador N.
$[Z_{abc}]$	Matriz de impedancia
$[Y_{abc}]$	Matriz de admitancia
$Z_{m,r}$	Impedancia entre los nodos M y R.
$Z_{m,app}$	Impedancia aparente calculada con las mediciones de tensión y corriente del nodo M.
$Z_{m,r}^m$	Impedancia modificada entre el nodo M y R
$R_f$	Resistencia de falta
$V_m$	Tensión en el nodo M.
$V_{(a)m}$	Tensión en la fase A del nodo M.
$I_{m,r}$	Corriente que sale del nodo M hacia el nodo R.
$I_{(a)m,r}$	Corriente de la fase A, que sale del nodo M hacia el nodo R.
$I_f$	Corriente que circula por la resistencia de falta.
$V_{(0,1,2)m}$	Tensión de secuencia cero, positiva o negativa respectivamente, en el nodo M.
$I_{(0,1,2)m,r}$	Corriente de secuencia cero, positiva o negativa respectivamente, que sale del nodo M hacia el nodo R.
$[V'_{abc}]_x$	Tensión superimpuesta de las fases, a, b y c en el nodo X
$[I'_{abc}]_{x,f}$	Corriente superimpuesta de las fases, a, b y c que fluye entre el nodo X y el nodo F
A, B, C y D	Constantes del modelo de parámetros distribuidos para una línea larga
$\gamma_{m,r}$	Constante de propagación por unidad de longitud, de la sección de línea MR
$z_{m,r}^c$	Impedancia característica de la sección de línea MR
$G_r, B_r$	Constantes proporcionales a la conductancia y susceptancia respectivamente, de la carga en el nodo R (estimadas de los valores de prefalta)
$n_p, n_q$	Constantes para las componentes activa y reactiva de la carga.

### 2.3 Modelo generalizado del sistema de distribución para localización de faltas

Para el análisis de los diferentes métodos de localización de faltas, se presenta en este documento de tesis, un modelo unificado del sistema de distribución. Este modelo contiene todos los elementos utilizados por los autores aquí referenciados, para el desarrollo de su metodología en particular.

En la figura 2.1 se presenta un modelo unifilar generalizado en el cual los parámetros de las líneas se consideran como parámetros concentrados por unos autores, o distribuidos para otros, como se muestra en la siguiente sección. La presencia de cargas intermedias y su modelado, también es un aspecto que cada autor trata de forma diferente. Finalmente, una de las diferencias de los métodos es en la consideración de la presencia de circuitos laterales que se desprenden del alimentador principal. Para los circuitos reales, las medidas se tienen en la subestación (nodo M).

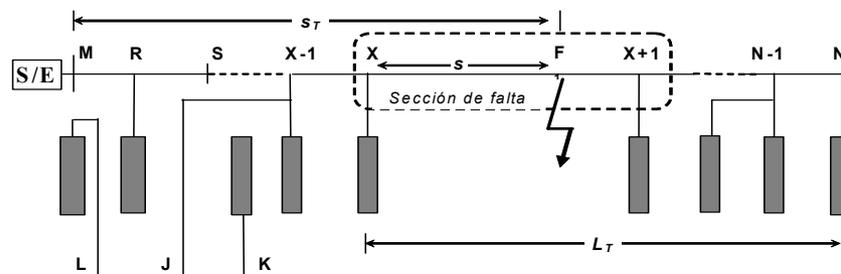


Figura 2.1: Modelo generalizado de un sistema de distribución

En la figura 2.2 se presenta un modelo trifilar generalizado para una sección de línea bajo falta, entre el nodo X y X+1 de un sistema de distribución, en donde  $R_f$  representa la resistencia de falta en el nodo F, localizado a una distancia  $s$  del nodo X.

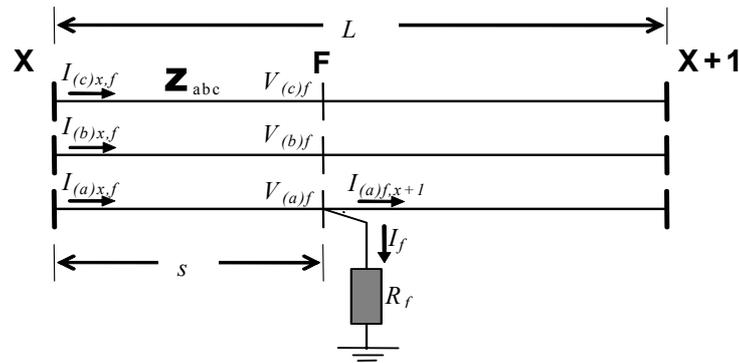


Figura 2.2: Modelo generalizado de una sección de línea

## 2.4 Métodos de localización de faltas basados en el modelo

En esta sección se presentan algunos de los métodos basados en la estimación de la impedancia o en las variaciones de la corriente y tensión de prefalta y falta, que han planteado y resuelto el problema de la localización de faltas. Se presenta un resumen general de la técnica y el método básico para hallar la distancia a la falta en p.u. de la distancia total de la sección de falta, así como la distancia total medida desde la subestación.

### 2.4.1 Método de la componente reactiva

#### 2.4.1.1 Resumen de la técnica

El método presentado en [WARR68], considera la localización de la distancia de falta entre los nodos **M** y **N**, tal como se presenta en la figura 2.1. Para encontrar la distancia  $s_T$ , se realiza la comparación de la reactancia estimada mediante los valores de tensión y de corriente a frecuencia fundamental medidos en el nodo **M** y la reactancia total de la línea entre **M** y **N**. La comparación de las reactancias atenúa el efecto de la resistencia de falta. El método de la componente reactiva ignora la presencia de cargas en las líneas.

#### 2.4.1.2 Estimación de la distancia a la falta

La impedancia aparente hasta la falta se estima a partir de las medidas de tensión y corriente tomadas bajo falta en el nodo **M**. Como consecuencia de la presencia de resistencia en la falta, el valor de la impedancia aparente es mayor que la impedancia de la línea desde la subestación hasta el punto de falta. Para el caso de sistemas alimentados desde un terminal, la reactancia desde la subestación hasta el sitio de falta permanece constante, con cualquier valor de resistencia de falta. En la figura 2.3 se presenta una falta con una resistencia  $R_f$  en la línea **M** y **N**, a un porcentaje  $s_T$  de la impedancia  $Z_{m,n}$ .  $Z_{m,app}$  es la impedancia aparente obtenida con los valores de tensión y de corriente medidos en **M**, mientras que  $Z_{m,n}$  representa la impedancia de la línea hasta el siguiente nodo.

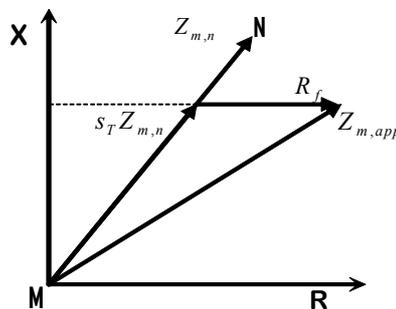


Figura 2.3. Efecto de la resistencia de falta para sistemas alimentados por un solo terminal

De la figura 2.3 se obtiene la ecuación que permite obtener el valor de la distancia en p.u. de la distancia a la falta, desde el nodo **M**, tal como se presenta en (2.1).

$$s_T = \frac{\text{Im}(Z_{m,app})}{\text{Im}(Z_{m,n})} \quad (2.1)$$

Las mediciones utilizadas para estimar la impedancia aparente, varían con el tipo de falta. En su consideración básica, este método de estimación se plantea para sistemas radiales, aunque también puede ser extendido a sistemas alimentados por los dos terminales de la línea.

## 2.4.2 Método de Srinivasan

### 2.4.2.1 Resumen de la técnica

El método propuesto en Srinivasan et. al. [SRIN89], utiliza una representación de parámetros distribuidos de la línea. A partir de esta representación se determinan las tensiones y corrientes en cada uno de los nodos del sistema. Para considerar el efecto de la carga, se tiene en cuenta un modelo dependiente de la tensión y del tipo de carga; a partir del flujo de cargas se halla la corriente en las líneas; y mediante la sustracción de la corriente de carga se obtienen las corrientes en la siguiente sección de la red.

El proceso de localización de faltas se realiza de forma iterativa. El valor inicial de la distancia al nodo en falta **F** se obtiene mediante la aproximación de la matriz de transmisión de línea larga a una línea corta. Posteriormente, y considerando que los valores de admitancia de carga son los mismos obtenidos antes de la falta, se estiman la corriente de falta, la tensión en el sitio de falta y la tensión en el nodo remoto **N**. Allí se concentran todas las cargas aguas abajo del nodo en falta. Con el valor de la tensión en cada nodo, se calcula la admitancia de las cargas, y con este valor, se estima un nuevo valor de la distancia a la falta. El proceso iterativo finaliza cuando el error sea inferior a un valor de tolerancia deseado.

### 2.4.2.2 Estimación de distancia a la falta

El método considera el modelo de línea larga con la matriz de transmisión  $T(x)$ , que permite establecer la relación entre las tensiones y las corrientes en los extremos de una sección de línea. Por ejemplo, para la sección **MR**, de longitud  $x$ , se presenta en la ecuación (2.2).

$$\begin{bmatrix} V_m \\ I_{m,r} \end{bmatrix} = T(x) \begin{bmatrix} V_r \\ I_{r,m} \end{bmatrix} = \begin{bmatrix} D_{m,r} & -B_{m,r} \\ C_{m,r} & -A_{m,r} \end{bmatrix} \begin{bmatrix} V_r \\ I_{r,m} \end{bmatrix} = \begin{bmatrix} \cosh(x\gamma_{m,r}) & -z_{m,r}^c \sinh(x\gamma_{m,r}) \\ \frac{\sinh(x\gamma_{m,r})}{z_{m,r}^c} & -\cosh(x\gamma_{m,r}) \end{bmatrix} \begin{bmatrix} V_r \\ I_{r,m} \end{bmatrix} \quad (2.2)$$

Donde  $\gamma_{m,r}$  es la constante de propagación y  $z_{m,r}^c$  es la impedancia característica de la sección **MR**.

Para la deducción de la distancia en caso de falta monofásica, se utiliza la ecuación (2.3). Como las faltas en líneas eléctricas se consideran puramente resistivas, la relación para la estimación de la distancia se obtiene de la parte imaginaria de la impedancia de falta.

$$\text{Im}(R_f) = \text{Im} \left( \frac{V_{(0)f} + V_{(1)f} + V_{(2)f}}{I_{(0)f} + I_{(1)f} + I_{(2)f}} \right) = 0 \quad (2.3)$$

En las dos ecuaciones anteriores,  $V$  e  $I$  son valores en el punto de falta **F** y los subíndices 0, 1 y 2 denotan las secuencias cero, positiva y negativa, respectivamente. Para hallar  $V$  e  $I$  en el punto de falta, se usa la matriz de transmisión  $T(s)$  entre las secciones **XF**, tal como se presenta en (2.4). El método asume **X** como el comienzo de la línea, **X+1** como el final de la línea y **F** el punto en falta.

$$\begin{bmatrix} V_f \\ I_{f,x} \end{bmatrix} = T(s) \begin{bmatrix} V_x \\ I_{x,f} \end{bmatrix} \quad (2.4)$$

La corriente que va de la falta al nodo **N** está dada por (2.5)

$$I_{f,n} = -I_{f,x} - I_f \quad (2.5)$$

A partir de (2.4) y (2.5) se pueden obtener los valores de tensión y corriente en el nodo final **N**, tal como se presenta en (2.6).

$$\begin{bmatrix} V_n \\ I_{n,f} \end{bmatrix} = T(L_T - s) \begin{bmatrix} V_f \\ I_{f,n} \end{bmatrix} \quad (2.6)$$

$L_T$  es la longitud entre el nodo **X** y nodo **N**.

A partir de las ecuaciones (2.4), (2.5) y (2.6), se obtiene la relación entre la tensión y la corriente en el nodo  $\mathbf{X}$  y en  $\mathbf{N}$ , tal como se presenta en (2.7).

$$\begin{bmatrix} \mathbf{V}_n \\ \mathbf{I}_{n,f} \end{bmatrix} = \mathbf{T}(s) \begin{bmatrix} \mathbf{V}_x \\ \mathbf{I}_{x,f} \end{bmatrix} + \mathbf{T}(\mathbf{L}_T - s) \begin{bmatrix} 0 \\ -\mathbf{I}_f \end{bmatrix} \quad (2.7)$$

El método sugiere un modelo de carga general para el nodo cualquiera  $r$  dependiente de la tensión, y los parámetros de la carga, tal como el que se presenta en (2.8).

$$\mathbf{Y}_r = G_r \left| \frac{\mathbf{V}_r}{\mathbf{V}_0} \right|^{np-2} - jB_r \left| \frac{\mathbf{V}_r}{\mathbf{V}_0} \right|^{nq-2} \quad (2.8)$$

$V_r$  es la tensión en el nodo  $\mathbf{R}$ ,  $Y_r$  es la admitancia de carga,  $G_r$  y  $B_r$  son las constantes proporcionales a la conductancia y susceptancia, y  $np$  y  $nq$  son valores que modelan el comportamiento de las componentes activa y reactiva de la carga.

Reemplazando la corriente  $I_{n,f}$  por el negativo del producto entre la tensión y la admitancia, la ecuación (2.7) se transforma en (2.9).

$$\begin{bmatrix} \mathbf{V}_n \\ \mathbf{I}_f \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{T}_{12}(\mathbf{L}_T - s) \\ -\mathbf{Y}_n & \mathbf{T}_{22}(\mathbf{L}_T - s) \end{bmatrix}^{-1} \mathbf{T}(\mathbf{L}_T) \begin{bmatrix} \mathbf{V}_x \\ \mathbf{I}_{x,f} \end{bmatrix} \quad (2.9)$$

Adicionalmente, de (2.4) se obtiene la ecuación (2.10) para la tensión en el nodo de la falta  $\mathbf{F}$ .

$$\mathbf{V}_f = \mathbf{T}_{11}(s) \mathbf{V}_x + \mathbf{T}_{12}(s) \mathbf{I}_{x,f} \quad (2.10)$$

Para la localización de la falta, a partir de los registros de tensión y corriente medidos en la subestación, y utilizando los valores de secuencia, se calcula inicialmente la constante de propagación y la impedancia de sobretensión de la línea. Las constantes de la carga se calculan con los valores de prefalta. A partir de un valor inicial asumido de  $s$  se calculan las componentes de secuencia de la corriente y tensión en el nodo en falta así como la tensión en el nodo  $\mathbf{N}$ , a partir de (2.9) y (2.10). Con el nuevo valor de la tensión en el nodo  $\mathbf{N}$  se calcula la admitancia de carga bajo condiciones de falta. Finalmente, se comprueba que la ecuación que corresponde al tipo de falta, en este caso falta monofásica (2.3), sea satisfecha. Si esto no ocurre se actualiza el valor de  $s$  y se vuelven a calcular todos los valores, pero con el valor actualizado de la admitancia en el nodo  $\mathbf{N}$ . El proceso termina cuando se satisface la ecuación (2.3), con un error menor al valor de tolerancia seleccionado.

Finalmente, el método no considera las cargas que existen entre el nodo de medida  $\mathbf{M}$  y el nodo en falta  $\mathbf{F}$ , por lo que hay que evaluar todos los escenarios posibles y hacer una selección lógica de todos los resultados para obtener la estimación correcta de la distancia de a la falta.

## 2.4.3 Método de Girgis

### 2.4.3.1 Resumen de la técnica

La técnica de localización propuesta por Girgis et al [GIRG93], usa el cambio de la magnitud del fundamental de corriente para determinar el tipo de falta. Conocido el tipo de falta, se calcula la impedancia aparente con la selección adecuada de las tensiones y corrientes utilizadas. La impedancia aparente así calculada sirve para plantear una ecuación compleja, que tiene como variables la resistencia de la falta y la distancia a la cual ésta ocurrió. El método considera las cargas y actualiza las tensiones en cada nodo mediante un flujo de carga simple, con un modelo de carga de impedancia constante. Para el caso de laterales, la técnica está deducida considerando el hecho que en esas líneas sólo pueden ocurrir faltas monofásicas.

### 2.4.3.2 Estimación de distancia de falta

Para el caso de la falta monofásica en el nodo  $\mathbf{F}$  y partiendo de las ecuaciones de tensión de secuencia en la falta presentadas en (2.11), y la relación dada por la conexión de la redes de secuencia que se presenta en (2.12), se obtiene que el valor de la tensión de la fase faltada (fase A), en el nodo  $\mathbf{X}$ , está dada por la ecuación (2.13). Para obtener esta última ecuación se asume que la impedancia de secuencia positiva es igual a la de secuencia negativa, lo cual es válido para líneas de transmisión.

$$V_{(0)f} = V_{(0)x} - I_{(0)x,f} Z_0 \quad V_{(1)f} = V_{(1)x} - I_{(1)x,f} Z_1 \quad V_{(2)f} = V_{(2)x} - I_{(2)x,f} Z_2 \quad (2.11)$$

$$V_{(0)f} + V_{(1)f} + V_{(2)f} = 3I_{(0)f} R_f \quad (2.12)$$

$$V_{(a)x} = (I_{(a)x,f} + k I_{(0)x,f}) Z_1 + 3I_{(0)f} R_f \quad (2.13)$$

En este caso,  $Z_1$ ,  $Z_2$ , y  $Z_0$  son las componentes de secuencia positiva, negativa y cero de la impedancia de la línea, vistas desde el punto de medida hasta la falta **F**. Como no hay cargas intermedias, para la figura 2.2 se asume que se puede determinar la tensión en el nodo **X**. La constante  $k$  está dada por (2.14).

$$k = \frac{Z_0 + Z_1}{Z_1} \quad (2.14)$$

A partir de la tensión en el punto de medida, se puede plantear la ecuación para la distancia a la falta como se presenta en (2.15)

$$Z_{app} = \frac{V_{selec}}{I_{selec}} = s Z_1 + \frac{I_{comp} R_f}{(I_a + k I_0)} \quad (2.15)$$

De la ecuación (2.15), se puede obtener el valor de  $s$  y de  $R_f$ , por la separación en parte real e imaginaria. Los valores de  $V_{selec}$ ,  $I_{selec}$  e  $I_{comp}$  se presentan en la tabla 2.1 para cada tipo de falta

Tipo de falta	$V_{selec}$	$I_{selec}$	$I_{comp}$
Monofásica, fase A	$V_a$	$I_a + k I_0$	$3 I_0$
Monofásica, fase B	$V_b$	$I_b + k I_0$	$3 I_0$
Monofásica, fase C	$V_c$	$I_c + k I_0$	$3 I_0$
Línea - línea AB	$V_a - V_b$	$I_a - I_b$	$\Delta I_a - \Delta I_b$
Línea - línea BC	$V_b - V_c$	$I_b - I_c$	$\Delta I_b - \Delta I_c$
Línea - línea CA	$V_c - V_a$	$I_c - I_a$	$\Delta I_c - \Delta I_a$
Trifásica	Igual que para faltas línea-línea y línea-línea-tierra		

Tabla 2.1. Tensiones y corrientes seleccionadas de un sistema trifásico para el cálculo de la impedancia aparente

El método considera un modelo de carga de impedancia constante, tal como se presenta en (2.16).

$$[Y_L] = \begin{bmatrix} Y_a & 0 & 0 \\ 0 & Y_b & 0 \\ 0 & 0 & Y_b \end{bmatrix} \quad (2.16)$$

Cuando se consideran cargas intermedias entre el punto de medida y la localización de la falta, se deben actualizar los valores de tensión en cada nodo, con el valor en el nodo anterior menos la caída de tensión en la línea. La nueva corriente se obtiene de la sustracción de la corriente por la línea menos la corriente de la carga en el nodo.

## 2.4.4 Método de Zhu

### 2.4.4.1 Resumen de la técnica

La técnica presentada en Zhu et. al. [ZHU97] propone un análisis matricial de la línea para obtener una ecuación en función de la distancia, la resistencia y las corrientes en el nodo en falta **F**. Como se plantea una ecuación compleja con tres incógnitas, ésta se resuelve mediante un proceso iterativo que parte de asumir que la corriente en el nodo en falta está dada por la diferencia entre la corriente de falta y la corriente de prefalta. La impedancia de la línea se representa por una matriz que considera la impedancia mutua. Para tener en cuenta las cargas, éstas se modelan dependientes de la tensión. Los valores de las tensiones en los nodos antes del nodo **F**, se calculan como el resultado de restar la caída en la línea, del valor de la tensión en el nodo anterior. Las corrientes de línea de una sección, se calculan como la corriente de línea de la sección anterior, menos la corriente de la carga en el nodo entre las secciones analizadas.

Debido a inexactitudes en la estimación de los parámetros del sistema y en las mediciones realizadas, se considera que el método da un intervalo de falta, esto es, hace un extensión

probabilística. Adicionalmente, usa una base de datos de faltas donde se almacenan los registros de activación de los elementos de protección, para reducir el problema de la múltiple estimación.

#### 2.4.4.2 Estimación de distancia de falta

El algoritmo de estimación de la distancia, se basa en la solución iterativa de la ecuación que describe la falta en estado estable. A partir de la figura 2.2 y para una falta monofásica en la fase A, la tensión en el nodo **X** está dada por (2.17).

$$V_{(a)x} = s(Z_{aa}I_{(a)x,f} + Z_{ab}I_{(b)x,f} + Z_{ac}I_{(c)x,f}) + I_f R_f \quad (2.17)$$

La corriente en el nodo en falta se calcula como la corriente estimada del nodo **X** al nodo **F** menos la corriente que sale del nodo **F** al nodo **X+1**. De acuerdo con la figura 2.2, y con las condiciones de falla monofásica, la corriente está dada por (2.18).

$$I_f = I_{(a)x,f} - I_{(a)f,x+1} \quad (2.18)$$

Como el sistema de distribución es de topología radial, la corriente del nodo **F** al nodo **X+1** sólo depende de las tensiones en el nodo de falta. Las tensiones en el nodo de falta están dadas por la ecuación (2.19).

$$\begin{bmatrix} V_{(a)f} \\ V_{(b)f} \\ V_{(c)f} \end{bmatrix} = \begin{bmatrix} V_{(a)x} \\ V_{(b)x} \\ V_{(c)x} \end{bmatrix} - s \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} \begin{bmatrix} I_{(a)x,f} \\ I_{(b)x,f} \\ I_{(c)x,f} \end{bmatrix} \quad (2.19)$$

El procedimiento para la estimación de la distancia al sitio de falta es iterativo e inicia considerando un valor de la corriente  $I_f$  igual a la corriente de falta menos la corriente de prefalta. Con este valor se puede resolver la ecuación (2.17), para la distancia y la resistencia de falta. Con el valor de la distancia de falta se puede determinar el valor de la tensión en el nodo **F**. Conociendo la tensión y utilizando un algoritmo de flujo radial se puede estimar la nueva corriente de falta, con la que se inicia otra iteración. El proceso termina cuando la diferencia entre los valores de la distancia al punto de falta entre una iteración y otra, sea inferior a un valor de tolerancia predefinido.

Para considerar el efecto de la carga, ésta se modela con una ecuación de corriente dependiente de la tensión en el nodo de carga **P**, tal como se presenta en (2.20).

$$I_L = I_r \left| \frac{V_p}{V_{p,nominal}} \right|^{np-2} + jI_i \left| \frac{V_p}{V_{p,nominal}} \right|^{nq-2} \quad (2.20)$$

Donde las componentes  $I_r$  e  $I_i$  son las componentes activa y reactiva de la carga a la tensión nominal.

Cuando se consideran cargas intermedias entre el punto de medida y la localización de la falta, se deben actualizar los valores de tensión en cada nodo, con el valor en el nodo anterior menos la caída en la línea. La nueva corriente se obtiene de la corriente por la línea menos la corriente de la carga en el nodo.

#### 2.4.5 Método de Aggarwal

##### 2.4.5.1 Resumen de la técnica

La metodología propuesta en Aggarwal et. al. [AGGA97] se fundamenta en la determinación de los valores de las componentes superimpuestas de tensión y corriente en cualquier punto de la línea, a partir de las medidas en la subestación. Se define como componente superimpuesta a la diferencia entre los valores de las señales de corriente y tensión de falta y de prefalta. La idea principal se fundamenta en que el valor de las componentes superimpuestas en las fases que no están en falta, toman su valor mínimo en el nodo de la falta **F**. Debido a que se parte de este principio, este método no puede estimar la distancia ante una falta trifásica.

El método se basa entonces en hacer un barrido de la sección, asumiendo una distancia de la falta y calculando las componentes superimpuestas de corriente en el punto de falta. A medida que el valor

de la componente superimpuesta disminuye, la distancia a la falta asumida se hace más parecida a la distancia real. El proceso iterativo termina cuando se halla un valor mínimo de la componente superimpuesta. Si se tienen varias secciones de línea, se repite este procedimiento iterativo para cada una de ellas, considerando las cargas que existen en cada nodo. Este método detecta cuando hay falta y no necesita la determinación del tipo de falta como entrada y por el contrario ofrece esta información como salida.

#### 2.4.5.2 Estimación de distancia de falta

A partir de la figura 2.2 que representa la sección bajo falta, se desarrolla el algoritmo de localización con un modelo de parámetros concentrados de la línea, la tensión  $V_f$  en el punto de falta inicialmente asumido a una distancia  $\beta$  desde el nodo  $\mathbf{X}$ , se puede relacionar con la tensión medida en la subestación por la ecuación (2.21).

$$\begin{bmatrix} V_{(a)ff\beta} \\ V_{(b)ff\beta} \\ V_{(c)ff\beta} \end{bmatrix} = \beta \begin{bmatrix} z_s & z_m & z_m \\ z_m & z_s & z_m \\ z_m & z_m & z_s \end{bmatrix} \begin{bmatrix} I_{(a)x,ff\beta} \\ I_{(b)x,ff\beta} \\ I_{(c)x,ff\beta} \end{bmatrix} + \begin{bmatrix} V_{(a)x} \\ V_{(b)x} \\ V_{(c)x} \end{bmatrix} = \beta [Z_{abc}]_{x,x+1} \begin{bmatrix} I_{(a)x,ff\beta} \\ I_{(b)x,ff\beta} \\ I_{(c)x,ff\beta} \end{bmatrix} + \begin{bmatrix} V_{(a)x} \\ V_{(b)x} \\ V_{(c)x} \end{bmatrix} \quad (2.21)$$

$z_s$  y  $z_m$  son la impedancia propia y mutua de la línea por unidad de longitud. La tensión superimpuesta en el punto de falta asumido se puede expresar tal como se presenta en (2.22).

$$[V'_{a,b,c}]_{ff\beta} = [V_{(a,b,c)ff\beta} \text{ falta} - V_{(a,b,c)ff\beta} \text{ prefalta}] \quad (2.22)$$

La tensión de prefalta en  $\beta$  se obtiene con la ecuación (2.23)

$$[V_{a,b,c}]_{ff\beta} \text{ prefalta} = [Z_{\beta}][I_{a,b,c}]_{x,ff\beta} \text{ prefalta} + [V_{a,b,c}]_{x} \text{ prefalta} \quad (2.23)$$

A partir de las ecuaciones anteriores se obtienen los valores de las tensiones y las corrientes superimpuestas en el extremo de la sección (nodo  $\mathbf{X}$ ), tal como se presenta en las ecuaciones (2.24) y (2.25).

$$[V'_{a,b,c}]_x = [V_{(a,b,c)x} \text{ falta} - V_{(a,b,c)x} \text{ prefalta}] \quad (2.24)$$

$$[I'_{a,b,c}]_{x,f} = [V_{(a,b,c)x,f} \text{ falta} - V_{(a,b,c)x,f} \text{ prefalta}] \quad (2.25)$$

Ahora el valor de la componente superimpuesta de corriente entrando al nodo  $\mathbf{X+1}$  está dado por la relación (2.26)

$$\begin{bmatrix} I'_{(a)f,x+1} \\ I'_{(b)f,x+1} \\ I'_{(c)f,x+1} \end{bmatrix} = \left[ (L - \beta) [Z_{abc}]_{x,x+1} + [Z_{abc}]_{x+1,n} \right]^{-1} \begin{bmatrix} V'_{(a)ff\beta} \\ V'_{(b)ff\beta} \\ V'_{(c)ff\beta} \end{bmatrix} \quad (2.26)$$

La matriz  $[Z_{abc}]_{x+1,n}$  representa la impedancia vista desde el nodo  $\mathbf{X+1}$  hasta el nodo  $\mathbf{N}$ . A partir de las ecuaciones (2.25) y (2.26) se puede conformar el modelo de las componentes superimpuestas de corriente en el nodo de falta  $\mathbf{F}$ , que se presenta en la figura 2.4 y es modelado por la ecuación 2.27.

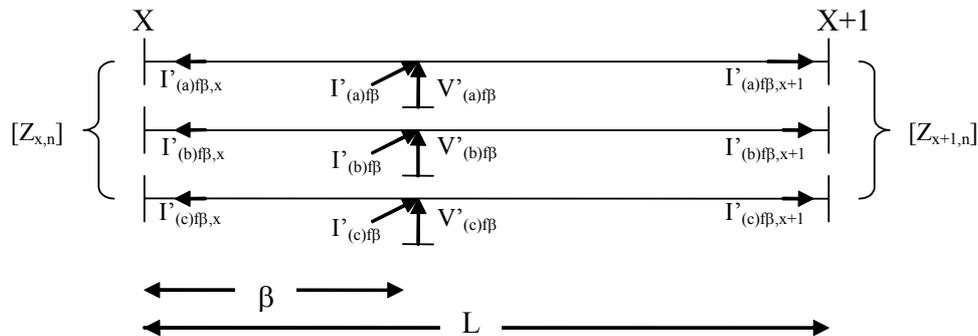


Figura 2.4 Circuito de las componentes superimpuestas

$$[I'_{a,b,c}]_{f\beta} = [I'_{a,b,c}]_{f\beta,x} + [I'_{a,b,c}]_{f\beta,x+1} \quad (2.27)$$

A partir del modelo de la figura 2.4 se obtiene la localización de la falta mediante la evaluación de las componentes superimpuestas según (2.27).

La inclusión de las cargas se hace considerando las impedancias equivalentes de líneas más cargas vistas desde los extremos  $\mathbf{X}$  hasta la subestación y desde  $\mathbf{X}+1$  hasta el nodo  $\mathbf{N}$ , que es el final del alimentador. En cuanto a la representación de las cargas, se considera que hay cargas monofásicas y trifásicas al final del transformador. En general se considera un modelo de admitancia dependiente de la tensión en el nodo, tal como se presenta en (2.28).  $V_L$  es la tensión en el nodo donde se encuentra la carga y  $fp$  es su factor de potencia.  $K$  es una constante que expresa si la carga es trifásica o monofásica (3 y 1 respectivamente)

$$Z_{carga} = K |V_L|^2 \angle \cos^{-1} fp \quad (2.28)$$

Para resolver el problema de la estimación en un alimentador, el proceso iterativo se tiene que realizar para cada punto en cada sección de línea, por lo tanto es complejo e impone altos requerimientos computacionales.

## 2.4.6 Método de Ratan Das

### 2.4.6.1 Resumen de la técnica

Metodológicamente, la técnica propuesta en [DAS98] propone inicialmente, la detección e identificación del tipo de falta. Posteriormente, y a partir de los valores del fundamental de tensión y corriente de prefalta y falta, se hace una estimación de la sección bajo falta entre los nodos  $\mathbf{X}$  y  $\mathbf{X}+1$ . A continuación, el sistema con ramales, se convierte en un sistema radial equivalente, que incluye la sección bajo falta. Finalmente, para cada tipo de falta y de forma iterativa se calculan las tensiones y las corrientes en cada uno de los nodos y en el sitio de falta, en un procedimiento iterativo que se realiza partiendo de los valores de prefalta de la admitancia de carga, para estimar la distancia de falta desde el nodo  $\mathbf{X}$  hasta el sitio de falta  $s$ , dentro de la sección bajo falta. Con el valor de  $s$  se actualizan los valores de tensión y con éstos el valor de las admitancias de carga, iterando hasta que el error en la estimación de  $s$  sea menor a un valor de tolerancia predeterminado.

### 2.4.6.2 Estimación de la distancia de falta

Para la estimación de la distancia a la falta se siguen los siguientes pasos:

#### a. Detección y determinación del tipo de falta

La corriente de línea durante la falta contiene la componente de carga y también la componente de falta. Si una o más de las corrientes de línea son más grandes que una corriente umbral, se asume que ocurrió una falta y si la magnitud de la corriente de secuencia cero es mayor que un valor de umbral se asume que una o dos fases se encuentran en corto circuito con tierra. A partir del análisis anterior, se determina el tipo de falta mediante comparaciones sucesivas de las corrientes medidas en cada fase con su valor de umbral.

#### b. Estimación de la sección faltada – Análisis para el caso monofásico

La estimación de la sección bajo falta se realiza por la comparación de la reactancia estimada con los valores de tensión y de corriente de falta medidos en la subestación (reactancia aparente) y la calculada a partir del análisis de las componentes simétricas y los parámetros del sistema, despreciando los efectos de las cargas (reactancia modificada).

Para el caso de una falta monofásica (fase A a tierra), la reactancia aparente se presenta en (2.29)

$$X_{(1)m,app} = \text{Im} (Z_{(1)m,app}) = \text{Im} \left( \frac{V_{(a)m}}{I_{(a)m,r}} \right) \quad (2.29)$$

Donde:

$Z_{(1)m,app}$  Impedancia aparente de secuencia positiva, desde la subestación (nodo  $\mathbf{M}$ ), hasta la falta.

$X_{(1)m,app}$  Reactancia aparente de secuencia positiva, desde la subestación (nodo  $\mathbf{M}$ ), hasta la falta

La reactancia modificada para la primera sección entre el nodo  $\mathbf{M}$  y el nodo  $\mathbf{R}$  se obtiene con (2.30)

$$X_{m,r}^m = X_{(1)m,r} + \frac{X_{(0)m,r} - X_{(1)m,r}}{3} \quad (2.30)$$

Donde:

- $X_{m,r}^m$  Reactancia modificada entre el nodo **M** y **R**  
 $X_{(0)m,r}$  Reactancia de secuencia cero entre **M** y el **R**  
 $X_{(1)m,r}$  Reactancia de secuencia positiva entre **M** y **R**.

Si la reactancia aparente es mayor que la reactancia modificada, entonces la falta está localizada más allá del nodo **R**. La reactancia de la segunda sección se calcula utilizando la ecuación (2.30), y se adiciona a la primera sección para obtener la reactancia modificada total. Si la nueva reactancia modificada es menor que la reactancia aparente, la falta está situada más allá de las primeras dos secciones. Se continúa el proceso iterativo hasta que la reactancia aparente total sea menor que la reactancia modificada. Así la localización de la falta se obtiene en la sección **X** y **X+1**.

### c. Obtención de un sistema radial equivalente

Establecida la localización de la falta, en la sección comprendida entre los nodos **X** y **X+1**, la carga en los laterales se concentra en la unión con la línea. Las cargas ubicadas luego de la sección bajo falta, se concentran en el nodo final **N**. El sistema modificado se muestra en la figura 2.5.

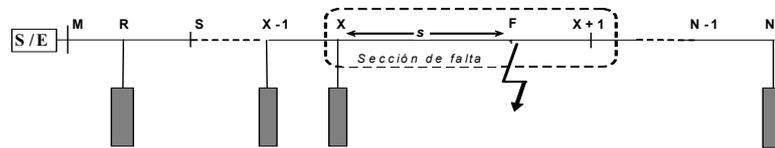


Figura. 2.5. Sistema radial equivalente para una falta en F

### d. Estimación de las cargas

Los efectos de las cargas se tienen en cuenta mediante la compensación de las corrientes aplicando el método de superposición. Se asume que las cargas en un nodo dependen de la tensión en ese nodo tal como se presenta en (2.31). Las constantes de la carga se estiman a partir de la tensión y la corriente de pre-falta.

$$Y_r = G_r |V_r|^{n_p - 2} + j B_r |V_r|^{n_q - 2} \quad (2.31)$$

Donde:

- $V_r$ : Tensión en un nodo cualquiera R  
 $Y_r$ : Admitancia de carga en R  
 $G_r, B_r$ : Constantes proporcionales a la conductancia y susceptancia respectivamente, de la carga en el nodo R (estimadas de los valores de prefalta)  
 $n_p, n_q$ : Constantes para las componentes activa y reactiva de la carga.

Una carga monofásica se modela como una falta línea-tierra, con impedancia de falta igual a la de carga. Una carga bifásica se modela como una falta fase-fase, con la impedancia de falta igual a la de carga. Las cargas trifásicas se asumen como cargas equilibradas y se modelan como faltas trifásicas equilibradas. Para todos los tipos de cargas se usan las redes de secuencia.

Las constantes  $G_r$  y  $B_r$  para cada carga, se calculan con la tensión de prefalta, la admitancia de carga y las constantes de la respuesta de la carga  $n_p$  y  $n_q$ , mostradas a continuación:

$n_p = n_q = 0$  : Modelo de carga a potencia constante.

$n_p = n_q = 1$  : Modelo de carga de corriente constante.

$n_p = n_q = 2$  : Modelo de carga de impedancia constante.

La estimación se realiza con los valores de tensión y de corriente de prefalta en cada nodo.

### e. Estimación de las tensiones y corrientes de prefalta en los nodos

Las tensiones y corrientes en el sistema radial se obtienen usando un modelo de parámetros distribuidos de línea larga, que se simplifica debido a la corta longitud de las líneas de distribución, tal como se presenta en la ecuación (2.32).

$$\begin{bmatrix} V_r \\ I_{r,m} \end{bmatrix} = \begin{bmatrix} 1 & -B_{mr} \\ C_{m,r} & -1 \end{bmatrix} \begin{bmatrix} V_m \\ I_{m,r} \end{bmatrix} \quad (2.32)$$

Las constantes de la línea se presentan en (2.33)

$$B_{m,r} = z_{mr}^c \sinh(\gamma_{m,r} L_{m,r}) \quad C_{mr} = \frac{\sinh(\gamma_{m,r} L_{m,r})}{z_{m,r}^c} \quad (2.33)$$

Donde:

- $V_m, V_r$  Tensiones en los nodos **M** y **R**  
 $I_{m,r}, I_{r,m}$  Corrientes de **M** a **R**, y de **R** a **M** respectivamente.  
 $\gamma_{m,r}$  Constante de propagación por unidad de longitud.  
 $L_{m,r}$  Longitud de la sección.  
 $z_{m,r}^c$  Impedancia característica de la sección.

#### f. Estimación de tensiones y corrientes en el extremo remoto **N** y en la falta **F**

La tensión y la corriente de secuencia en el nodo **X** se calculan antes que en los nodos **F** y **N**. La tensión y la corriente de secuencia en **R** durante la falta se calculan con las medidas tomadas en el nodo **M** y la ecuación (2.32). Las corrientes de secuencia en la carga en el nodo **R**, durante la falta se obtienen a partir de un modelo fase-tierra, fase-fase o trifásico de carga. Las admitancias durante la falta se calculan con (2.31). Las corrientes en el nodo siguiente se obtienen restando la corriente de carga de cada nodo. Este procedimiento se repite hasta el nodo **X**.

Las tensiones y las corrientes en el nodo **X** se calculan con la fase **A** como referencia. Sin embargo, dependiendo del tipo de falta, la tensión y la corriente se deben convertir a la referencia real, como se presenta en la tabla 2.2.

Referencia Fase A	Referencia Fase B	Referencia Fase C
A a tierra	B a tierra	C a tierra
B Y C a tierra	C y A a tierra	A y B a tierra
B a C	C a A	A a B
Trifásica balanceada		

Tabla 2.2. Fase de referencia para los diferentes tipos de faltas

Las tensiones y las corrientes en el nodo **F** se pueden expresar en función de sus valores en el nodo **X**, como se presenta en (2.34).

$$\begin{bmatrix} V_f \\ I_{f,x} \end{bmatrix} = \begin{bmatrix} 1 & -sB_{x,x+1} \\ s \times C_{x,x+1} & -1 \end{bmatrix} \begin{bmatrix} V_x \\ I_{x,f} \end{bmatrix} \quad (2.34)$$

Donde  $s$  es la distancia desde **X** hasta **F**, expresada como fracción de la longitud entre los nodos **X** y **X+1**. Las tensiones y corrientes de secuencia en los nodos **N** y **F** durante la falta están relacionados por la ecuación (2.35).

$$\begin{bmatrix} V_{x+1} \\ -I_{x+1,f} \end{bmatrix} = \begin{bmatrix} De & -Be \\ Ce & -Ae \end{bmatrix} \begin{bmatrix} 1 & -(1-s)B_{x,x+1} \\ -(1-s)C_{x,x+1} & 1 \end{bmatrix} \begin{bmatrix} V_f \\ I_{f,x+1} \end{bmatrix} \quad (2.35)$$

A partir de esta expresión, se obtiene una nueva para las componentes de secuencia de corriente y tensión en la falta **F** y tensión en el nodo **N**, dependientes del parámetro  $s$  y de constantes asociadas a la carga, tal como se presenta en las ecuaciones (2.36), (2.37) y (2.38).

$$V_{(0)f} = V_{(0)x} - sB_{(0)x,x+1}I_{(0)x,f} \quad V_{(1)f} = V_{(1)x} - sB_{(1)x,x+1}I_{(1)x,f} \quad V_{(2)f} = V_{(2)x} - sB_{(2)x,x+1}I_{(2)x,f} \quad (2.36)$$

$$\begin{aligned} I_{(0)f} &= \frac{1}{K'_{ov} + sK'_{0w}} \left\{ (K'_{oq} + sK'_{0r})V_{0x} + (K'_{ov} + sK'_{0u})I_{(0)x,f} \right\} \\ I_{(1)f} &= \frac{1}{K'_{1v} + sK'_{1w}} \left\{ (K'_{1q} + sK'_{1r})V_{1x} + (K'_{1v} + sK'_{1u})I_{(1)x,f} \right\} \\ I_{(2)f} &= \frac{1}{K'_{2v} + sK'_{2w}} \left\{ (K'_{2q} + sK'_{2r})V_{2x} + (K'_{2v} + sK'_{2u})I_{(2)x,f} \right\} \end{aligned} \quad (2.37)$$

$$\begin{aligned}
V_{(0)n} &= \frac{1}{K'_{ov} + sK'_{0w}} \left\{ (K_{om} + sK_{0n})V_{0x} + sK_{0p} I_{(0)x,f} \right\} \\
V_{(1)n} &= \frac{1}{K'_{1v} + sK'_{1w}} \left\{ (K_{1m} + sK_{1n})V_{1x} + sK_{1p} I_{(1)x,f} \right\} \\
V_{(2)n} &= \frac{1}{K'_{2v} + sK'_{2w}} \left\{ (K_{2m} + sK_{2n})V_{2x} + sK_{2p} I_{(2)x,f} \right\}
\end{aligned} \tag{2.38}$$

### g. Estimación de la localización de la falta

En el análisis de faltas monofásicas, la distancia  $s$  desde  $\mathbf{X}$  hasta la falta  $\mathbf{F}$ , como una parte de la longitud de la línea del nodo  $\mathbf{X}$  hasta el nodo  $\mathbf{X}+1$ , se obtiene como se presenta en esta sección.

La ecuación básica para la reactancia de falta, en caso de falta monofásica está dada por (2.39).

$$\text{Im} \left( \frac{V_{(0)f} + V_{(1)f} + V_{(2)f}}{I_{(0)f} + I_{(1)f} + I_{(2)f}} \right) = 0 \tag{2.39}$$

Donde  $V_{(0)f}$ ,  $V_{(1)f}$ ,  $V_{(2)f}$ ,  $I_{(0)f}$ ,  $I_{(1)f}$  e  $I_{(2)f}$  son las tensiones y corrientes de secuencia cero, positiva y negativa en el nodo en falta  $\mathbf{F}$ .

A partir de las ecuaciones (2.36), (2.37) y (2.38) en (2.39), se obtiene la ecuación (2.40) para  $s$ , en términos de constantes que dependen también de la admitancia y de las tensiones y corrientes de secuencia.

$$s = \frac{K''_{AR} \times K''_{CI} - K''_{AI} \times K''_{CR}}{(K''_{CR} \times K''_{BI} - K''_{CI} \times K''_{BR}) + (K''_{DR} \times K''_{AI} - K''_{DI} \times K''_{AR})} \tag{2.40}$$

Como la admitancia depende de la tensión, se debe desarrollar el proceso. El valor de  $s$  se calcula utilizando el siguiente procedimiento:

- i. Calcular las constantes  $K$  que no son dependientes de admitancia.
- ii. Realizar una primera estimación de las constantes  $K'$ , que son dependientes de la admitancia, usando la admitancia de prefallo de la carga consolidada en  $\mathbf{N}$ .
- iii. Calcular las nuevas constantes  $K''$  a partir de las constantes  $K$  y  $K'$  y las tensiones y corrientes de secuencia.
- iv. Calcular una estimación inicial de  $s$ , usando la ecuación (2.40).
- v. Calcular la tensión en  $\mathbf{N}$  usando la ecuación (2.38)
- vi. A partir del valor de tensión obtenido, se calcula el nuevo valor de admitancia para el nodo  $\mathbf{N}$ , usando la ecuación (2.31) para cada secuencia.
- vii. Actualizar las constantes  $K'$  a partir de los nuevos valores de las admitancias en  $\mathbf{N}$ .
- viii. Actualizar los valores de las constantes  $K''$ , con los datos obtenidos en el paso **vii**.
- ix. Calcular el nuevo valor de  $s$  usando la ecuación (2.40).
- x. Si la solución converge, es decir, hay una diferencia entre la estimación anterior y la actual, menor que un margen predefinido, ésta es la distancia del sitio de falta en p.u. desde el nodo  $\mathbf{X}$ . Si no converge, se repite el proceso empezando en el paso **v**.

## 2.4.7 Método de Novosel

### 2.4.7.1 Resumen de la técnica

El método propuesto en Novosel et. al. [NOVO98], se basa en la simplificación de un circuito de distribución con cargas, a un circuito con las cargas concentradas en el nodo final. Esta simplificación se realiza considerando que la impedancia de la carga es mucho más grande que la impedancia de la línea, por lo tanto los errores en la estimación de la distancia no son considerables. A partir de esta simplificación, y con el mismo concepto del método de la componente reactiva se plantea una ecuación para un circuito superimpuesto equivalente que considera las variaciones de la tensión y la corriente en el estado estable de la falta y la prefalta. A

partir de este circuito se plantea una ecuación compleja en términos de la resistencia y la corriente de la falta y la distancia en p.u. de la misma desde el punto de medida.

Como se tiene una ecuación con aparentemente tres variables, este método plantea una solución iterativa asumiendo el valor de una de las variables. También, a partir de un análisis más detallado, se puede encontrar una ecuación que permite resolver una ecuación cuadrática en términos de la distancia a la falta, resolviendo así el problema en forma directa.

### 2.4.7.2 Estimación de distancia de falta

Para la estimación de la distancia a la falta, el circuito general de la figura 2.1 se convierte en un circuito equivalente tal como se presenta en la figura 2.6.

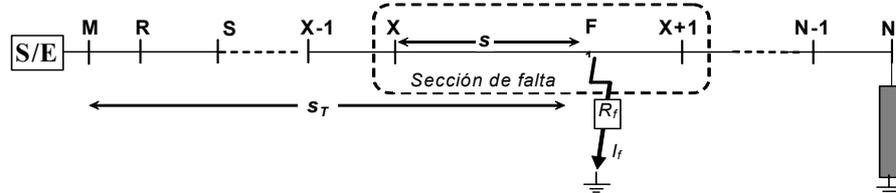


Figura 2.6. Modelo simplificado del sistema de distribución

La aproximación del circuito se realiza considerando que las impedancias de las cargas intermedias son mucho mayores que las impedancias de las líneas y por tanto el error de la aproximación no es significativo.

Para calcular el valor de la distancia de falta  $s_T$  (desde la subestación), se plantea la ecuación para la impedancia aparente medida durante la falta (2.41).

$$Z_{m.app} = \left( \frac{V_m}{I_{m,r}} \right) = s_T Z_{L1} + R_f \left( \frac{I_f}{I_{m,r}} \right) \quad (2.41)$$

Ahora, para resolver el circuito se plantea un nuevo circuito con las componentes superimpuestas que reflejan el cambio de prefalta a falta. El circuito de las componentes superimpuestas se presenta en la figura 2.7.

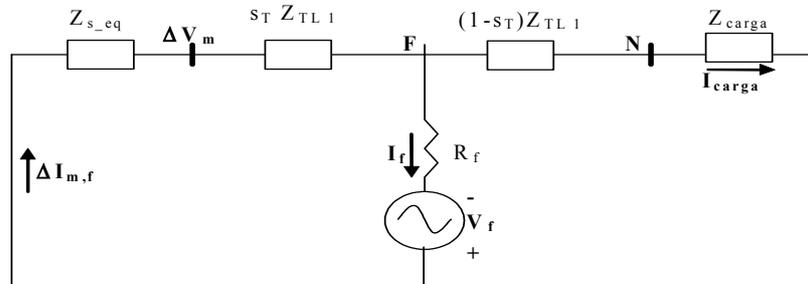


Fig. 2.7. Circuito de las componentes superimpuestas.

La figura 2.7 presenta el sistema superimpuesto con la impedancia de la fuente dada por la ecuación (2.42).

$$Z_{s,eq} = -\frac{\Delta V_m}{\Delta I_{m,f}} = -\frac{V_{m\text{ falla}} - V_{m\text{ prefalla}}}{I_{r,f\text{ falla}} - I_{m,f\text{ prefalla}}} \quad (2.42)$$

En (2.42) se usa la secuencia positiva. En redes no balanceadas se pueden usar los valores de secuencia negativa, si hay gran diferencia entre los valores de prefalta y falta en la secuencia positiva. A partir de los valores de impedancia de la fuente y la carga, se estima la localización de la falta ya sea con el método iterativo o con el directo. La corriente de prefalta de la fuente es igual a la corriente de la carga, tal como se presenta en (2.43).

$$\Delta I_{m,f} = I_{m,f\text{ falla}} - I_{carga} \quad (2.43)$$

De la figura 2.6, los cambios causados en la corriente de la fuente se pueden modelar como se presenta en (2.44), donde  $ds$  es el factor de distribución que describe el cambio de la corriente de la

fuente y la corriente de carga de posfalla.

$$ds = \frac{\Delta I_{m,f}}{I_f} \quad (2.44)$$

Adicionalmente, se define el factor  $ns$  tal como se presenta en (2.45).

$$ns = \frac{I_{m,f \text{ falla}}}{\Delta I_{m,f}} = |ns| \angle \gamma_s \quad (2.45)$$

Reemplazando (2.44) y (2.45) en (2.41) se obtiene (2.46).

$$Z_{med} = s_T Z_{L1} + R_f \left( \frac{1}{ds} \right) \left( \frac{1}{ns} \right) \quad (2.46)$$

Ambos métodos, el iterativo y el directo, se basan en la solución de la ecuación (2.46) que depende de tres incógnitas, la distancia a la falta, la resistencia de la falta y la corriente por la falta incluida en el término  $ds$ . La solución iterativa parte de suponer que el ángulo de  $ns$  es igual a cero. Para la solución directa se propone una ecuación cuadrática en términos de la distancia  $s_T$ .

La carga se tiene en cuenta mediante un modelo de impedancia constante, que se calcula a partir de los valores de prefalta en  $\mathbf{M}$ .

Para considerar los diferentes tipos de faltas, los valores medidos de tensión y de corriente son reemplazados por los que se presentan en la tabla 2.3.

Tipo de falta	Tensión de falta en M ( $V_{m \text{ falta}}$ )	Corriente de falta en M ( $I_{m,f \text{ falta}}$ )	Corriente Superimpuesta $\Delta I_{m,f}$
a-g	Va	Ia+kI <sub>0</sub>	3*ΔI1a
b-g	Vb	Ib+kI <sub>0</sub>	3*ΔI1b
c-g	Vc	Ic+kI <sub>0</sub>	3*ΔI1c
b-c	Vb-Vc	Ib-Ic	ΔIb-ΔIc
a-b	Va-Vb	Ia-Ib	ΔIa-ΔIb
c-a	Vc-Va	Ic-Ia	ΔIc-ΔIa
b-c-g	Vb-Vc	Ib-Ic	ΔIb-ΔIc
a-b-g	Va-Vb	Ia-Ib	ΔIa-ΔIb
c-a-g	Vc-Va	Ic-Ia	ΔIc-ΔIa
a-b-c	Vb-Vc	Ib-Ic	ΔIb-ΔIc
a-b-c-g	Vb-Vc	Ib-Ic	ΔIb-ΔIc

Tabla 2.3. Tensiones y corrientes para diferentes tipos de falta

El valor de  $k$  está dado por la ecuación 2.14.

## 2.4.8 Método de Yang

### 2.4.8.1 Resumen de la técnica

El método propuesto en [YANG98], se plantea para líneas homogéneas sin cargas intermedias, y mediante el análisis del circuito del sistema, se determina la impedancia vista desde la subestación, en función de la distancia  $a$ , la resistencia  $de$ , y la corriente por la falta. Como el método está planteado para determinar la distancia a la falta usando medidas en un solo terminal, pero con una línea alimentada por ambos extremos y un sistema balanceado, la corriente por la falta se relaciona con la corriente medida en uno de los extremos con la corriente de secuencia negativa.

El método así planteado resuelve de manera directa la distancia a la falta, considerando faltas monofásicas y polifásicas. Finalmente, y puesto que no considera cargas intermedias, el método no propone ninguna alternativa para su modelado o agrupamiento.

### 2.4.8.2 Estimación de distancia de falta

Considerando el circuito de la figura 2.2, para una falta monofásica en la fase A entre  $\mathbf{X}$  y  $\mathbf{X}+1$ , la tensión en el nodo  $\mathbf{X}$  de la fase A se presenta en (2.47).

$$V_{(a)x} = sZ_{L1} I_{(a)x,f} + R_f I_f \quad (2.47)$$

Donde  $I_{(a)x,f}$  es la corriente estimada en el nodo **X** para la fase A. La impedancia de falta medida desde el nodo **X** a la falta está dada por (2.48)

$$Z_{x,app} = \frac{V_x}{I_{x,f}} \quad (2.48)$$

De (2.47) y (2.48) se obtiene (2.49)

$$Z_{x,app} = \frac{V_x}{I_{x,f}} = \frac{s \cdot Z_{L1} \cdot I_{x,f} + R_f \cdot I_f}{I_{x,f}} = s \cdot Z_{L1} + \frac{R_f \cdot I_f}{I_{x,f}} \quad (2.49)$$

En los sistemas de transmisión se deben considerar las corrientes inducidas en las fases por efectos de acoplamiento mutuo. Por lo tanto, la corriente en cualquier fase debe compensarse por la corriente que las otras fases inducen en ella, donde para el caso de falta monofásica la corriente de línea se compensa utilizando el factor  $k$  (2.14), tal como se presenta en (2.50).

$$I_{x,f} = I_x + k I_{x0} \quad (2.50)$$

La corriente de línea  $I_{x,f}$ , puede expresarse como la suma de la corriente que aporta la fuente A a la falta ( $I_{Af}$ ), y la corriente de carga ( $I_{Ldx}$ ), como se muestra en (2.51).

$$I_{x,f} = (I_{Af} + I_{Ldx}) + k \cdot I_{x0} \quad (2.51)$$

Como la corriente de falta  $I_f$ , no se puede determinar con medidas directas, debido a la presencia de una fuente equivalente en el nodo **X** y otra en el nodo **X+1**, se recurre a una sustitución por medio de las corrientes de secuencia, que para el caso de falta monofásica se puede representar como (2.52).

$$I_f = 3I_{(1)f} = 3I_{(2)f} = 3I_{(0)f} \quad (2.52)$$

Las corrientes de secuencia en la falta, se pueden expresar de manera proporcional a las corrientes de falta de la línea como se presenta en (2.53).

$$I_{f1} = \frac{I_{x1}}{K_1} \quad I_{f2} = \frac{I_{x2}}{K_2} \quad I_{f0} = \frac{I_{x0}}{K_0} \quad (2.53)$$

Donde los valores de las  $Ks$  representan la contribución de cada fuente a la falta, para cada secuencia. Como se analiza el caso de un circuito de distribución, no existe fuente en el nodo **X+1**, entonces el valor de  $K$  se convierte en una razón de impedancias.

El método asume un sistema balanceado, y por lo tanto la mejor opción es reemplazar la corriente de falta por su equivalente en esta secuencia, por considerar que esa opción es independiente de la corriente de carga, y por tanto (2.49) se reemplaza por (2.54).

$$Z_x = s \cdot Z_{L1} + 3 \frac{I_{x2}}{I_{x,f}} \cdot \frac{R_f}{K_2} \quad (2.54)$$

La razón de corrientes de (2.54) se puede expresar como (2.55). De otra parte, por la separación en la parte real e imaginaria y la división de las dos expresiones se obtiene (2.56), usada para obtener (2.57), para la distancia a la falta.

$$\frac{I_{x2}}{I_{x,f}} = \left| \frac{I_{x2}}{I_{x,f}} \right| e^{-j\alpha_x} \quad (2.55)$$

$$\frac{X_x - sX_{L1}}{R_x - sR_{L1}} = -\tan \alpha_x \quad (2.56)$$

$$s = \frac{X_x + R_x \tan \alpha_x}{X_{L1} + R_{L1} \tan \alpha_x} \quad (2.57)$$

El principio de localización de falta monofásica se aplica para faltas bifásicas o trifásicas, redefiniendo los valores de  $V_x$  e  $I_x$  para múltiples fases, y también que  $\Delta I_x$  reemplaza a la corriente de secuencia negativa en la ecuación para  $\alpha_x$ , como se presenta en la tabla 2.4.

Tipo de falta	$V_x$	$I_{x,f}$	$\Delta I_x$
a-b ó a-b-g	$V_{ab} = V_a - V_b$	$I_{ab}' = I_a - I_b$	$I_{ab}' - (I_{Lda} - I_{Ldb})$
b-c ó b-c-g	$V_{bc} = V_b - V_c$	$I_{bc}' = I_b - I_c$	$I_{bc}' - (I_{Ldb} - I_{Ldc})$
c-a ó c-a-g	$V_{ca} = V_c - V_a$	$I_{ca}' = I_c - I_a$	$I_{ca}' - (I_{Ldc} - I_{Lda})$
a-b-c ó a-b-c-g	Cualquiera de las anteriores		

Tabla 2.4. Sustituciones para faltas en múltiples fases

## 2.4.9 Método de Saha

### 2.4.9.1 Resumen de la técnica

Esta técnica propuesta en Saha et. al. [SAHA02], localiza la distancia de falta mediante la impedancia aparente, calculada a partir de las corrientes y las tensiones en la subestación. Inicialmente se estima la sección donde ocurrió la falta con la impedancia aparente usada para calcular la impedancia equivalente desde la subestación hasta el siguiente nodo (**R**), mediante el arreglo serie paralelo de las impedancias de línea y de carga, asumiendo que la falta en **R** tiene resistencia cero. El proceso se repite hasta cuando el valor de la impedancia equivalente sea menor que cero, caso en el cual se asume que la falta ocurrió justo después del último nodo evaluado. El método propone una fórmula recursiva que sirve para el cálculo de esta impedancia equivalente. Luego de determinar la sección bajo falta, se plantea la estimación de la distancia desde el nodo **X** hasta la falta **F**, mediante el principio planteado por el método de la componente reactiva.

### 2.4.9.2 Estimación de la sección bajo falta

El método presenta la formulación utilizada para el caso donde se tengan medidas de la corriente total de la subestación o para el caso donde se tengan las medidas de la corriente directamente en el alimentador bajo falta. A partir de estos valores de corriente, se determina la ecuación para la impedancia del lazo en falta, vista desde el *i*-ésimo nodo hasta el sitio donde ocurrió la falta, asumiendo resistencia de falta igual a cero. La ecuación recursiva del cálculo de la impedancia del lazo en falta representa los arreglos serie paralelo de las impedancias del sistema. Las impedancias serie están representadas por las de la línea, mientras que las paralelo son las impedancias debidas a las cargas entre la subestación y la falta. En la figura 2.8 se presenta el esquema básico utilizado para el cálculo, mientras que en la ecuación (2.58) se presenta la fórmula recursiva.

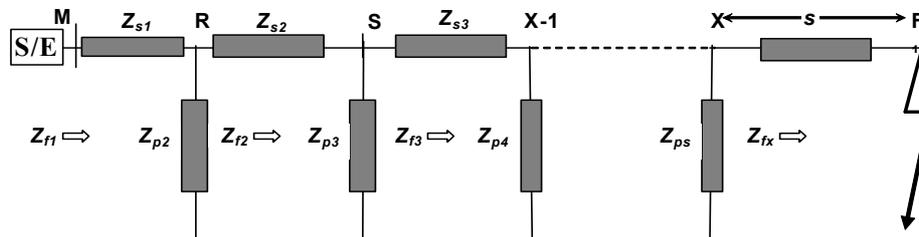


Figura 2.8. Modelo simplificado del sistema de potencia utilizado para hallar la impedancia equivalente vista desde la subestación hasta un nodo cualquiera K

$$Z_{fi} = \frac{Z_{pi} (Z_{fi-1} + Z_{si-1})}{Z_{pi} - Z_{fi-1} + Z_{si-1}} \quad (2.58)$$

En (2.58),  $Z_{si-1}$  representa la impedancia del segmento de línea entre los nodos consecutivos, mientras que  $Z_{pi}$  representa la impedancia de las cargas conectadas al nodo *i*, entre la subestación y el sitio de falta.  $Z_{fi}$  es la impedancia obtenida con los valores de tensión y corriente durante la falta.

La sección bajo falta se obtiene cuando el valor de la impedancia  $Z_{fi}$  calculada con la ecuación (2.58), pasa de un valor positivo en el nodo  $X$  a un valor negativo en el nodo  $X+1$ . La sección bajo falta es entonces la que se encuentra entre los nodos  $X$  y  $X+1$ .

### 2.4.9.3 Estimación de distancia a la falta

Después de encontrar la sección de línea donde se encuentra la falta, se hace una búsqueda detallada de la distancia en p.u. a la falta  $F$ . El modelo de la línea entre el nodo  $X$  y  $X+1$  se muestra en la figura 2.9 y representa la falta con una resistencia  $R_f$ , para el caso de falta fase-fase.

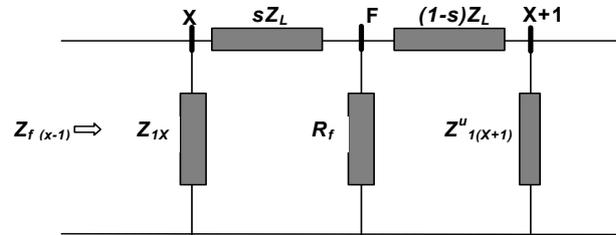


Figura 2.9. Modelo simplificado del sistema de potencia utilizado para hallar la impedancia equivalente vista desde la subestación hasta un nodo cualquiera K

De la figura 2.9,  $Z^u_{1(X+1)}$  representa la impedancia equivalente desde el nodo  $X+1$  hasta el último nodo del alimentador bajo falta,  $R_f$  es la resistencia de falta y  $Z_L$  es la impedancia total de la sección de línea  $X$ ,  $X+1$ .

De la ecuación de la impedancia equivalente del sistema se obtiene la ecuación de la resistencia de falta, para la cual su parte imaginaria se iguala a cero, tal como se presenta en (2.59).

$$\text{Im}(R_f) = s^2 \text{Im}\left(\frac{Z_L^2 (Z_{fi-1} - Z_{1X})}{M}\right) - s \text{Im}(Z_L) - \text{Im}\left(\frac{Z_{fi-1} Z_{1X} (Z_L - Z^u_{1(X+1)})}{M}\right) = 0 \quad (2.59)$$

$$M = (Z_{fi-1} - Z_{1X})(Z_L + Z^u_{1(X+1)}) - Z_{fi-1} Z_{1X}$$

El valor de  $s$  se obtiene del valor positivo de solución de la ecuación cuadrática (2.59). El valor de la distancia de falta desde la subestación, es la suma de la distancia hasta el nodo  $X$  más la correspondiente porción  $s$  de la distancia entre el nodo  $X$  y  $X+1$ .

## 2.4.10 Método de Choi

### 2.4.10.1 Resumen de la técnica

El algoritmo de localización de faltas propuesto en Choi et al. [CHOI04], está basado en el análisis directo del circuito trifásico para el sistema de distribución desbalanceado. Para resolver la complejidad de las ecuaciones matriciales se incorpora el lema de la matriz inversa, como un concepto de gran ayuda en la determinación de la ecuación de la distancia a la falta. Mediante este análisis, se obtiene una ecuación cuadrática que permite obtener la distancia a la falta, considerando que las cargas del circuito están ubicadas en el nodo final.

Como aspecto importante, el documento donde se presenta el método señala la imposibilidad de aplicar la teoría de las componentes simétricas sobre sistemas desbalanceados, debido a la dependencia entre las componentes positiva, negativa y cero. Como aporte adicional de esta tesis, se presenta el análisis para faltas de fase, no desarrollado por el autor a la fecha de publicación.

### 2.4.10.2 Estimación de distancia de falta

Para una falta entre los nodos  $X$  y  $X+1$ , el valor de la tensión en el nodo  $X$  está dado por (2.60).

$$V_{(abc)X} = (1-s) Z_{(abc)L} I_{(abc)X,f} + V_{(abc)f} \quad (2.60)$$

Donde:

$V_{(abc)X} = [V_{(a)X} \ V_{(b)X} \ V_{(c)X}]$  es el vector de tensiones de fase.

$I_{(abc)X,f} = [V_{(a)X,f} \ V_{(b)X,f} \ V_{(c)X,f}]$  es el vector de corrientes de fase entre el nodo  $X$  y el nodo  $F$ .

$V_{(abc)f} = [V_{(a)f} \ V_{(b)f} \ V_{(c)f}]$  es el vector de tensiones de fase en el punto de falta y  $Z_{Labc}$  es la matriz de impedancia de la línea  $\mathbf{X}$ ,  $\mathbf{X}+1$ .

Para el caso de la falta monofásica, la tensión de la fase A se expresa como se muestra en (2.61).

$$V_{(a)x} = (1-s)(Z_{Laa}I_{(a)x,f} + Z_{Lab}I_{(b)x,f} + Z_{Lac}I_{(c)x,f}) + I_f R_f \quad (2.61)$$

Como en el punto de falta  $\mathbf{F}$ , existen dos circuitos en paralelo, uno asociado a la falta y el otro con la carga. Para considerar un análisis matricial, la admitancia de falta está dada por (2.62).

$$Y_f = \begin{bmatrix} \frac{1}{R_f} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (2.62)$$

La admitancia de carga  $Y_L$  se puede obtener tomando el inverso de la impedancia de la conexión serie de la línea entre  $\mathbf{F}$  y  $\mathbf{X}+1$  y la carga en el nodo  $\mathbf{X}+1$ , tal como se presenta en (2.63).

$$Y_L = (sZ_{Labc} + Z_{(X+1)abc})^{-1} \quad (2.63)$$

La matriz de impedancia de carga  $Z_{(X+1)abc}$  se presenta en (2.64).

$$Z_{(X+1)abc} = \begin{bmatrix} Z_{raa} & Z_{rab} & Z_{rac} \\ Z_{rba} & Z_{rbb} & Z_{rbc} \\ Z_{rca} & Z_{rcb} & Z_{rcc} \end{bmatrix} \quad (2.64)$$

Mediante la ley de distribución de corriente de un circuito paralelo se obtiene el valor de la corriente en la falta con (2.65).

$$I_f = Y_f [Y_f + (sZ_{Labc} + Z_{(X+1)abc})^{-1}]^{-1} I_{x,f} \quad (2.65)$$

Para la solución para la corriente de falta, ecuación (2.65) se puede simplificar aplicando el lema de la matriz inversa (2.66), para obtener una nueva ecuación que se presenta en (2.67).

$$(A^{-1} + D * C * B)^{-1} = A - A * B(C^{-1} + D * A * B)^{-1} D * A \quad (2.66)$$

$$I_f = \frac{1}{R_f + s \cdot Z_{Laa} + Z_{raa}} \begin{bmatrix} (s \cdot Z_{Laa} + Z_{raa}) & (s \cdot Z_{Lab} + Z_{rab}) & (s \cdot Z_{Lac} + Z_{rac}) \end{bmatrix} \begin{bmatrix} I_{(a)x,f} \\ I_{(b)x,f} \\ I_{(c)x,f} \end{bmatrix} \quad (2.67)$$

Sustituyendo la ecuación (2.67) en la ecuación (2.61), se obtiene (2.68).

$$V_{sa} = (1-s) \frac{(Z_{Laa}I_{(a)x,f} + Z_{Lab}I_{(b)x,f} + Z_{Lac}I_{(c)x,f}) + R_f [(sZ_{Laa} + Z_{raa})I_{(a)x,f} + (sZ_{Lab} + Z_{rab})I_{(b)x,f} + (sZ_{Lac} + Z_{rac})I_{(a)x,f}]}{R_f + sZ_{Laa} + Z_{raa}} \quad (2.68)$$

La ecuación anterior es una ecuación compleja que tiene como variables la distancia y la resistencia de la falta. Mediante la separación en parte real e imaginaria, se puede resolver el sistema.

## 2.5 Comparación de los MBM para localización faltas

Como variables elegidas para desarrollar la comparación de los métodos se eligieron las siguientes: a) Algoritmo para detectar y determinar el tipo de falta; b) Valores eficaces de tensión y corriente de prefalta, falta y postfalta requeridos para la localización; c) Uso de componentes simétricas, d) Uso de componentes de fase, e) Modelo de la línea utilizado, f) Modelo de la carga utilizado en el planteamiento del método, g) Consideración de la no homogeneidad del circuito, h) Consideración de sistemas desbalanceados, i) Análisis de laterales en el sistema, j) Consideración de la presencia de cargas intermedias entre la falta y el nodo de medida, y finalmente, k) Información adicional utilizada por el método. Esta comparación se presenta en la tabla 2.5.

Aspecto analizado			Método de localización de faltas									
			[WARR68]	[SRIN89]	[GIRG93]	[ZHU97]	[AGGA97]	[DAS98]	[NOVO98]	[YANG89]	[SAHA02]	[CHOI04]
Detección e identificación de la falta			-	-	√	√	√	√	√	-	-	-
Procesamiento de la señal	Valor eficaz de tensión y corriente	Prefalta	-	√	-	√	√	√	√	√	√	-
		Falta	√	√	√	√	√	√	√	√	√	√
		Posfalta	-	-	-	√	-	-	-	-	-	-
	Componentes simétricas		√	√	√	-	-	√	√	√	√	-
	Componentes de fase		-	-	-	√	√	-	-	-	-	√
Sistema de distribución	Modelo de línea		Corta	Larga	Corta	Corta	Corta	Larga	Corta	Corta	Corta	Corta
	Modelo de carga		-	$Z=f(V)$	$Z=cte$	$Z=f(V)$	$Z=f(V)$	$Z=f(V)$	$Z=cte$	-	$Z=cte$	-
	No homogeneidad		-	-	√	√	√	√	-	-	√	√
	Sistemas desbalanceados		-	-	-	√	√	√	-	-	-	√
	Laterales		-	-	√	√	√	√	-	-	√	-
	Cargas intermedias		-	√	√	√	√	√	√	-	√	√
Información adicional			-	-	-	PM, CP	-	SBFI	-	-	-	-

Tabla 2.5. Comparación de los MBM para la localización de faltas.

De la tabla se tiene que PM es modelado probabilístico y CP son las patrones de corriente antes durante y después de la falta, requeridos para estimar la región probable de la falta en [ZHU97]. Adicionalmente, SBFI son los indicadores de falta basados en “software”, ubicados a lo largo del sistema de distribución para detectar el paso de corrientes de falta y así reducir el problema de la múltiple estimación.

En el capítulo 5 de esta tesis, se presentan las pruebas de los métodos aquí presentados, para a partir de ellas, determinar cual es el de mejor desempeño. Éste se seleccionará para el desarrollo del sistema híbrido de localización de fallas.

## 2.6 Conclusiones

En este capítulo se presentaron algunas técnicas de localización de faltas, que usan los valores del fundamental de tensión y de corriente medidos en un solo terminal de la línea, junto con los parámetros de la red de distribución. En este documento, son conocidos como Métodos Basados en el Modelo (MBM).

Existen muchos MBM que permiten la estimación de la distancia eléctrica, desde la subestación hasta el sitio de falta, muy aproximada a la real, siempre que los parámetros de la red sean perfectamente conocidos. Esto muestra la alta dependencia de un buen modelo del sistema de distribución y por tanto la exactitud de los métodos está estrechamente relacionada con el conocimiento de la red.

Adicionalmente, los MBM estiman una distancia eléctrica desde la subestación hasta el sitio de falta. Como los sistemas de distribución son altamente ramificados, la distancia eléctrica estimada, puede coincidir con muchos sitios en los diferentes ramales del sistema. Este problema es conocido como el de múltiple estimación del sitio de falta.

Finalmente, muchas de las referencias bibliográficas recientes, muestran la necesidad de integrar los MBM con otro tipo de métodos que usen información diferente, de tal manera que permita reducir los dos problemas mencionados. Este es un campo nuevo de investigación, motivado especialmente por los requerimientos del nuevo mercado eléctrico.

### 3 Métodos de Clasificación basados en el Conocimiento (MCBC)

#### 3.1 Introducción

A partir del análisis presentado en el capítulo anterior, en este capítulo se presenta el enfoque del problema de localización de faltas desde el ámbito de las técnicas de clasificación, iniciando para ello con la fundamentación asociada a la teoría de conocimiento.

Así, muchas de las actividades humanas consideradas como “inteligentes” se basan en la explotación de la información, los hechos, las experiencias y los conocimientos más o menos específicos de un ámbito particular. En consecuencia, una parte importante de las labores de investigación y desarrollo, consiste en la concepción de formalismos que permiten el desarrollo de “sistemas basados en conocimiento” (SBC), y específicamente el estudio de las distintas maneras de definir y crear sus bases [SANT98].

El proceso de conversión de conocimientos en un formato ordenado y útil se denomina “representación de conocimientos”. Una vez que el conocimiento se representa adecuadamente, se puede utilizar en un sistema inteligente, que con el empleo de herramientas de análisis, tratamiento y manipulación automática pueden inducir o deducir nuevos conocimientos.

La literatura especializada en el análisis de la información y su tratamiento y conversión en conocimiento útil, lo clasifica en dos grandes grupos: empírico y teórico [SANT98].

**a. Conocimientos teóricos:** Modelan el conocimiento que se tiene acerca de un tema a través de una teoría correspondiente con el problema planteado. Son “tratados” que se desarrollan a partir del análisis de los conocimientos básicos y representan una generalización de lo empírico, y habitualmente se representan por estructuras simbólicas como normas de producción, modelos matemáticos, redes semánticas u objetos estructurados, entre otros.

**b. Conocimientos empíricos:** Son conocimientos experimentales, esto es, representan el conjunto de casos prácticos observados sobre un tema (conjunto de ejemplos), pero no tienen una relación plenamente validada con el fenómeno. Son conocimientos puros que no se han tratado, analizado o modificado y representan los resultados de experiencias o los ejemplos de casos prácticos sin transformaciones y que no son fácilmente explicables en una primera instancia.

En general, los conocimientos de los que se dispone para abordar un problema particular son de tipo empírico y teórico, y forman conjuntos que se intersectan o se complementan. En la mayoría de casos, los conocimientos disponibles sobre un problema en particular, no son totalmente correctos ni completos; por esta razón la información disponible debe explotarse de la mejor forma posible para entender mejor los fenómenos que representan [MORA05-d].

Para el caso particular de la localización de faltas, y en concordancia con lo expuesto, la utilización de los conocimientos teóricos asociados al problema se realiza mediante los Métodos Basados en el Modelo (MBM), tales como los que se presentan en el capítulo 2. Éstos utilizan relaciones matemáticas plenamente validadas y reconocidas por la teoría general de circuitos eléctricos, para obtener la impedancia de un circuito a partir de las medidas de tensión y corriente tomadas en la subestación de distribución.

La utilización de los conocimientos empíricos se puede hacer mediante herramientas que establezcan una relación entre las medidas de tensión y corriente y el fenómeno real presentado en un sistema. Para el caso específico de localización de faltas, se utiliza la información adicional al valor eficaz, para establecer una relación con la localización de la falta, mediante la utilización de los Métodos de Clasificación basados en el Conocimiento (MCBC). La información adicional o entradas de los MCBC está compuesta por las características de carga, activación de protecciones, comportamiento de los huecos de tensión, comportamiento del perfil de corriente, entre otros. Esta información se caracteriza por lo que en este documento se denomina como “descriptores”. Las salidas de los MCBC son el tipo y la región donde se presentó la falta en el sistema de distribución. Con el problema planteado de esta forma, se encuentra que una de las mejores aproximaciones para

establecer la relación propuesta se obtiene con los métodos conocidos genéricamente como “clasificadores”. Estos métodos establecen una relación, normalmente no lineal, entre un conjunto de entrada (descriptores), y un conjunto de salida (tipo y región de la falta).

En este capítulo se presentan los conceptos básicos sobre la teoría de clasificación y su aplicación al problema de localización de faltas en sistemas de distribución, como elemento básico de los MCBC. Adicionalmente, se presentan los fundamentos básicos de las dos técnicas seleccionadas para abordar el problema planteado en esta tesis.

### **3.2 Algunas aplicaciones de sistemas basados en el conocimiento a la localización de faltas en redes de distribución**

Antes de entrar en el tema de los métodos de clasificación y su aplicación a la localización de faltas, se presenta en este numeral una síntesis algunos de los trabajos más relevantes donde se presenta la aplicación de las metodologías propias del campo de la inteligencia artificial. Esta presentación tiene como intención, remitir al lector a otras aplicaciones e implementaciones realizadas en el área de localización de faltas.

Un sistema basado en el conocimiento para localizar interrupciones del servicio usando datos comprensibles de las llamadas de los clientes, sistemas automáticos de medida remotos y sistemas de control supervisado y adquisición de datos (SCADAS), se presenta en [LIU02]. Este algoritmo utiliza la capacidad de lectura bajo demanda del sistema de medida, el cual permite al centro de control comunicarse remotamente con los controles del medidor. El algoritmo propuesto tiene dos partes importantes, el procedimiento de escalamiento de la interrupción y el de solicitud de información a los medidores (meter polling). El escalamiento involucra la búsqueda de la región de acuerdo con la información comprensiva de la interrupción. El procedimiento de solicitud de información a los medidores incluye el diseño del esquema de solicitud de información y la confirmación de la localización de la interrupción de acuerdo con los resultados del análisis de la información recibida. El algoritmo fue diseñado para trabajar adecuadamente durante una o varias situaciones de interrupción del servicio.

En [CHIE02] se propone una estructura de red Bayesiana en base al conocimiento experto y a los datos históricos para el diagnóstico de faltas en alimentadores de distribución. Esta red imita la relación causal entre el equipo bajo falta y las observaciones durante las interrupciones del servicio del alimentador, tales como distribución regional de las llamadas de reporte de problema de los clientes, observaciones anormales de los alimentadores expresadas en las llamadas, y los eventos climáticos en la región en la que se encuentra la línea. El sistema desarrollado fue usado para simular varios escenarios de falta del alimentador y para entrenar equipos de mantenimiento e inspección de alimentadores. El cuestionario desarrollado puede ser rápidamente adaptado para asegurar las fortalezas de las reglas de inferencia en otras redes bayesianas.

En [HUAN02], se presenta una propuesta de optimización basado en algoritmos inmunes (IA based), para la estimación de la sección bajo falta en un sistema de distribución. A través de la teoría de evolución inmunológica, un anticuerpo que más se asemeje al antígeno correspondiente se convierte en la solución. Para la aplicación de este método al problema de estimación, cada sección del modelo del sistema de potencia puede ser considerada como un “anticuerpo”.

En [MAHA04], la aplicación de las redes neuronales de función de base radial (RBF), se utiliza para la clasificación y localización de faltas en líneas de transmisión. Las muestras instantáneas de tensión y corriente, se utilizan como entradas de una red neuronal artificial. Para la clasificación de la falta, únicamente las muestras de falta y prefalta de la corriente son suficientes, para la localización de requieren las muestras de prefalta y falta de ambos, tensión y corriente de las tres fases. Para validar la propuesta, se realizaron estudios de simulación en dos modelos de sistema de potencia, uno de ellos es una línea de transmisión alimentada por un solo terminal, mientras que el otro es una línea alimentada por ambos terminales. El modelo fue sometido a diferentes tipos de faltas a condiciones de operación diferentes para las variaciones en la localización de la falta, ángulo de inserción de la misma y resistencia de falta.

Los equipos de mantenimiento usualmente identifican las faltas utilizando las llamadas de los clientes quienes reportan problemas, observaciones anormales en los alimentadores que han sido

reportados en las llamadas. Usualmente cada falta es registrada en una tabla que incluye la hora, la fecha, la dirección, el equipo faltado, las causas o accidentes, entre los más importantes. Estas bases de datos acumulan una gran cantidad de información durante años. El estudio propuesto en [PENG04] está orientado al uso de la “rough set theory” (herramienta matemática para manipular conjuntos de información vaga o incompleta), como una herramienta de minería de datos para obtener patrones y reglas útiles para el diagnóstico del equipo en falta y su localización. En particular, los datos históricos de alimentadores de distribución de sistemas reales de Taiwán Power Company fueron usados para la validación de la metodología propuesta.

En [JARV94], se presenta una propuesta para el diagnóstico de las faltas en redes de distribución, mediante el uso de conocimiento heurístico de los operadores del centro de control y de la información obtenida de las bases de datos de la red y los sistemas SCADA. Sin embargo, la naturaleza de este conocimiento heurístico es inexacta e incierta. También, la información obtenida del sistema de control remoto contiene incertezas y puede ser incorrecto, conflictivo o inadecuado. Este artículo propone un método basado en la teoría de conjuntos difusos para tratar la incerteza involucrada en el proceso de localización de faltas en las redes de distribución.

En [HSU90] se propone un sistema experto basado en reglas, para localizar las faltas en un sistema de distribución. Un conjunto de reglas heurísticas son compiladas a partir de la experiencia de los operadores y se encuentran embebidas en la base de reglas. Para localizar faltas en sistemas de distribución, un mecanismo de inferencia se desarrollado para que desempeñe razonamientos deductivos en las reglas de la base de conocimiento. El mecanismo de inferencia comprende tres partes principales: Un método dinámico de búsqueda, aproximación de “backtracking”(algoritmo usado por lenguajes de programación lógica tales como Prolog para encontrar los posibles caminos para alcanzar una meta ) y la operación de intersección de conjuntos.

Un algoritmo para basado en las señales de corriente medidas en la subestación de distribución, el conocimiento de los elementos de protección y las redes ANFIS se presenta en [MORA06-b]. Como resultado se muestra que este método permite localizar la zona bajo falta pro el uso de patrones de corriente mediante la elaboración de reglas “if –then” y el uso del modulo neuronal. En esta aplicación se utilizan las características asociadas a las variaciones de l señal de corriente, como resultado de la operación de los dispositivos de protección del circuito tales como reconectores, seccionalizadores y fusibles. Esto es lo que se denomina como “firma” del sistema” y que se presentará en el siguiente capítulo.

Cada uno de estos métodos tiene sus ventajas y desventajas. Entre las más importantes están el requerimiento de gran cantidad y diversidad de fuentes de información tal como llamadas de clientes, mediciones del sistema SCADA y condiciones climáticas entre otros, como se plantea en [LIU02] [CHIE02] [HUAN02] [PENG04]. Otros sistemas plantean el uso del conocimiento experto de los operadores para desarrollar sistemas basados en reglas [JARV94] [HSU90]. Un método muy interesante, que utiliza solo las medidas de tensión y de corriente, y que muestra su efectividad para localizar faltas en redes de transmisión es el que se presenta en [MAHA04]. Un método aplicado a las redes de distribución se presenta en [MORA06-b], el cual a pesar de usar solo información de los patrones de corrientes registrados bajo condiciones de falta, está limitado a sistemas pequeños y la determinación de una zona en falta, que debe ser muy grande para satisfacer las condiciones de precisión del método.

En la tabla 3.1 se presentan algunas de las principales características de los métodos analizados

Método	Aplicación	Información requerida	Técnica utilizada	Información de salida
[LIU02]	Distribución	Llamadas de clientes, medidas del sistema e información remita del sistema SCADA.	Sistemas expertos	Región o equipo en falta
[CHIE02]	Distribución	Datos históricos de faltas, llamadas de clientes, conocimiento de los expertos	Redes bayesianas	Equipo en falta
[HUAN02]	Distribución	Activación de elementos de protección, estado de interruptores de la subestación	Algoritmos inmunes	Alimentadores, equipos en falta
[MAHA04]	Transmisión	Medidas de tensión y corriente en la subestación	Redes neuronales RBF	Tipo de falta y distancia de falta

Tabla 3.1.a. Comparación de los métodos analizados basados en el conocimiento para localización de faltas.

Método	Aplicación	Información requerida	Técnica utilizada	Información de salida
[PENG04]	Distribución	Parámetros del circuito, Medidas de V e I, Condiciones climáticas, posible causa de la falta.	“Rough set theory”	Equipo en falta
[JARV94]	Distribución	Parámetros y topología del sistema, Medidas de tensión y corriente, Información de SCADA (Corriente da falta, activación de detectores de falta, condiciones climáticas), Información de activación de protecciones.	Teoría de conjuntos difusos	Distancia de falta
[HSU90]	Distribución	Parámetros y topología de la red, histórico de faltas y llamadas de los clientes.	Sistema basado en reglas	Región en falta
[MORA06-b]	Distribución	Medidas de corriente.	Sistema basado en reglas	Región en falta

Tabla 3.1.b. Comparación de los métodos analizados basados en el conocimiento para localización de faltas.

### 3.3 Definición de clasificación

La predicción es una de las aplicaciones más importantes del procesamiento de la información contenida en las bases de datos. Su objetivo es predecir el valor de una variable  $y$ , que es la salida o respuesta del sistema a partir de valores conocidos de las variables de entrada  $\mathbf{x} = \{x_1, \dots, x_n\}$ , que se conocerán como conjunto de descriptores. Si la salida del sistema es un valor numérico,  $y \in \mathbb{R}$ , la predicción es conocida con el nombre de estimación. De otra parte, si la variable de salida toma su valor de un conjunto finito  $\mathbf{C} = \{C_1, \dots, C_m\}$  se tiene un problema de clasificación.

Los elementos del conjunto de la variable de salida de un clasificador, toman sus valores con el nombre de clases. Estas clases son etiquetadas con cualidades que representan situaciones o conceptos significativos del sistema analizado. Las variables de entrada  $\{x_1, \dots, x_n\}$ , describen una situación del sistema u objeto  $\mathbf{x}$  al que se le asigna una clase. Así cada objeto está descrito por un número finito de atributos o cada atributo  $x_i$  toma su valor de un conjunto de  $\mathbf{D}_i$  (espacio de descripción). Los atributos podrán ser cualitativos o cuantitativos y  $\mathbf{D}_i$  estará en un intervalo abierto de  $\mathbb{R}$ . El conjunto cartesiano del espacio de descripción  $\Omega = \mathbf{D}_1 \times \mathbf{D}_2 \times \dots \times \mathbf{D}_n$ , representa el contexto en el cual se puede describir un objeto.

Considerando lo anterior, la definición formal de un clasificador se presenta como: Sea un conjunto de objetos  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , y un conjunto de clases  $\mathbf{C} = \{C_1, \dots, C_m\}$ . Si se le asigna a cada objeto  $\mathbf{x}$  una clase  $C_j$ , el resultado será una partición  $\mathbf{P} = \{P_{C_1}, \dots, P_{C_m}\}$ , del universo de descripción, donde el objeto está asignado a una clase. Así entonces, un clasificador es una regla de decisión  $R: \Omega \rightarrow \mathbf{P}$  ó  $R(\mathbf{x}_i) = P_{C_j}$ , si el objeto  $\mathbf{x}_i$  está incluido en la clase  $C_j$ .

La clasificación comprende típicamente dos etapas, una de aprendizaje (entrenamiento y validación cruzada) y otra de aplicación. Primero el clasificador aprende la partición en clases del universo de descripción del sistema, a partir de un conjunto de datos de entrenamiento. A continuación éste se valida a partir de otro conjunto de datos conocido como conjunto de validación. Finalmente y luego de comprobar su validez, el clasificador se utiliza para predecir la clase a la cual pertenece un nuevo objeto.

El aprendizaje consiste en la estimación los parámetros o las relaciones entre los valores de los atributos, que permiten que el clasificador se adapte a un problema específico. Dentro de las formas como un clasificador aprende, podemos encontrar dos paradigmas principales: Aprendizaje supervisado y no supervisado.

En el aprendizaje supervisado, la base de datos está formada por un conjunto de ejemplos de aprendizaje  $\mathbf{T} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$ . En este tipo de aprendizaje se conoce plenamente a que clase corresponden los objetos usados en el conjunto de entrenamiento.

De otra parte si existe una base de datos de la cual solamente se conoce el valor de los descriptores de cada objeto, entonces el conjunto de entrenamiento será  $\mathbf{T} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ . En este caso se utiliza una estrategia de aprendizaje no supervisado para establecer una relación entre las entradas y las salidas. Las clases que se originan en este tipo de aprendizaje se tienen que relacionar con situaciones significativas del problema analizado.

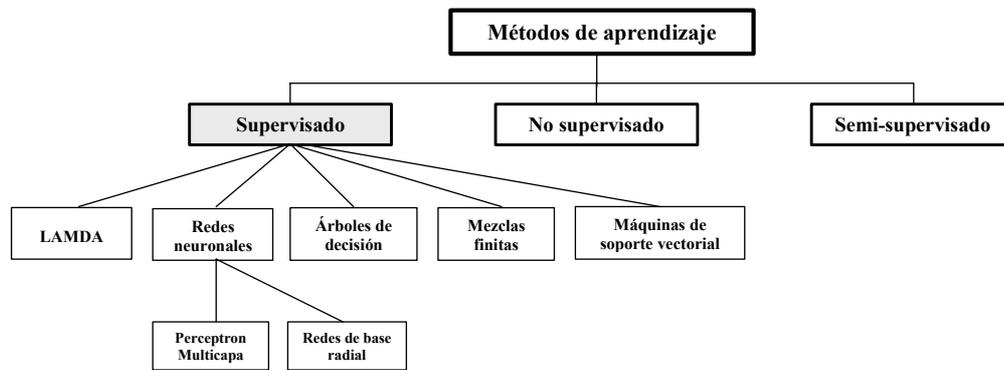
### 3.4 Técnicas de clasificación

La descripción de las técnicas de clasificación usualmente se puede asociar a los métodos de aprendizaje utilizados, tres de los cuales se usan para el ajuste de los parámetros de un clasificador. Estos métodos son el aprendizaje supervisado, el no supervisado o agrupamiento (también conocido como aprendizaje sin instructor) y el agrupamiento semi-supervisado [EVER93].

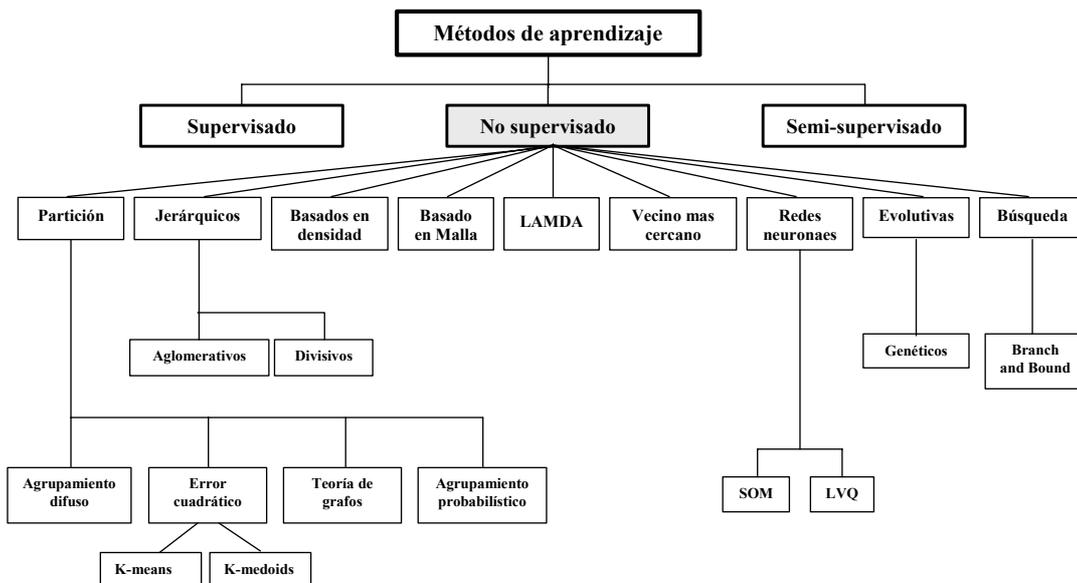
En la aplicación del aprendizaje supervisado a la detección de fallos en sistemas de energía eléctrica, se requiere de bases de datos que contengan las señales de tensión y corriente medidas, asociadas a la localización de la falta, para una cantidad de eventos considerable. Cuando no se pueda obtener una base de datos con las características mencionadas, existe la opción de utilizar metodologías de clasificación no supervisadas. Éstas exploran las interrelaciones existentes entre un conjunto de datos para interpretar su estructura y así resolver el problema.

Una de las alternativas mixta es el aprendizaje semi-supervisado, el cual se vale de alguna información (por ejemplo, datos ya clasificados). En lugar de utilizar esta información en la parte final de la clasificación con el fin de validar los datos, se puede realizar un tipo de seguimiento guiado o ajuste al proceso de agrupamiento. Esta metodología semi-supervisada, a diferencia de las metodologías no supervisadas, es muy reciente y son pocos los métodos publicados hasta ahora.

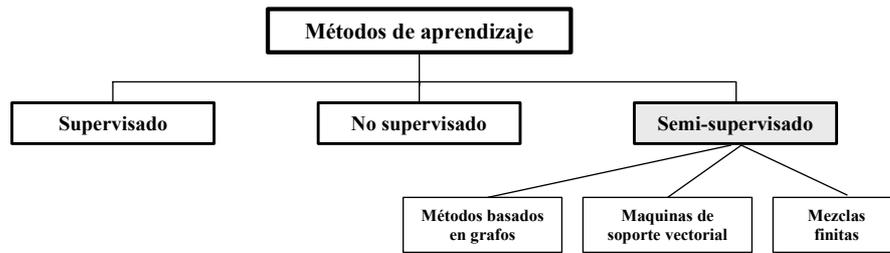
En la figura 3.1 se presenta un resumen de los principales métodos de aprendizaje y los algoritmos considerados como representativos y más comúnmente referenciados de cada categoría.



a)



b)



c)

Figura 3.1. Principales técnicas de clasificación, según su estructura típica de aprendizaje: a) Supervisados, b) no supervisados y c) semisupervisados

En el anexo B, se presenta una referencia de algunos de los algoritmos de clasificación más usados y referenciados en las bases bibliográficas.

### 3.5 Criterios de evaluación del método de clasificación

Una de las preguntas más frecuentes es cómo elegir una técnica de clasificación específica teniendo en cuenta el gran número de métodos reportados en la literatura especializada. El criterio más natural para evaluar un algoritmo de predicción está asociado a la determinación de la diferencia entre el valor estimado y el valor real de salida del sistema, o error de predicción. No obstante el error de predicción para la clasificación no siempre es fácil de establecer.

Para determinar la diferencia de la estimación es posible definir varias funciones de distancia entre el valor estimado y el valor real, pero en la clasificación, no es posible determinar una relación de orden para los conjuntos de posibles valores de salida. Así, existe una sola función de distancia general entre dos valores de salida: si los dos valores son iguales ( $C_j = C_r$ ), o si son diferentes ( $C_j \neq C_r$ ); esta distancia se conoce comúnmente como distancia 0/1 [WAIS00].

A menudo, en caso de aprendizaje supervisado, las técnicas simples de estimación del error son reemplazadas por otras más sofisticadas. La medida del error de clasificación, normalmente se utiliza en los algoritmos de aprendizaje, usan normalmente una distancia distinta a la 0/1, como por ejemplo la distancia euclidiana. Sin embargo, en [FRIE97] se muestra analíticamente la razón por la cual a pesar de la debilidad de la estimación de los resultados obtenidos por métodos simples como el del “vecino más cercano”, con éste se tienen resultados similares que los obtenidos por técnicas de estimación de distancia más sofisticadas.

De otra parte, diferentes aplicaciones prácticas ponen de manifiesto que para la base de datos real, los algoritmos más simples tienen a menudo resultados equivalentes o mejores que otras técnicas más sofisticadas. [LIM00] compara 33 algoritmos de clasificación, en los prototipos de investigación y de programas comerciales. En sus conclusiones, la discriminación lineal se encuentra entre las técnicas que presenta mejores resultados globales en términos de error de predicción.

De acuerdo con lo anterior, el interés no siempre debe centrarse en desarrollar técnicas más sofisticadas de clasificación, sino por el contrario en determinar las características propias de un método que pueden ser más aptas que las de otros para ciertas aplicaciones. En primer lugar las características inherentes a la base de datos utilizada delimitan los algoritmos a aplicar. Por otra parte el límite máximo del error de predicción considerado como aceptable es otro factor para la elección de la técnica de clasificación.

Algunos de los factores que pueden determinar la utilización de un método de clasificación son:

**a. Tipo de atributos de la base de datos:** Una base de datos puede estar conformada por atributos únicamente cualitativos o únicamente cuantitativos, o por una combinación de atributos cuantitativos y cualitativos. Un gran número de técnicas de clasificación utilizan un único tipo de atributos, por ejemplo las redes neuronales trabajan únicamente con atributos numéricos.

**b. Interpretabilidad:** Para algunas aplicaciones, el interés de la clasificación es el descubrimiento de la relación entre diferentes variables o las causas de ciertos comportamientos.

Asimismo, para aplicaciones delicadas (tales como en el ámbito médico), las reglas de decisión generadas deben ser comprensibles, de tal manera que los expertos las puedan validar. La interpretabilidad de las reglas de decisión, es en muchos ámbitos, una condición indispensable en la elección de la técnica de clasificación.

**c. Velocidad de aprendizaje y/o de reconocimiento:** Ciertos métodos de clasificación utilizan algoritmos para el aprendizaje que toman un tiempo de cálculo considerable. Así pues la complejidad del algoritmo o el volumen del cálculo determinan la aplicabilidad de un método específico. Para el caso de aplicaciones que están en un proceso funcionando en tiempo real, la velocidad de reconocimiento es fundamental.

**d. Aprendizaje secuencial:** La mayoría de los algoritmos de clasificación necesitan de todos los datos del conjunto de aprendizaje al mismo tiempo. El entrenamiento de tales algoritmos no puede aplicarse sobre bases de datos que no contengan toda la información del proceso. Debido al desarrollo de los medios de adquisición, las bases de datos son cada vez más completas, y por tanto, requieren de algoritmos que puedan tratar la información de manera secuencial.

**e. Grado de aceptación:** Por definición, el resultado de la clasificación es siempre una partición estricta del universo de descripción. Sin embargo, en ciertos métodos, pueden determinar no solamente la clase a la cual el objeto pertenece, sino también el grado de aceptación del objeto a esta clase y a otras clases existentes. Esto da una información adicional que puede ser importante para el usuario cuando el sistema analizado presente incertidumbres considerables, o cuando el sistema debido a su naturaleza, evolucione de una clase a otra de manera gradual.

Para el caso de las aplicaciones presentadas en esta tesis, la precisión del método está directamente asociada a la cantidad de datos que fueron bien clasificados, respecto al número total de datos.

A partir de la discusión presentada, y considerando la naturaleza del problema, se propone utilizar las Máquinas de Soporte Vectorial (SVM), en la localización de faltas en sistemas de distribución de energía eléctrica, debido que es una técnica que combina muy bien la teoría de la generalización del aprendizaje, con la optimización en la selección de los elementos de conformación de clases. Adicionalmente, mediante la definición adecuada de los “kernels”, esta técnica ha mostrado que resuelve muy bien los problemas no lineales.

Una de las primeras aplicaciones de las SVM en el campo del diagnóstico de faltas es presentada en [WEIW03]. En [GE04] se presenta la aplicación de las SVM para el diagnóstico de faltas en procesos de estampado en hojas metálicas. Según los autores se probaron dos caso diferentes, en los cuales obtuvieron un éxito superior al 95.6 % cuando los mejores resultados obtenidos con las redes neuronales habían sido del 93.3%. Adicionalmente se reporta una disminución considerable en el conjunto de entrenamiento, respecto a la aplicación con redes neuronales. Adicionalmente, dos interesantes aplicaciones de las SVM en el campo del diagnóstico de faltas en transformadores eléctricos, se presentan en [DON04] y [GANY05]. En estas referencias, se presenta por primera vez el uso de esta técnica en el diagnóstico de transformadores, para clasificar características obtenidas a partir del análisis de los gases disueltos en el aceite con el fin de identificar cuatro tipos de faltas. En [XU05] las SVM fueron aplicadas satisfactoriamente en el diagnóstico de faltas en un motor de combustión. En esta aplicación, las SVM se utilizan en combinación de la teoría de conjuntos con incertezas (Rough set theory). Este problema no es de fácil solución debido a la compleja estructura del motor de combustión. Finalmente en [SHUI04] se reporta la aplicación de las SVM en un problema similar, en el cual éstas son aplicadas para el diagnóstico de faltas mediante el análisis de las señales de vibración de un sistema electromecánico de puertas elevadoras.

De igual manera y como metodología de contraste, se propone el algoritmo de clasificación LAMDA, el cual se fundamente en el modelado estadístico del problema. En esta estrategia, se conforman grupos o clases prototipo, con las cuales se compara cada uno de los objetos a clasificar.

Esta metodología tiene la ventaja de que admite como entradas a elementos cualitativos y cuantitativos.

LAMDA ha sido utilizado como herramienta de diagnóstico en muchas aplicaciones. En [AGUA98] esta técnica se utilizó en el campo del diagnóstico en procesos industriales. En [WAIS00] se presenta una aplicación a la supervisión de procesos para una estación de tratamiento de aguas residuales. En [MORA04-a], se presenta una aplicación de la técnica a la predicción de fenómenos de inestabilidad en horno de fundición de acero. Más recientemente, en [KEMP06] se presenta esta técnica como una alternativa a la supervisión de procesos mediante la elaboración de máquinas de estado basado en LAMDA.

Como elementos de contraste, que no se presentan en esta tesis, también se utilizaron las redes ANFIS [MORA06-b] y las mezclas finitas [MORA06-h][CORM06]. Estas aplicaciones se presentan a nivel de referencias que por limitación de espacio físico en el texto no se incluyeron.

### 3.6 Máquinas de Soporte Vectorial

#### 3.6.1. Generalidades

Uno de los MCBC que utiliza características significativas o patrones para determinar la solución del problema de localización de faltas son las Máquinas de Soporte Vectorial (SVM). Éstas fueron desarrolladas por Vladimir Vapnik usando herramientas estadísticas, de optimización y de las máquinas de aprendizaje [VAPN00].

Las SVM son una consecuencia práctica de la teoría de aprendizaje y su estudio es útil por características como la ausencia de mínimos locales, la alta capacidad de generalización y el control y uso de funciones de núcleo o *kernel*. Las SVM, a diferencia de otros métodos de clasificación, tienen la ventaja de no requerir ningún tipo de hipótesis sobre la densidad de probabilidad de los rasgos, y de ser convenientes en términos de la dimensionalidad del problema, entre otros.

La arquitectura de las SVM sólo depende de un parámetro de penalización denotado como  $C$  y la función *kernel* (incluyendo sus parámetros). En el caso de la Función de Base Radial (RBF), existe sólo un parámetro a configurar denotado como  $\sigma$  [BURG98]. De esta manera se evitan requerimientos sobre parámetros exclusivos de arquitectura, tales como número de nodos, capas y tipo de conexión entre capas.

En esta sección se presenta una breve introducción a la teoría de las SVM, se describen sus diferentes parámetros y se presentan sus variantes para el caso de multclasificación.

#### 3.6.2. Metodología de SVM

Considérense  $n$  datos de entrenamiento  $N$  dimensionales  $(\bar{x}_i)$ , con su respectiva etiqueta  $(y_i)$ , tal como se presenta en (3.1)

$$x_i \in \mathfrak{R}^N \quad y \in \{+1, -1\} \quad (3.1)$$

Se busca estimar una función  $f$  tal que para una entrada en  $\mathfrak{R}^N$  produzca una salida en  $\{\pm 1\}$ , según se presenta en (3.2)

$$f: \quad \mathfrak{R}^N \rightarrow \{+1, -1\} \quad (3.2)$$

Así se puede clasificar correctamente un nuevo dato, considerando que  $y = f(\bar{x})$  para este nuevo dato es generado con la misma distribución de probabilidad  $P(\bar{x}_i, y_i)$ , de los datos de entrenamiento.

Como no se imponen restricciones en la función que se escoge, se pueden cometer errores en la estimación, ya que aunque se realice un buen entrenamiento, no necesariamente tiene una buena generalización para datos desconocidos. Por tanto, el aprendizaje perfecto no es posible y la minimización del error de entrenamiento no implica que haya error en la prueba.

### 3.6.2.1. Análisis del caso linealmente separable

#### a. Hiperplano clasificador óptimo

Los clasificadores con SVM se fundamentan en la obtención de hiperplanos que separen los datos de entrenamiento en dos subgrupos. Entre cada una de las clases etiquetadas como  $\{-1, +1\}$ , existe un único hiperplano óptimo de separación (OSH). Se busca que la distancia entre el hiperplano óptimo y el patrón de entrenamiento más cercano sea máxima, con la intención de forzar la generalización de la máquina de aprendizaje [BURG98].

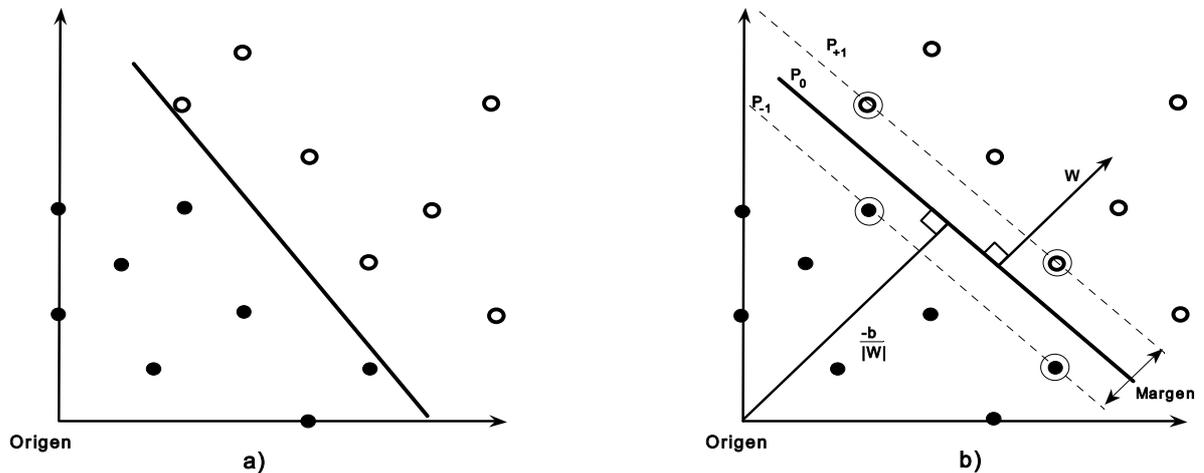


Figura 3.2. Hiperplanos que separan correctamente un conjunto de datos. a) Hiperplano de separación de datos. b) OSH con un mayor margen de separación entre clases.

El hiperplano óptimo de separación (OSH) de la figura 3.2, se presenta en (3.3)<sup>1</sup>:

$$g(\vec{x}) = (\vec{w} \cdot \vec{x}) + b = 0 \quad (3.3)$$

Para maximizar el margen, se proponen dos planos paralelos que contienen los puntos más cercanos al OSH. Si  $1/|w|$  es la distancia entre el punto más cercano al OSH, las ecuaciones de los planos están dadas por (3.4) y (3.5)<sup>2</sup>:

$$p_{+1} = (\vec{w} \cdot \vec{x}) + b = +1 \quad (3.4)$$

$$p_{-1} = (\vec{w} \cdot \vec{x}) + b = -1 \quad (3.5)$$

El margen definido como la distancia perpendicular entre (3.4) y (3.5) está dado por (3.6)

$$\begin{aligned} [(\vec{w} \cdot \vec{x}_{+1}) + b] - [(\vec{w} \cdot \vec{x}_{-1}) + b] &= +1 - (-1) \\ \vec{w} \cdot (\vec{x}_{+1} - \vec{x}_{-1}) &= 2 \\ \frac{\vec{w}}{\|\vec{w}\|} \cdot (\vec{x}_{+1} - \vec{x}_{-1}) &= \frac{2}{\|\vec{w}\|} \end{aligned} \quad (3.6)$$

La distancia entre el OSH y el origen de coordenadas está dada por (3.7)

$$\text{Si } (\vec{w} \cdot \vec{x}_i) + b = 0 \text{ y } \vec{x}_i = 0 \Rightarrow \frac{(\vec{w} \cdot \vec{x}_i) + b}{\|\vec{w}\|} = \frac{b}{\|\vec{w}\|} \quad (3.7)$$

Por definición, entre los dos hiperplanos no deben existir datos de entrenamiento y por tanto los datos deben cumplir con (3.8) y (3.9).

<sup>1</sup> Un plano  $N$  dimensional de la forma  $a_n x_n + a_{n-1} x_{n-1} + \dots + a_1 x_1 + a_0 = 0$ , se puede reescribir de la forma  $(\vec{w} \cdot \vec{x}) + b = 0$ .

$\vec{w} = [a_n, a_{n-1}, \dots, a_1]$  es un vector perpendicular al hiperplano,  $\vec{x} = [x_n, x_{n-1}, \dots, x_1]$ ,  $b = a_0$

<sup>2</sup> La distancia perpendicular que hay entre cualquier punto a un plano es  $\frac{(\vec{w} \cdot \vec{x}) + b}{\|\vec{w}\|}$

$$(\bar{w} \cdot \bar{x}) + b \geq +1 \quad \text{para } y_i = +1 \quad (3.8)$$

$$(\bar{w} \cdot \bar{x}) + b \leq -1 \quad \text{para } y_i = -1 \quad (3.9)$$

Luego la función decisión  $f_{w,b}(\bar{x}_i) = y_i$ , corresponde al signo que resulta de evaluar un dato en la ecuación del OSH (3.3), tal como se presenta en (3.10).

$$f_{w,b}(\bar{x}_i) = \text{sign}[g(\bar{x}_i)] = \text{sign}[(\bar{w} \cdot \bar{x}_i) + b] \quad (3.10)$$

Combinando (3.8) y (3.9) se obtiene (3.11).

$$y_i (\bar{w} \cdot \bar{x} + b) \geq 1 \quad (3.11)$$

Si existe un hiperplano que satisfaga (3.11), se dice que los datos son linealmente separables. Para encontrar el OSH se debe maximizar el margen (3.6), teniendo en cuenta la restricción (3.11), lo que es equivalente a resolver el problema planteado en (3.12) y (3.13).

$$\min_{w,b} \frac{1}{2} (\bar{w} \cdot \bar{w}) \quad (3.12)$$

$$\text{Sujeto a } y_i (\bar{w} \cdot \bar{x}_i + b) \geq 1, \forall i \quad (3.13)$$

La función (3.12) se llama función objetivo, y junto con (3.13) se conforma un problema de optimización cuadrático con restricciones. Los problemas de este tipo se tratan introduciendo el método de multiplicadores de *Lagrange*.

#### b. Solución al problema de optimización con restricciones

Para las restricciones de la forma  $R_i \geq 0$ , cada restricción se multiplica por  $\alpha_i$  (un multiplicador de Lagrange positivo), y se restan de la función objetivo, para así formar la *función de Lagrange* presentada en (3.14) [BURG98].

$$L(\bar{w}, \bar{\alpha}, b) = \frac{1}{2} \|\bar{w}\|^2 - \sum_{i=1}^n \alpha_i [y_i (\bar{w} \cdot \bar{x}_i + b) - 1] \quad (3.14)$$

La función de Lagrange (3.14) se debe minimizar con respecto a las variables primarias  $w$  y  $b$ , y maximizada sobre los  $\alpha_i$  para encontrar el punto de silla [SCHO02]. Para el caso de  $y_i (\bar{w} \cdot \bar{x}_i + b) - 1 > 0$ , el correspondiente  $\alpha_i$  debe ser cero, debido a que éste es el valor que maximiza a (3.14).

Los  $\alpha_i$  diferentes de cero son para el caso en que  $y_i (\bar{w} \cdot \bar{x}_i + b) - 1 = 0$ , que corresponden a los patrones de entrenamiento que quedan sobre los hiperplanos paralelos al OSH dados por la ecuaciones (3.4) y (3.5). Este último enunciado corresponde a las condiciones de *Karush-Kuhn-Tucker* o condiciones complementarias de optimalidad [KUHN51], presentadas de (3.15) a (3.17).

$$\alpha_i [y_i (\bar{w} \cdot \bar{x}_i + b) - 1] = 0, \forall i \quad (3.15)$$

$$\frac{\partial}{\partial w} L(\bar{w}, \bar{\alpha}, b) = 0 \quad \Rightarrow \quad \bar{w} = \sum_{i=1}^n \alpha_i y_i \bar{x}_i \quad (3.16)$$

$$\frac{\partial}{\partial b} L(\bar{w}, \bar{\alpha}, b) = 0 \quad \Rightarrow \quad \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.17)$$

La solución de  $\bar{w}$  en (3.16) queda en función de un subconjunto de patrones de entrenamiento cuyo multiplicador de Lagrange es diferente de cero. Es decir, el soporte de  $\bar{w}$  está en los patrones de entrenamiento más cercanos al OSH. De aquí el nombre de *Máquinas de Soporte Vectorial*.

Reemplazando (3.16) y (3.17) en (3.14), se eliminan las variables primarias  $\bar{w}$  y  $b$  llegando así al *problema de optimización dual de Wolfe*, el cual se resuelve mediante (3.18).

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\bar{x}_i \cdot \bar{x}_j) \quad (3.18)$$

$$\text{sujeto a } \alpha_i \geq 0, \forall i \text{ y } \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.19)$$

Este problema puede ser resuelto con métodos de programación cuadrática estándar [BURG98]. Una vez obtenido el vector  $\vec{\alpha}$ , se pueden obtener los parámetros  $\vec{w}$  y  $b$  con (3.16) y (3.15) respectivamente:

$$\vec{w} = \sum_{i=1}^n \alpha_i y_i \vec{x}_i \quad b = y_i - \vec{w} \cdot \vec{x}_i$$

Ahora, la ecuación del OSH y la función de decisión se pueden expresar como se presentan en (3.20) y (3.21).

$$g(\vec{x}) = \sum_{i=1}^n [\alpha_i y_i (\vec{x}_i \cdot \vec{x})] + b \tag{3.20}$$

$$f(\vec{x}) = \text{sign} \left( \sum_{i=1}^n [\alpha_i y_i (\vec{x}_i \cdot \vec{x})] + b \right) \tag{3.21}$$

La utilización de los multiplicadores de Lagrange se debe a dos razones fundamentales:

- Las restricciones (3.13) están en función de  $\alpha_i$  (3.19), que las hace más fácil de resolver.
- En la reformulación del problema, los datos de entrenamiento  $\vec{x}_i$  sólo aparecen en forma de productos punto entre ellos mismos (3.18), (3.20) y (3.21).

### 3.6.2.2. Análisis del caso linealmente no separable

#### a. Planteamiento general

La implementación de clasificadores basados en SVMs desarrolladas como se presenta en las secciones anteriores, puede tener altos errores, debido a que en la práctica no necesariamente existe un hiperplano separador, y si existe, no siempre es la mejor solución para el problema de clasificación. Cuando existen datos erróneos, ruido o alto traslapamiento de clases en los datos de entrenamiento, ésta solución puede no ser la mejor.

Debido a lo anterior, se propone otra alternativa que busque el mejor hiperplano tolerando ruido en los datos de entrenamiento. Una solución aparente consiste en encontrar el hiperplano que conduzca al menor número de errores de entrenamiento, pero desafortunadamente, esto se convierte en un problema combinatorial difícil de aproximar. Cortes y Vapnik [CORT95], proponen un planteamiento diferente para las SVMs, basándose en [BENN92], para permitir la posibilidad de ejemplos que violen la restricción (3.13), por la consideración de variables de relajación (slack), como se presenta en la figura 3.3.

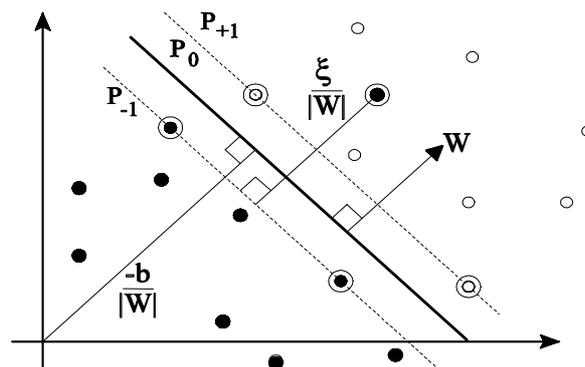


Figura 3.3. Hiperplano lineal clasificador para el caso no separable.

Donde:

$$\xi_i \geq 0, \quad \forall i \tag{3.22}$$

La nueva restricción se presenta en (3.23).

$$y_i (\vec{w} \cdot \vec{x}_i + b) \geq 1 - \xi_i, \quad \forall i \tag{3.23}$$

Luego se encuentra el clasificador que mejor generaliza, controlando su capacidad de clasificación (con  $\|w\|$ ), y el límite superior del número de errores de entrenamiento ( $\sum_{i=1}^n \xi_i$ ). Una posible forma de obtener el hiperplano óptimo con margen débil es minimizando la función (3.24) sujeta a (3.25).

$$\min_{w,b} \frac{1}{2} (\bar{w} \cdot \bar{w}) + C \sum_{i=1}^n \xi_i \quad (3.24)$$

$$\text{sujeto a: } y_i (\bar{w} \cdot \bar{x}_i + b) \geq 1 - \xi_i, \forall i \quad (3.25)$$

Un alto valor del parámetro  $C$  corresponde a una alta penalización a los errores. Con los multiplicadores de Lagrange el problema se transforma en (3.26), sujeto a (3.27).

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\bar{x}_i \cdot \bar{x}_j) \quad (3.26)$$

$$\text{sujeto a: } 0 \leq \alpha_i \leq C, \forall i \text{ y } \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.27)$$

La solución es (3.16). El hiperplano separador solución se puede expresar como (3.20), y la función de decisión como (3.21).

### b. Máquinas de soporte no lineales

El principio de las SVM no lineales consiste en “mapear” o establecer una relación entre el espacio de entrada y un espacio de representación de dimensión alta, a través de una función no lineal elegida a priori, tal como se presenta en la figura 3.4 [BOSE92].

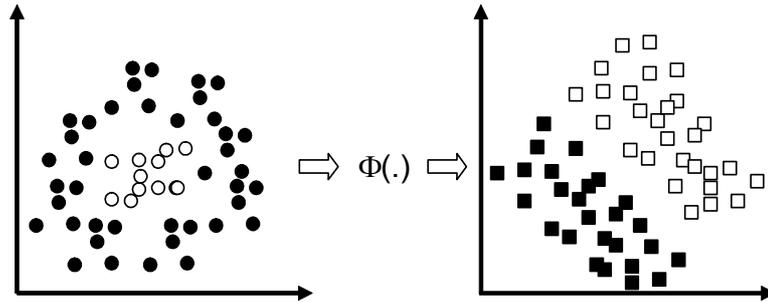


Figura 3.4. La SVM no lineal establece una relación del espacio de entrada con otro de representación de dimensión alta.

Por medio de una función ( $\Phi$ ) se trazan los datos de entrada ( $\bar{x}_i \in \mathcal{R}^N$ ) a algún espacio de mayor dimensión y con producto punto definido (3.28). Este espacio se llama *espacio característico* ( $F$ ).

$$\Phi: \mathcal{R}^N \rightarrow F \quad (3.28)$$

Así, la función (3.20) que depende del producto punto de los vectores en el espacio de entrada, pasa a una función que depende del producto punto de los vectores en el espacio característico, como se muestra en (3.29).

$$g(\bar{x}) = \sum_{i=1}^n [\alpha_i y_i (\Phi(\bar{x}_i) \cdot \Phi(\bar{x}))] + b \quad (3.29)$$

Entonces se define una función que sea el producto punto de los vectores en el espacio característico, presentada en (3.30).

$$k(\bar{x}_i, \bar{x}) = \Phi(\bar{x}_i) \cdot \Phi(\bar{x}) \quad (3.30)$$

Debido a que ( $F$ ) es de alta dimensión, el lado derecho de la ecuación (3.30) es costoso en términos computacionales. Sin embargo existe una función *kernel* ( $k$ ), que puede evaluarse eficazmente y se puede demostrar que corresponde a un trazado de ( $\Phi$ ) en un espacio que abarca todos los productos punto. Un ejemplo para el *kernel* polinomial se muestra en (3.31), (3.32) y (3.33).

$$k(\bar{x}, \bar{y}) = (\bar{x} \cdot \bar{y})^d \text{ con } (\bar{x}, \bar{y}) \in \mathcal{R}^2 \text{ y } d = 2 \quad (3.31)$$

Se tiene que:

$$\begin{aligned}
 k(\vec{x}, \vec{y}) &= \left[ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right]^2 \\
 k(\vec{x}, \vec{y}) &= [x_1 y_1 + x_2 y_2]^2 \\
 k(\vec{x}, \vec{y}) &= x_1^2 y_1^2 + 2x_1 y_1 x_2 y_2 + x_2^2 y_2^2
 \end{aligned} \tag{3.32}$$

(3.32) se puede escribir como (3.33).

$$k(\vec{x}, \vec{y}) = \left[ \begin{pmatrix} x_1 \\ \sqrt{2}x_1 x_2 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \sqrt{2}y_1 y_2 \\ y_2 \end{pmatrix} \right] = \Phi(\vec{x}) \cdot \Phi(\vec{y}) \tag{3.33}$$

Con la función *kernel* no se necesita definir explícitamente la función  $\Phi$ , ya que ésta entrega directamente al resultado del producto punto, que es lo que realmente interesa para la aplicación de clasificación. Generalizando, se puede probar que por cada función *kernel* que presente una matriz definida positiva, se puede construir una función  $\Phi$  que cumpla con (3.30). Algunos de los *kernel* más utilizados son el polinomial (3.34), función de base radial (RBF) (3.35) y sigmoide (3.36).

$$k(\vec{x}, \vec{y}) = (\vec{x} \cdot \vec{y} + c)^d \quad \text{para } c > 0 \tag{3.34}$$

$$k(\vec{x}, \vec{y}) = e^{-\frac{|\vec{x}-\vec{y}|^2}{2\sigma^2}} \tag{3.35}$$

$$k(\vec{x}, \vec{y}) = \tanh(\kappa(\vec{x}, \vec{y}) + \theta) \tag{3.36}$$

Incluyendo la función *kernel* se puede reescribir (3.26) sujeta a (3.27), como (3.37)

$$\begin{aligned}
 \max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j k(\vec{x}_i, \vec{x}_j) \\
 \text{sujeto a: } 0 \leq \alpha_i \leq C, \forall i \text{ y } \sum_{i=1}^n \alpha_i y_i = 0
 \end{aligned} \tag{3.37}$$

Las ecuaciones del OSH (3.20) y función decisión (3.21) se reescriben como (3.38) y (3.39)

$$g(\vec{x}) = \sum_{i=1}^n [\alpha_i y_i k(\vec{x}_i, \vec{x})] + b \tag{3.38}$$

$$f(\vec{x}) = \text{sign} \left( \sum_{i=1}^n [\alpha_i y_i k(\vec{x}_i, \vec{x})] + b \right) \tag{3.39}$$

### 3.6.3. Multclasificación con SVM

El planteamiento anterior sólo aborda el problema de la biclasificación (clases con valores de  $\pm 1$ ), pero muchos problemas son de más de dos clases ( $y_i \in \{1, 2, \dots, l\}$ ,  $l > 2$ ). Para resolver el problema de multclasificación con máquinas de vectores de soporte se admiten dos tipos de arquitectura, las multclasificadoras y las biclasificadoras generalizadas.

#### 3.5.1.1. Máquina multclasificadora

Esta máquina construye una función clasificadora global, considerando todas las posibles clases. Se intenta resolver el problema, modificando la función objetivo de la SVM de tal manera que permita calcular una única máquina capaz de trabajar con todas las clases a la vez. La función objetivo para este caso se presenta en (3.40) y (3.41) [WEST99].

$$\min_{w_r, b_r, \xi^r} \frac{1}{2} \sum_{r=1}^{\ell} \|\bar{w}\|^2 + \frac{C}{m} \sum_{i=1}^m \sum_{r \neq y_i} \xi_i^r \quad (3.40)$$

$$\text{sujeto a: } (\bar{w}_{y_i} \cdot \bar{x}_i + b_{y_i}) \geq \bar{w}_r \cdot \bar{x}_i + b_r + 2 - \xi_i^r, \quad \xi_i^r \geq 0, \quad (3.41)$$

$$m \in \{1, \dots, \ell\}, \quad y_i \in \{1, \dots, \ell\}$$

La principal desventaja de esta generalización se encuentra en el problema de optimización cuadrático, donde hay  $n$  ( $l-1$ ) variables sujetas a  $2n$  ( $l-1$ )+ $2l$  restricciones de desigualdad. Esta desventaja se ha confirmado con experimentos donde se muestra que este tipo de solución es lenta. Además, no hay referencias de técnicas que mejoren la robustez del sistema o de estudios sobre la cota de error [ANGU01].

### 3.5.1.2. Máquinas biclasificadoras generalizadas

Este tipo de máquina construye una función clasificadora global a partir de un conjunto de funciones biclasificadoras. Existen técnicas de descomposición y reconstrucción que permiten a las SVM biclasificadoras manejar problemas de multclasificación con mayor simplicidad y/o menor tiempo de respuesta que una SVM generalizada a multclasificación.

#### a. Arquitecturas de descomposición

En el esquema de descomposición estándar se construyen  $m$  máquinas biclasificadoras, en paralelo que son entrenados sobre modificaciones del conjunto de aprendizaje. Para esto se crea una matriz de descomposición, donde los elementos de unas clases son asignados a salidas positivas, los de otras a salidas negativas. Algunas arquitecturas de descomposición se presentan a continuación.

- **Uno contra el resto**

Conocido como 1-v-r (*one-versus-rest*), este esquema se basa en la idea de que si existe un grupo de  $n$  datos de entrenamiento donde existen  $l$  clases ( $l > 2$ ), se puede tener un grupo de  $m$  clasificadores binarios (donde  $m=l$ ), cada uno entrenado para separar una clase del resto de clases existentes ( $l-1$ ).

La descomposición, considera que existe un grupo de datos ( $n_j$ ) que pertenecen a la  $j$ -ésima clase ( $j \in \{1, \dots, l\}$ ), a los que se asigna una etiqueta positiva ( $t_j = +1$ ), y al resto de datos ( $n_r = n - n_j$ ) se les dará una etiqueta negativa ( $t_r = -1$ ) para el entrenamiento de la  $i$ -ésima SVM. Así se crea una matriz de descomposición ( $D_{1-v-r}$ ) de  $m$  filas y  $l$  columnas que se presenta en (3.42).

$$D_{i,j} = \begin{cases} +1 & \text{si } n_h \in n_j \\ -1 & \text{si } n_h \in n_r \end{cases} \quad (3.42)$$

Por ejemplo, para una máquina de clasificación multiclase (1-v-r) con  $l=5$ , se obtiene  $m=5$ . Entonces la correspondiente matriz de descomposición está dada por (3.43).

$$D_{1-v-r} = \begin{pmatrix} +1 & -1 & -1 & -1 & -1 \\ -1 & +1 & -1 & -1 & -1 \\ -1 & -1 & +1 & -1 & -1 \\ -1 & -1 & -1 & +1 & -1 \\ -1 & -1 & -1 & -1 & +1 \end{pmatrix} \quad (3.43)$$

En esta arquitectura propuesta por [CORT95], el tiempo de entrenamiento es proporcional al número de clases, y debido a que el entrenamiento de cada biclasificador es con el conjunto de datos de entrenamiento completo su costo computacional es alto.

- **Uno contra uno**

Conocido como 1-v-1 (*one-versus-one*), se realiza implementando  $m=l(l-1)/2$  clasificadores binarios.

El entrenamiento de la  $i$ -ésima SVM se realiza con sólo 2 de las  $l$  clases existentes ( $l > 2$ ) en el grupo de  $n$  datos de entrenamiento. Se asigna la etiqueta positiva ( $t_j = +1$ ) a los datos ( $n_j$ ) que pertenecen al subgrupo de datos de la clase  $j$  ( $j \in \{1, \dots, l\}$ ), y la etiqueta negativa ( $t_p = -1$ ) a los

datos ( $n_p$ ) que pertenecen al subgrupo de datos de la clase  $p$  ( $p \in \{1, \dots, l\}$  y  $p \neq j$ ). Los demás datos ( $n_r = n - n_j - n_p$ ) no se utilizan en el entrenamiento de la  $i$ -ésima SVM y por lo tanto son etiquetados con cero ( $t_r = 0$ ), creándose la matriz de descomposición ( $D_{1-v-1}$ ) presentada en (3.44).

$$D_{i,j} = \begin{cases} +1 & \text{si } n_h \in n_j \\ -1 & \text{si } n_h \in n_p \\ 0 & \text{si } n_h \in n_r \end{cases} \quad (3.44)$$

Por ejemplo, para una máquina de clasificación multiclase (1-v-1) con  $l=5$ , se obtiene  $m=10$ . La correspondiente matriz de descomposición se presenta en (3.45).

$$D_{1-v-1} = \begin{pmatrix} +1 & -1 & 0 & 0 & 0 \\ +1 & 0 & -1 & 0 & 0 \\ +1 & 0 & 0 & -1 & 0 \\ +1 & 0 & 0 & 0 & -1 \\ 0 & +1 & -1 & 0 & 0 \\ 0 & +1 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 & -1 \\ 0 & 0 & +1 & -1 & 0 \\ 0 & 0 & +1 & 0 & -1 \\ 0 & 0 & 0 & +1 & -1 \end{pmatrix} \quad (3.45)$$

#### • Arquitectura de descomposición ECOC

La técnica ECOC (del inglés *Error Correcting Output Codes*) [DIET95], utiliza la codificación estándar para obtener robustez contra fallos en las máquinas biclasificadoras.

Se denomina codificación estándar a cada una de las posibles particiones de todo el conjunto de clases  $y_i \in \{1, \dots, l\}$  en problemas de biclasificación que asignan etiquetas positivas  $t_p = +1$  a los patrones de entrenamiento  $n_j$  de un cierto subconjunto de clases  $Y_j$ , y etiquetas negativas  $t_p = -1$  a los patrones de entrenamiento  $n_r$  representantes del resto de clases  $Y_r$ .

La descomposición debe ser tan diferente como sea posible en términos de la distancia de Hamming<sup>3</sup> para añadir redundancia, en este caso  $m=2^{l-1}-1$ .

Como ejemplo, para una máquina de clasificación multiclase (ECOC) con  $l=4$ , se obtiene  $m=7$ , y la correspondiente matriz de descomposición es (3.46).

$$D_{ECOC} = \begin{pmatrix} +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & -1 & +1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & +1 & -1 & +1 \\ +1 & +1 & +1 & -1 \end{pmatrix} \quad (3.46)$$

#### b. Métodos de reconstrucción

Cada máquina biclasificadora entrenada emite una respuesta en forma numérica  $z_i = g_i(\vec{x})$  a una entrada  $\vec{x}$ . La información más importante en esta respuesta, en principio, se encuentra en el signo  $s^i = f_i(\vec{x}) = \text{sign}(g_i(\vec{x}))$  que adopta la función de decisión. En la determinación de la respuesta final del método de reconstrucción de la máquina de aprendizaje multiclase se consideran los siguientes elementos:

- Las predicciones numéricas parciales de los nodos de dicotomía,  $z_i = g_i(\vec{x})$
- El signo de las predicciones numéricas,  $s^i = f_i(\vec{x}) = \text{sign}(g_i(\vec{x}))$

<sup>3</sup> La distancia de Hamming se define como el número de bits que tienen que cambiarse para transformar una palabra de código válida en otra palabra de código válida.

- c. Un elemento intérprete de las predicciones numéricas y binarias,  $\Theta(z^i, s^i)$ , con el fin de asignar o no, una o varias clases como posible respuesta de clasificación a una entrada  $\bar{x}$ .
- d. Un elemento  $\Psi(\Theta(z^1, s^1), \dots, \Theta(z^m, s^m))$  de combinación de las predicciones, que tenga o pueda tener en consideración con las predicciones numéricas, sus signos y/o la clase o clases asignadas.

- **Esquemas de votación**

Son la forma de reconstrucción más habitual y consideran sólo el signo de las predicciones de todas las máquinas biclasificadoras. Estos signos se interpretan en función de las clases implicadas en las máquinas biclasificadoras utilizado en el esquema de descomposición.

- a. i-ésimo *1-v-r* máquina biclasificadora

$$\theta(s^i) = \begin{cases} y_i & \text{si } s^i = +1 \\ 0 & \text{si } s^i = -1 \end{cases} \quad (3.47)$$

- b. i-ésimo *1-v-1* máquina biclasificadora

$$\theta(s^i) = \begin{cases} y_j & \text{si } s^i = +1 \\ y_p & \text{si } s^i = -1 \end{cases} \quad (3.48)$$

- c. i-ésimo *ECOC* máquina biclasificadora

$$\theta(s^i) = \begin{cases} y_j & \text{si } s^i = +1 \\ y_r & \text{si } s^i = -1 \end{cases} \quad (3.49)$$

Tras la interpretación de las predicciones, el elemento de combinación realiza un recuento del número de clases votadas, acción de la que toma el nombre de esquema de reconstrucción, que posee diferentes variantes. A continuación se presentan algunas de estas posibilidades para las arquitecturas de descomposición:

Votación por unanimidad: se determina como respuesta, aquella única clase que haya obtenido todos los votos posibles en las predicciones.

Votación por mayoría absoluta: se determina como respuesta final, aquella única clase que haya obtenido más de la mitad de los votos posibles.

Votación por mayoría simple: se determina como respuesta final, aquella única clase que haya obtenido más votos que el resto de clases.

### 3.5.2. Razón para utilizar SVM en localización de faltas

Las SVM son una herramienta importante que permite hacer uso de la teoría de optimización y la de generalización en el proceso de aprendizaje. A partir de éstas, se asume una alta eficiencia en el proceso de clasificación de la zona de falta, en redes eléctricas. Asimismo, la inclusión de funciones “kernel” no definidas explícitamente, permite resolver problemas no linealmente separables en el espacio de las variables de entrada, lo cual le da una potencia adicional a esta metodología.

Adicionalmente, el éxito en otros tipos de aplicaciones complejas, muestra que es una técnica de gran potencialidad para ser explorada mediante la solución del problema de localización de faltas.

## 3.7 Clasificación basada en LAMDA

### 3.6.1 Generalidades

El método LAMDA (Learning Algorithm for Multivariate Data Analysis) o Algoritmo de aprendizaje para el análisis de datos multivariantes, es una estrategia de clasificación desarrollada por N. Píera [PIER87] basada en una idea original de J. Aguilar [AGUI82]. Esta ha sido aplicada por varias aplicaciones, tal como en el SYCLARE [DESR87], y más recientemente en el desarrollo

de LAMDA 2 [AGUA98] y LINDA [WAIS00]. La formación y el reconocimiento de clases en este método están basados en la atribución de un objeto a una clase a partir de la norma heurística llamada *adecuación máxima*.

El método LAMDA representa un sistema de clases o de conceptos mediante la conexión lógica de toda la información marginal disponible. De esta manera la adecuación global de un objeto a una clase se calcula a partir de la adecuación marginal de cada atributo, y según una regla heurística de adecuación máxima, el objeto pertenecerá a la clase que presente el mayor grado de adecuación.

LAMDA es una metodología de clasificación conceptual, debido a que los objetos no clasificados son confrontados a un prototipo o concepto formado por cada una de las clases existentes. El carácter difuso de los prototipos modela la imprecisión en la formación de conceptos.

La diversidad de problemas abordadas por el método LAMDA es muy amplia e incluye el reconocimiento de formas [AGUI82], las aplicaciones médicas [CHAN89] y el diagnóstico de procesos industriales [AGUA98], entre otros. Esta flexibilidad radica en varias propiedades del método como la posibilidad de realizar un aprendizaje supervisado y no supervisado; la utilización de atributos cualitativos y/o cuantitativos; y la posibilidad de realizar un aprendizaje secuencial. La complejidad del algoritmo de aprendizaje y el reconocimiento, dependen directamente de funciones específicas adoptadas y para ciertas funciones el método LAMDA es muy potente frente al volumen de cálculo, tanto para el aprendizaje como para el reconocimiento.

Una propiedad importante que diferencia a LAMDA de otros métodos de clasificación consiste en la capacidad de modelar de manera natural la indistinguibilidad total, o la homogeneidad en el universo de descripción o ninguna de las anteriores. Esto es una clase especial denominada *clase no informativa* (NIC), que acepta todo los objetos que pueden estar contenidos en el universo de descripción con el mismo grado de adecuación. Según el principio de adecuación máxima esta clase representa un límite mínimo de asignación de un objeto a una clase considerada como significativa. Así, a diferencia de la mayoría de los algoritmos de clasificación, el límite mínimo de aceptación de un objeto a una clase no se fija arbitrariamente, sino que se determina por el propio contexto [AGUA98].

### 3.6.2 Metodología del LAMDA

A partir de un objeto  $x$ , descrito por un número finito de atributos  $[x_1, x_2, \dots, x_n]$ , que se desea confrontar con las diferentes clases  $C_j$ , existe una función de adecuación  $M_{ij}: D_i \times C \rightarrow [0, 1]$  denominada *grado de adecuación marginal* (GAM). Éste se calcula como una función del grado de adecuación de los atributos del objeto a los de la clase, y expresa la forma en que el espacio de descripción correspondiente es representado por la clase  $C_j$  [WAIS00].

El GAM es una función de pertenencia propia de la lógica difusa. Esta función expresa un grado de *adecuación* de un atributo a una clase, o la *inadecuación* de este atributo a esa clase. Entre estos dos valores extremos, existe un valor del atributo tal que a partir de esta única información sea imposible negar o aceptar la pertenencia de este objeto a esta clase. Esto es equivalente a una adecuación *neutra*.

El concepto de adecuación neutra es necesario en la representación de la indistinguibilidad en el algoritmo de clasificación. La expresión de una adecuación neutra, para todo valor en el espacio de descripción, es equivalente a la indistinguibilidad de una clase a partir de la información de este atributo. La clase NIC equivale a considerar indistinguibles todos los atributos.

A partir de lo anterior, se puede asociar a un objeto  $x$  para cada clase  $C_j$  un vector  $[M_{1,j}(x_1), \dots, M_{n,j}(x_n)]$  o  $M_{i,j}$  que expresa el grado de adecuación marginal del atributo  $x_i$  para la clase  $C_j$ . La información de este grado debe ser incorporada a fin de obtener un indicador que permita saber como un objeto no clasificado satisface las condiciones propias de la clase  $C_j$ . Este indicador es modelado por un operador lógico de agregación  $L: [0, 1]^n \rightarrow [0, 1]$  que es conocido como el *grado de adecuación global* (GAG) [WAIS00].

En lógica, la manera clásica de agregar la información se hace por la inclinación de operadores de conjunción y de disyunción. Si se utiliza el operador de intersección, un objeto tendrá una adecuación elevada a una clase solamente si todos los atributos del objeto tienen un grado de adecuación elevada a esta clase. Al contrario en el caso del operador de unión, si sólo uno de estos

atributos presenta un grado de adecuación marginal elevado, es suficiente para considerar este objeto adecuado a la clase correspondiente. Sin embargo, es normal encontrar situaciones en las cuales no se debe ser tan estricto para utilizar el operador intersección, ni tan permisivos para usar el operador de unión. Para esta situación, existen operadores lógicos de agregación que tienen este propósito. Un estado entre la unión y la intersección, puede variar la exigencia del método. El esquema general del cálculo de adecuación de un objeto  $x$  a una clase  $C_j$  se representa en la figura 3.5.

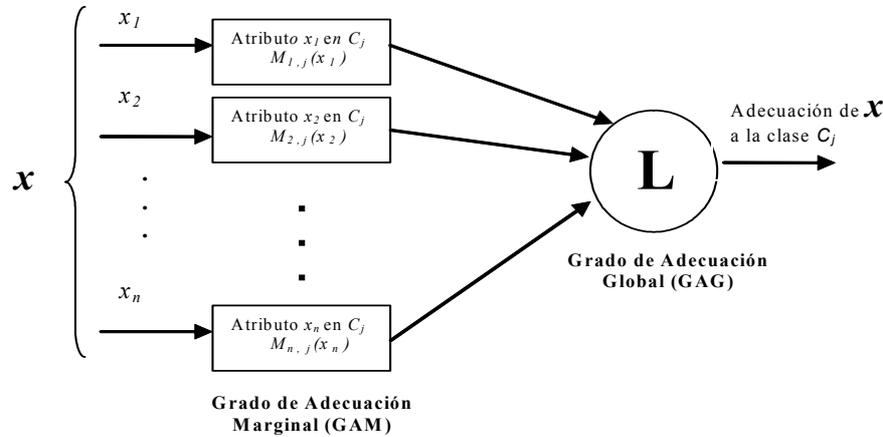


Figura 3.5. Esquema general del método LAMDA. Adecuación del objeto  $x$  a la clase  $C_j$

En consecuencia de lo anterior, el resultado final de la clasificación no es más que una partición clásica del universo de descripción, donde a cada objeto del mismo se le asigna un valor de pertenencia a cada una de las clases. A fin de obtener una partición clásica del universo de descripción, el objeto  $x$  será asignado a la clase donde el valor del GAG es máximo.

El aprendizaje consiste en extraer a partir de la información contenida en una base de datos conocida, las características que describen mejor cada concepto. En LAMDA, esto es traducido para la estimación de parámetros, definiendo las funciones de pertenencia de GAM. En este caso, la función de pertenencia de cada atributo es estimada independientemente de la información disponible sobre otros atributos. Los parámetros que representan una clase son estimados únicamente, a partir de los datos del conjunto de aprendizaje pertenecientes a esta clase. Esto implica que en LAMDA el aprendizaje de un concepto (o clase), se realiza únicamente a partir de la información de la que se dispone y no por la aplicación de otros conceptos preestablecidos [PIER87].

### 3.6.3 Función de pertenencia para el cálculo del DAM

La determinación de la función de pertenencia a partir de datos de entrada es una operación muy importante de la aplicación de la lógica difusa a cualquier situación real. Sin embargo, no existe una guía o regla que se pueda utilizar para escoger el mejor método de cálculo de estas funciones. Además no existe medida para evaluar la calidad de una función de pertenencia seleccionada. Otro inconveniente en la definición de la función de pertenencia es la falta de consenso con respecto a la interpretación de la función de pertenencia, tal como se presenta en [DUBO94].

Para delimitar los tipos de funciones de pertenencia usados por el método LAMDA, se establecen las siguientes restricciones:

- El grado de adecuación de un objeto a una función de pertenencia tiene una interpretación de naturaleza conceptual. Esto se debe interpretar como que la distancia normalizada al prototipo ideal es igual a uno menos el grado de adecuación a la función de pertenencia.
- Las funciones de pertenencia utilizadas para el GAM deben reflejar la adecuación del valor de un atributo a una clase con relación a la inadecuación. En LAMDA, los valores mínimos y máximos posibles de una función son 0 y 1, que significan una total adecuación del atributo o una total inadecuación, respectivamente. Entre los valores extremos de la función de pertenencia, un valor de adecuación neutro debe estar representado por un grado de pertenencia

bien definido y conocido. Las funciones de pertenencia utilizados en LAMDA son una generalización difusa de una lógica de tres valores (0,1,?).

- c. El modelo de la clase NIC para parámetros precisos, requiere que la función de pertenencia tenga una adecuación neutra en todo el espacio de descripción.

### 3.6.3.1 Función de pertenencia para los espacios de descripción cuantitativos

Las funciones de pertenencia utilizadas en LAMDA se basan en la estimación de una partición probabilística en {0, 1}, con una generalización difusa en el espacio [0, 1], tal como se propone en [AGUI80].

Considerando un atributo  $x$  del objeto  $\bar{x}$  que posee solamente la información relativa (la presencia ( $a(x) = 1$ ), o a la ausencia ( $a(x) = 0$ ) de uno de sus valores), el problema de clasificación se expresa como la probabilidad para que  $x \in C_j$ , si  $a(x) = 1$ , o la probabilidad de que  $x \notin C_j$  si  $a(x) = 0$ . Una forma aplicada a los espacios de descripción binarios, evalúa la función para el cálculo de GAM como una probabilidad Bayesiana, y está dada por (3.50).

$$GAM(x) = \rho^{a(x)} q^{1-a(x)} \tag{3.50}$$

Donde  $\rho$  es la probabilidad de que suceda el evento estocástico  $a(x) = 1$  y  $q$  es la probabilidad de fracaso por lo tanto  $q = 1 - \rho$ .

La función basada en la probabilidad binomial, contiene intrínsecamente un grado entre la adecuación y la inadecuación para el cual no importa el valor de  $a(x)$ . Este valor define la clase no informativa NIC y corresponde a  $q = \rho = 0,5$ .

En conclusión el GAM depende únicamente del parámetro  $\rho$  de la función de pertenencia y de los límites superior e inferior del espacio de descripción del atributo en cuestión. En consecuencia, el aprendizaje es más simple y la velocidad de cálculo es reducida [WAIS00]. En la figura 3.6 se muestra la función para el cálculo del GAM para diferentes valores.

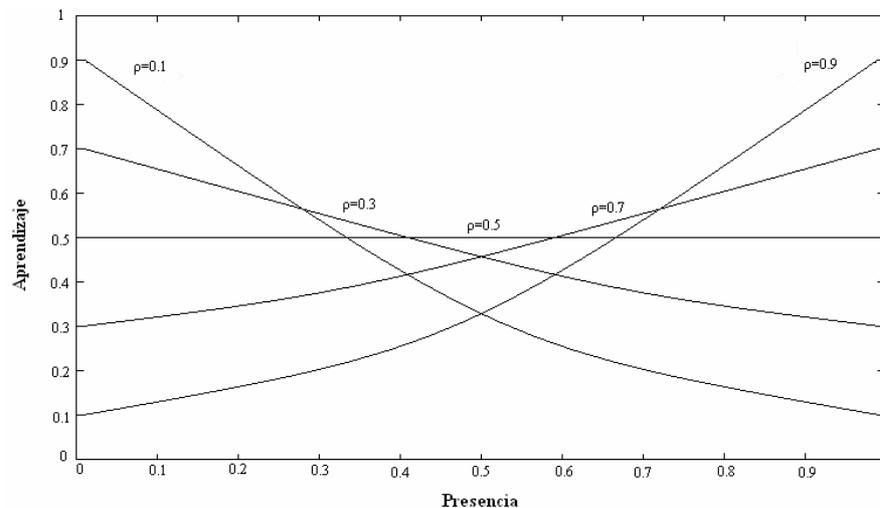


Figura. 3.6: Función para el cálculo de GAM para diferentes valores de  $\rho$  ( $q = 1 - \rho$ )

Otra forma alternativa para el cálculo del GAM, propuesta en [WAIS00], está fundamentada en la presunción de que  $a(x)$  es un grado de presencia por reporte de un prototipo ideal. De esta manera se define el grado de presencia como inversamente proporcional a la distancia normalizada entre  $x$  y un valor considerado como ideal o prototipo. Por lo tanto la presencia está dada por (3.51).

$$a(x_i) = a(x_i, c_{i,j}) = 1 - d(x_i, c_{i,j}) \tag{3.51}$$

Donde  $c_{i,j}$  es el centro de la clase  $C_j$ . El GAM se puede calcular como (3.52)

$$GAM_{i,j}(x_i) = \rho_{i,j}^{a(x_i, c_{i,j})} q_{i,j}^{1-a(x_i, c_{i,j})} \tag{3.52}$$

El hecho de considerar la función de presencia como una distancia a un valor porcentual conserva el número de parámetros determinantes de la clase. Adicionalmente, como  $d(x,c)$  es una distancia unidimensional. Todas las funciones monótonas y estrictamente crecientes pueden ser consideradas útiles. Por razones de simplicidad, las más utilizadas son el error absoluto entre  $x$  y  $c$  ( $d(x,c)=|x-c|$ ) o el error cuadrático ( $d(x,c)=(x-c)^2$ ).

### 3.6.3.2 Función de pertenencia para los espacios de descripción cualitativos

Los atributos como “Forma” y “Color” son dos ejemplos del espacio de descripción cualitativos, es decir, valores que puedan formar un conjunto discreto no ordenado  $D_i = \{q_1, \dots, q_m\}$ , los elementos de tal conjunto se llaman *modalidades*. Es necesario conocer el conjunto de modalidades que puede tomar cada atributo para que tenga un sentido; por ejemplo, si todos los objetos tienen la “forma” cuadrada, entonces este atributo no aporta ninguna información para propósitos de clasificación.

En este caso, se puede utilizar una función de pertenencia basada en histogramas. El conjunto de modalidades de un espacio de descripción está representado en forma de tabla donde cada modalidad está asociada a un valor  $\phi_k$  que indica la frecuencia de ésta en los objetos que pertenecen a la clase  $C_j$ . Para el cálculo del GAM, el valor cuantitativo del atributo  $q_k$  designa una máscara que cubre todas las otras modalidades, y solamente la frecuencia de la modalidad correspondiente se tiene en cuenta. Esto equivale a una función basada en la ley de probabilidad polinomial (3.53).

$$GAM(x_i) = \phi_1^{q_1} \cdot \phi_2^{q_2} \dots \phi_m^{q_m} \quad (3.53)$$

De tal forma que se cumple (3.54)

$$\begin{aligned} q_k &= 1 \quad \text{si } x_i = q_k \\ q_k &= 0 \quad \text{si } x_i \neq q_k \end{aligned} \quad (3.54)$$

El hecho de que todas las frecuencias sean iguales  $\phi_k = \phi_1 = 1/m$  equivale a que más atributos de adecuación a una modalidad específica sean de la clase NIC.

### 3.6.4 Aprendizaje

El aprendizaje consiste en la estimación de parámetros que mejor representan cada clase en el universo de descripción. El aprendizaje es entonces una optimización de las variables que definen una clase en el cálculo del GAM.

Para un espacio de descripción  $D_i$ , se define  $[x^i]_i = [x^0, x^1, \dots, x^n]$  como un vector de atributos en  $D_i$  de  $n$  objetos observados que pertenecen a la clase  $C_j$ . El valor de  $x^i$  es considerado como el valor del  $i$ -ésimo objeto para un atributo y una clase específica.

En principio no se tiene ninguna información a priori concerniente a la relación de los diferentes objetos, para que exista la hipótesis de independencia estocástica entre los diferentes observaciones  $x^i$ .

#### 3.6.4.1 Aprendizaje supervisado de atributos cuantitativos

Si el espacio de descripción es cuantitativo, el aprendizaje consiste en la estimación de dos parámetros que definen la clase:  $\hat{\rho}$  y  $\hat{c}$ . Al elegir el criterio de máxima probabilidad, la estimación óptima será tal que maximiza el criterio (3.55)

$$J(\hat{\rho}, \hat{c}) = \max_{\rho, c} (M(x^i)_{i=1}^n) \quad (3.55)$$

Ahora bien, al hacer la hipótesis de independencia de las diferentes observaciones, se tiene 3.56

$$J(\hat{\rho}, \hat{c}) = \max_{\rho, c} \left( \prod_{i=1}^n M(x^i) \right) \quad (3.56)$$

$$\prod_{i=1}^n M(x^i) = \rho^{\sum_{i=0}^n a(x^i, c)} q^{n - \sum_{i=0}^n a(x^i, c)}$$

Los valores extremos de esta función se encuentran estimando el gradiente (3.57)

$$\nabla \left( \prod_{i=1}^n M(x^i) \right) = 0 \tag{3.57}$$

Existen 4 soluciones posibles bajo condiciones de primer orden representados por la ecuación (3.57). Las dos primeras son  $\rho = 1$  y  $\rho = 0$ , que se descartan puesto que  $\rho$  está definido en el intervalo abierto  $(0, 1)$ . Una tercera solución es que  $\rho = g$  tal que  $g = n(\rho)$ , que caracteriza la clase NIC, pues es una solución trivial. La cuarta solución que se considera la óptima por el criterio de máxima variabilidad está determinada por la solución que caracteriza las condiciones (3.58).

$$\begin{aligned} \hat{\rho} &= \sqrt[n]{\frac{1}{n} \sum_{i=0}^n a(x^i, \hat{c})} \\ \sum_{i=0}^n \frac{d}{dc} a(x^i, \hat{c}) &= 0 \\ \hat{\rho} &> g \end{aligned} \tag{3.58}$$

Una ventaja importante del método LAMDA radica en la capacidad de representar la función de aprendizaje de manera secuencial, esto proviene del hecho que la estimación óptima de  $\rho$  es la media de las funciones de presencia y existe un algoritmo iterativo para su cálculo. La tabla 3.2 muestra la solución de la estimación óptima para las tres funciones de presencia propuestas.

Función de presencia	Función de aprendizaje	Función de recursiva
$x$	$\hat{\rho} = \frac{1}{T} \sum_{i=1}^T x^i$	$\hat{\rho}^{i+1} = \hat{\rho}^i - \frac{1}{i+1} (x^{i+1} - \hat{\rho}^i)$
$1 -  x - c $	$\hat{c} = \text{mediana}_{i=1}^T(x^i)$ $\hat{\rho} = 1 - \frac{1}{T} \sum_{i=1}^T  x^i - c $	-
$1 - (x - c)^2$	$\hat{c} = \frac{1}{T} \sum_{i=1}^T x^i$ $\hat{\rho} = 1 - \frac{1}{T} \sum_{i=1}^T (x^i - c)^2$	$\hat{c}^{i+1} = \hat{c}^i + \frac{1}{i+1} (x^{i+1} - \hat{c}^i)$ $\hat{\rho}^{i+1} = \hat{\rho}^i - \frac{1}{(i+1)} \left[ (x^{i+1} - \hat{c}^{i+1})^2 - \dots - (1 - \hat{\rho}^i) + \frac{1}{(i+1)^2} (x^{i+1} - \hat{c}^i)^2 \right]$

Tabla 3.2. Funciones de aprendizaje ( $q = 1 - \rho$ )

### 3.6.4.2 Aprendizaje supervisado de atributos cualitativos

Para un espacio de descripción cualitativo,  $\phi_k$  se calcula mediante la frecuencia de la modalidad correspondiente. Si  $n$  es el número de elementos que pertenecen a  $C_j$ , dentro de una base de datos de aprendizaje, es decir  $n_k$  es el número de elementos que pertenecen a la clase que está representada por la modalidad  $q_k$ ,  $\phi_k$  está determinado por  $n/n_k$ . Debido a que pueden existir problemas cuando una clase no tenga elementos, se tiene una corrección dada por  $\phi_k = (n_k + f) / ((n + f) \cdot m)$ , donde  $m$  es el número de modalidades del atributo y  $f$  para la mayoría de clasificadores Bayesianos es  $1/N$ , donde  $N$  es el número total de elementos del conjunto de aprendizaje.

Otra opción posible proviene de interpretar la frecuencia como un promedio de una función de presencia. Se define  $\delta_k = 1$  si una modalidad está presente dentro del atributo de un objeto  $l$  de la base de aprendizaje, y  $\delta_k = 0$  en otro caso. La frecuencia de la modalidad  $k$  para  $T$  objetos de la base de datos de aprendizaje está dada por (3.59).

$$\phi_k^T = \frac{\text{Número de } \delta = 1}{\text{Número total de elementos}} = \frac{\sum_{k=1}^T \delta_k}{T} \quad (3.59)$$

Entonces el aprendizaje está dado por (3.60).

$$\phi_k^{l+1} = \phi_k^l + \frac{1}{l+1} (\delta_k^{l+1} - \phi_k^{l+1}) \quad (3.60)$$

### 3.6.4.3 Aprendizaje no supervisado

Dos de las propiedades de las funciones utilizadas para el cálculo del GAM permiten que el método LAMDA pueda realizar un aprendizaje no supervisado: Las funciones secuenciales de aprendizaje y la clase no informativa. Esta última introduce un umbral de clasificación que determina naturalmente si es necesario o no crear una nueva clase. Este algoritmo de aprendizaje no supervisado lo desarrolló [PIER89]. Los objetos son tratados secuencialmente y confrontados con las clases existentes. Si un objeto se asigna a la clase NIC, se debe crear una nueva clase, usando los parámetros de este objeto como prototipo. En el caso contrario, el objeto modifica parámetros que definen la clase por medio de la función recursiva de aprendizaje. Inicialmente se considera que la NIC es la única clase existente. Así el número de clases no se impone de antemano pero se determina por el contexto entero.

El aprendizaje es dependiente del orden en el cual se introducen datos del conjunto de entrenamiento en el clasificador. De esta manera la formación de conceptos depende también del orden en el que se presenta la información. Esto presenta ventajas y desventajas. Así se introducen conceptos o clases reconocidas por LAMDA en un instante, para que puedan cambiar a medida que se introduce nueva información. Adicionalmente, los datos que pertenecieron a una clase, tienen la posibilidad de cambiar de clase.

Esta conducta tiene ventajas cuando el sistema analizado evoluciona a medida que se presentan los datos al clasificador (en orden cronológico). Sin embargo, representa un problema para el entrenamiento, si no es posible lograr una fase de estabilización o re-aprendizaje. Una vez todo el conjunto de entrenamiento es usado, se eliminan las clases que no tienen elementos. Los objetos que han cambiado de clase se reintroducen como nuevos datos de entrenamiento. Este procedimiento se repite hasta que ningún objeto cambie de clase y converja hacia un conjunto de clases estable (constante). No obstante como este procedimiento de re-aprendizaje puede ser prohibitivo en cuanto al tiempo de cálculo, se debe usar algún un criterio de parada dado por la variabilidad del objeto a las clases o el número de iteraciones.

### 3.6.5 Operadores lógicos de agregación para el cálculo del GAG

Luego del cálculo del GAM, el problema que se presenta es cómo agregar toda la información marginal para llegar a la adecuación global del objeto a esta clase. Debido que el GAM es un grado de pertenencia, la adecuación global se calcula en el contexto de conjuntos difusos, mediante los cuales se desarrollan las operaciones de intersección, unión negación y de promedio.

Para que una función pueda usarse como operador de agregación, ésta debe ser conmutativa y monótona. Por conmutabilidad se tiene que el orden en que los rasgos identificaban el objeto no debe afectar el reconocimiento. La monotonía hace referencia que si el valor del GAM, aumenta, también lo tiene que hacer el GAG.

A partir de los operadores de base, como las normas T, se han desarrollado nuevos operadores para imitar otros aspectos del razonamiento humano. LAMDA centra el estudio en dos posibles tipos del operador: los conectivos mixtos de compensación y las sumas simétricas para la función generadora factorizable.

#### 3.6.5.1 Conectivos mixtos de compensación

Dentro del contexto de la toma de decisiones, la compensación entre dos situaciones extremas se usa normalmente.

Para este caso, el clasificador más exigente sólo acepta un objeto en una clase, si todos sus atributos presentan un grado de adecuación lo suficientemente elevado a la clase. Desde un punto de vista lógico esta conducta corresponde a la intersección. Dentro de la teoría de conjuntos difusos, esta operación se logra usando una norma T.

De la misma manera, el clasificador menos exigente acepta un objeto en una clase, si por lo menos uno de sus atributos posee una pertenencia marginal elevada dentro de la clase. Desde un punto de vista lógico, esta conducta corresponde a la unión. En la teoría de conjuntos difusos este funcionamiento se logra por una co-norma T que es la función dual de la norma T respecto a una negación estricta.

Mientras se use un operador que permita tener la norma T y la co-norma T, se puede ajustar la exigencia del clasificador mediante el parámetro  $\alpha$  llamado **grado de exigencia**. De esta manera, a partir de un conectivo mixto, se pueden hacer particiones más o menos estrictas del espacio de descripción como se presenta en (3.61).

$$L_{\alpha}(x_1, \dots, x_n) = f^{-1}(\alpha f(T(x_1, \dots, x_n)) + (1 - \alpha) f(C(x_1, \dots, x_n))) \quad (3.61)$$

Para el caso de la clasificación no supervisada, el número de particiones del mismo espacio de aprendizaje tiene un comportamiento creciente de clases, a medida que se aumenta la exigencia.

Los más comunes conectivos mixtos de compensación derivados de funciones promedio son: el lineal (3.62), desarrollado y estudiado en [PIER89]; el geométrico (3.63), desarrollado por [ZIMM80]; el armónico (3.64), y el de raíces y potencias (3.65) [WAIS00].

$$L_{\alpha}(x_1, \dots, x_n) = \alpha T(x_1, \dots, x_n) + (1 - \alpha) C(x_1, \dots, x_n) \quad (3.62)$$

$$G_{\alpha}(x_1, \dots, x_n) = T(x_1, \dots, x_n)^{\alpha} * C(x_1, \dots, x_n)^{(1-\alpha)} \quad (3.63)$$

$$H_{\alpha}(x_1, \dots, x_n) = \frac{T(x_1, \dots, x_n) * C(x_1, \dots, x_n)}{\alpha C(x_1, \dots, x_n) + (1 - \alpha) T(x_1, \dots, x_n)} \quad (3.64)$$

$$R_{\alpha}^a(x_1, \dots, x_n) = \sqrt[a]{\alpha T(x_1, \dots, x_n)^a + (1 - \alpha) C(x_1, \dots, x_n)^a} \quad (3.65)$$

### 3.6.5.2 Sumas simétricas

Las sumas simétricas [SILV79] han sido desarrolladas con la intención de modelar la propiedad de compensación entre atributos. Este tipo de operadores permite combinar conjuntos difusos complementarios.

Contrario a las normas T y las co-normas T de los conectivos mixtos, las sumas simétricas son operadores que permiten compensar la presencia de atributos con grados de adecuación extremos para los valores del atributo. Ellos no se localizan obligatoriamente entre los conectivos máximo y mínimo, como es el caso de los promedios.

Una definición de una suma simétrica generalizada para n atributos está dada por (3.66).

$$L(x_1, \dots, x_n) = \frac{\prod_{i=1}^n G(x_i)}{\prod_{i=1}^n G(x_i) + \prod_{i=1}^n G(n(x_i))} \quad (3.66)$$

Donde  $n(x)$  es la negación de  $x$ . Considerando  $G(x) = x^{\frac{1}{n}}$ , donde  $n$  es la dimensión del vector  $\vec{x}$ . Si se conserva la negación clásica, y para vectores de atributos de la misma longitud, el operador de agregación es independiente ( $L(x, \dots, x) = x$ ). Esta elección corrige el valor del grado de pertenencia para que permanezca uniformemente repartido para una dimensión dada.

Considerando la familia de negaciones estrictas  $n(x) = \sqrt[1-\alpha]{1-x^{\alpha}}$  que para  $\alpha=1$ , incluye la negación clásica. Algunas transformaciones se obtienen del operador (3.67).

$$L(x_1, \dots, x_n) = \frac{1}{1 + \prod_{i=1}^n \left[ \frac{1 - x_i^\alpha}{x_i^\alpha} \right]^{\frac{1}{n\alpha}}} \quad (3.67)$$

### 3.6.6 Razón para utilizar LAMDA en localización de faltas

LAMDA es un método que ofrece características importantes para resolver el problema de la localización de faltas en sistemas de distribución. Por ser un método de clasificación, permite establecer una relación entre la entrada y la salida, de tal manera que un objeto  $x$  corresponde a un conjunto de descriptores extractados de las señales de tensión y corriente medidas antes, durante y después de la falta, y la clase  $C_j$  puede corresponder a la zona donde se presentó la falta.

Adicionalmente, en LAMDA cada clase definida después de la fase de entrenamiento, puede ser representada por un valor prototipo, que corresponde al objeto teórico con el más alto valor de adecuación global GAG. Este prototipo permite tener una explicación del significado físico que representa dicha clase.

En este documento, el método LAMDA se probará junto con las SVM, para determinar su potencial aplicación y éxito en la resolución del problema aquí planteado. Finalmente, este método servirá como elemento de comparación.

### 3.7 Conclusiones

El problema de localización de faltas en sistemas de distribución puede ser abordado por Métodos de Clasificación Basados en el Conocimiento - MCBC, conocidos genéricamente como “clasificadores”. Estos métodos tienen la capacidad de establecer una relación, normalmente no lineal, entre un conjunto de entrada (descriptores), y un conjunto de salida (tipo y región de la falta). En este capítulo se presentó un análisis de los diferentes métodos de clasificación y su posible aplicación al problema de localización de faltas.

Los métodos de clasificación se caracterizan porque tienen dos etapas básicas: el entrenamiento y la validación. El entrenamiento puede ser supervisado o no supervisado y es la etapa donde se crea una relación entre la entrada y la salida, a partir de un conjunto de datos que sirven como ejemplos. La validación consiste en la presentación de datos de entrada para que el clasificador presente una salida similar a la deseada. El error en la validación se estima a partir de la estimación de un grado de similitud entre la entrada y la salida.

Considerando el problema de localización de faltas, si se tiene una base de datos compuesta por registros de corriente y tensión, plenamente relacionada con el sitio de falta, se puede entrenar un clasificador para que realice la localización de la misma.

Existen muchas técnicas de clasificación, cada una con sus ventajas y desventajas. Para la adecuada selección de una técnica en una aplicación en particular, se deben verificar al menos las siguientes características: tipo de atributos disponibles en la base de datos de entrenamiento y validación, interpretabilidad de los resultados, velocidad de aprendizaje y/o de reconocimiento, posibilidad de realizar aprendizaje secuencial y grado de aceptación de la técnica.

Las técnicas LAMDA y SVM son una buena opción de solución del problema de localizar faltas en los sistemas de distribución. La metodología se desarrolla con base en la información monitorizada en los alimentadores de los circuitos. En el capítulo 5 se realizarán pruebas para verificar el comportamiento de cada algoritmo y así seleccionar el mejor para esta aplicación.

## 4 Caracterización de las señales de tensión y corriente para localización de faltas en sistemas de distribución

### 4.1 Introducción

La aplicación de los Métodos de Clasificación Basados en el Conocimiento (MCBC), al problema de localización de faltas en sistemas de distribución, se fundamenta en la necesidad de aprovechar el conocimiento empírico disponible en la información registrada en la subestación. A partir de la utilización de los MCBC, se pretende identificar una región probable de falta, así como el tipo de la misma.

La metodología a seguir se fundamenta en el tratamiento de las señales de tensión y de corriente medidas en la subestación de distribución, antes, durante y después de la falta. Posteriormente, se obtiene un conjunto de descriptores por cada registro de falta. Los diferentes descriptores servirán para elaborar los conjuntos de entrenamiento y validación de los MCBC propuestos. Finalmente, los MCBC validados servirán para la localización de la región probable de falta.

En este capítulo, se presentan inicialmente las definiciones de las características o descriptores obtenidos de las señales de tensión y corriente en la subestación. A continuación se presenta la obtención de cada descriptor, haciendo referencia a la técnica utilizada. Finalmente, se presentan las conclusiones del capítulo.

### 4.2 Definición de los descriptores

Un descriptor es una característica que aporta información significativa sobre un evento. En este caso particular, los descriptores se obtienen de las señales de corriente y tensión, registradas antes y durante la falta, en la subestación de distribución. Así definidos, para el caso del estudio aquí propuesto, los descriptores proporcionan toda la información necesaria para el aprendizaje de los MCBC usados para la localización de la zona más probable de ocurrencia de la falta.

La definición de los descriptores básicos, utilizados para localización de faltas mediante los MCBC, se presenta a continuación:

- a. Profundidad del hueco de tensión ( $dV$ ): Definida como la variación en el valor eficaz de la tensión entre los estados estables de falta y de pre-falta.
- b. Variación de la magnitud de corriente ( $dI$ ): Similar a los huecos de tensión, está definida como la variación en el valor eficaz de la corriente entre los estados estables de falta y de pre-falta.
- c. Variación de la potencia ( $dS$ ): Este descriptor está definido como la variación entre la potencia aparente calculada en estado estable, entre falta y pre-falta.
- d. Reactancia de falta ( $Xf$ ): Definida como la parte imaginaria de la impedancia, obtenida con los valores de tensión y de corriente en estado estable de falta.
- e. Frecuencia del transitorio ( $f$ ): Valor obtenidos a partir del análisis transitorio de la señal de tensión.
- f. Valor propio máximo de la matriz de correlación de corrientes ( $Sv$ ): Definido a partir de la transformación de las muestras de corriente tomadas antes y durante la falta, mediante la transformada Clark-Concordia.

Como complemento a esta sección, se presenta el descriptor conocido como firma del sistema. Con este nombre genérico se conocen las características asociadas a la variación del registro de corriente, causado por la operación de los elementos de protección. Mediante el análisis de las señales de corriente y tensión, se puede relacionar este descriptor con la distancia a la falta [KRIS89][LEWI87]. Este descriptor no se usa como entrada a los clasificadores, pero se presenta aquí dada su importancia para la elaboración de reglas *if-then*, como lo propone el autor como parte de su trabajo exploratorio en [MORA06-b] y en [MORA06-e].

### 4.3 Caracterización del sistema mediante la obtención de descriptores

En esta sección se presenta la caracterización de las señales de tensión y de corriente medidas en la subestación antes, durante y después de una falta en el sistema de distribución. La caracterización se fundamenta en la utilización de la señal transitoria y en estado estable [MORA05-c].

La variación de los valores eficaces de tensión y corriente entre falta y pre-falta, permite obtener fácilmente descriptores para relacionar la distancia de falta con la magnitud de la variación [MORA06-a].

De manera similar, a partir de la variación de la potencia reactiva se puede medir el cambio de la carga y la variación de las constantes del modelo de carga, visto desde la subestación, para después relacionarlos con la distancia a la localización de la falta [IEEE93][SEPA96][LAKE89].

El valor de la reactancia vista desde la subestación, se puede obtener como consecuencia directa de la aplicación de los Métodos Basados en el Modelo [DAS98][MORA05-a].

La frecuencia para el circuito equivalente entre la subestación y el sitio de falta pueden obtenerse usando las transformadas Wavelet y Fourier [GREE91][MART02] [OPPE97].

Finalmente, a través del análisis de las corrientes de línea en una ventana que incluye datos de antes, durante y después de la falta, se pueden obtener el valor máximo de la matriz de correlación mediante el uso de la transformada Clark-Concordia, así como su relación con la distancia al sitio de falta, que en este documento se define como valor propio [SOUS02][SOUS03].

Como resultado de esta caracterización, es posible obtener descriptores para usar como entradas para los MCBC. Mediante el uso de estos descriptores y los métodos clasificadores se pretende localizar la zona de falta en el sistema de distribución.

#### 4.3.1 Profundidad del hueco de tensión ( $dV$ )

El descriptor asociado a la variación del valor eficaz, entre los estados estables de falta y pre-falta, se le conoce como hueco, cuando se trata de la señal de tensión [BOLL00] [OLGU05].

Para obtener esta variación se requiere del valor eficaz del fundamental de pre-falta y de falta, y el descriptor corresponde a la diferencia de estas magnitudes. En la figura 4.1 se presentan de forma gráfica los descriptores para cada una de las fases.

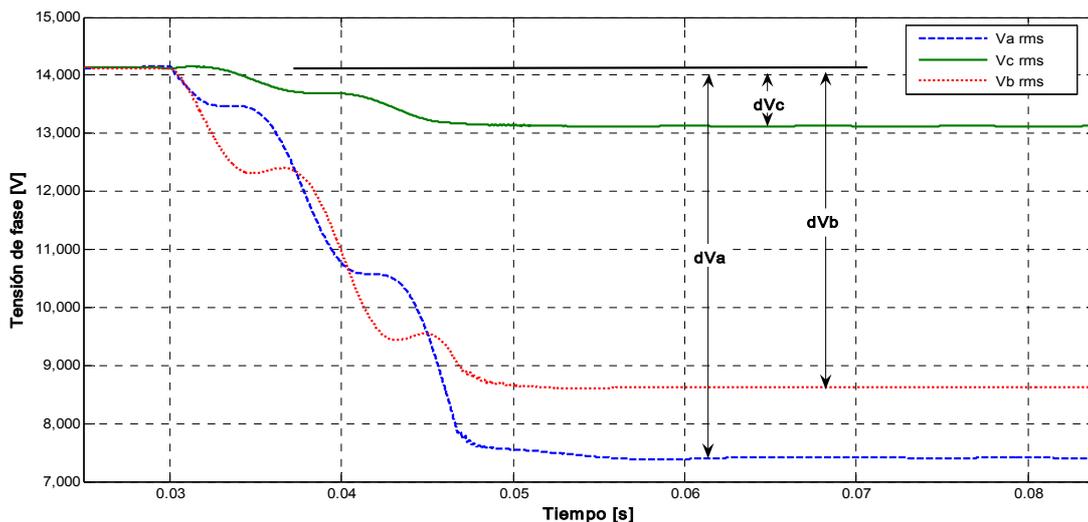


Figura 4.1. Descriptores de la variación del valor eficaz de la tensión

Tal como se presenta en la figura 4.1, se utilizan los tres valores, uno para cada fase. A partir de aquí, y en este documento, siempre que se utilice  $dV$  se está haciendo referencia a la terna conformada por  $dV_a$ ,  $dV_b$  y  $dV_c$ .

El uso de las tres señales está justificado ya que la magnitud de la variación del valor eficaz de la tensión, puede ser el mismo para faltas en dos sitios diferentes. Así, una falta monofásica lejana del

punto de medida con una resistencia de falta baja, puede tener la misma variación de tensión en la fase en falta que para el caso que ésta ocurra en un nodo cercano y con una resistencia de falta alta.

La situación anterior se muestra en las figuras 4.3 y 4.4. En la figura 4.3 se presenta el comportamiento de las señales de tensión para el caso de una falta monofásica A-T en el sistema de prueba de la figura 4.2, en la barra 4 y con una resistencia de falta de  $26\Omega$ . En la figura 4.4 se muestran el valor eficaz de la tensión para el caso de una falta monofásica A-T, en la barra 11, con una resistencia de falta de  $4\Omega$ .

Para el caso del ejemplo presentado, se muestra como la magnitud del cambio del valor eficaz de la fase en falta ( $dV_a$ ), es muy parecido. En las fases que no están en falta, esta variación es sensiblemente diferente ( $dV_b$  y  $dV_c$ ).

Los descriptores asociados a los huecos de tensión se pueden obtener para señales de fase ( $dV$ ) y de línea ( $dVL$ ).

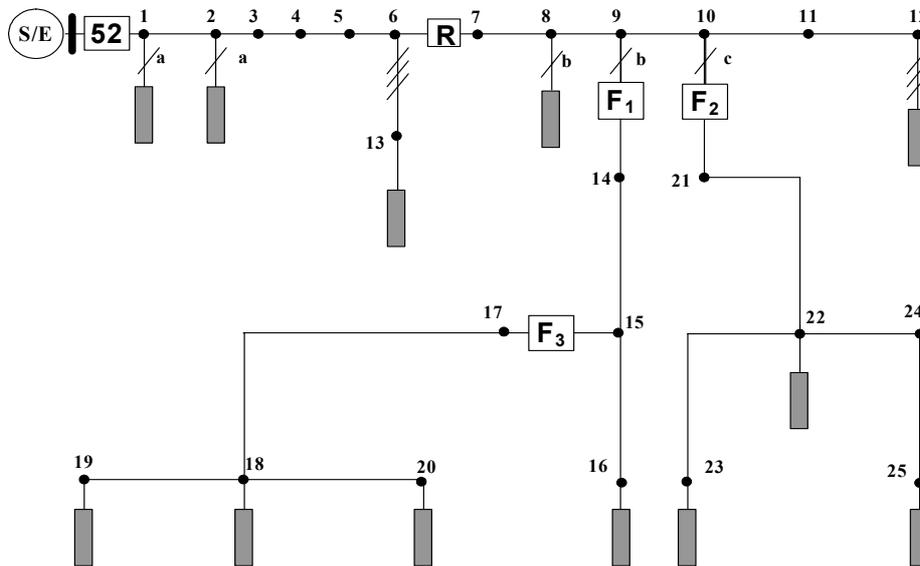


Figura 4.2. Sistema de distribución seleccionado para pruebas.

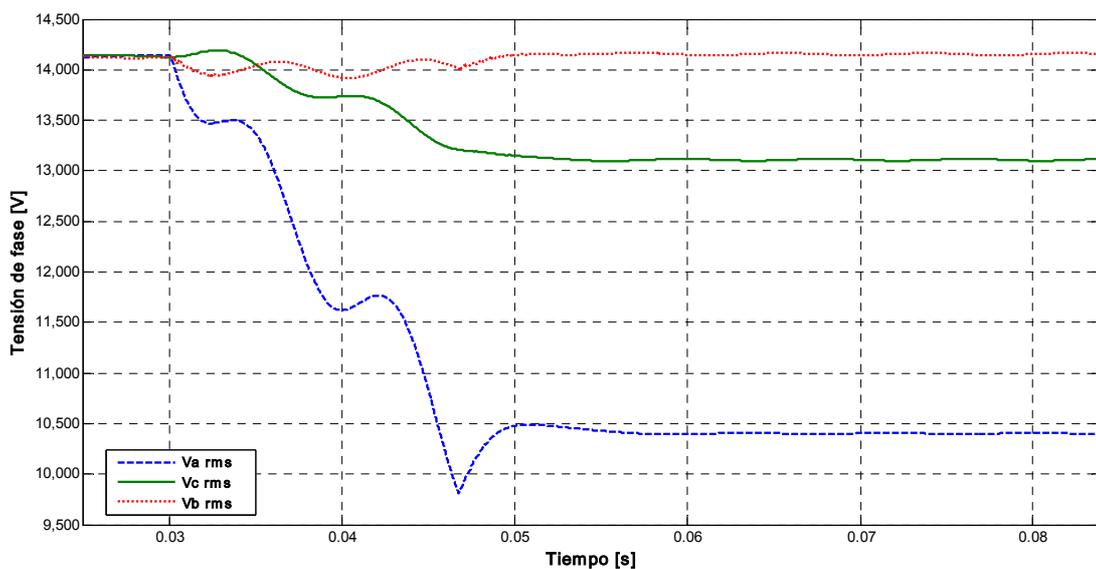


Figura 4.3. Señales de tensión para el caso de una falta monofásica A-T, en la barra 4 y con una resistencia de falta de  $26\Omega$

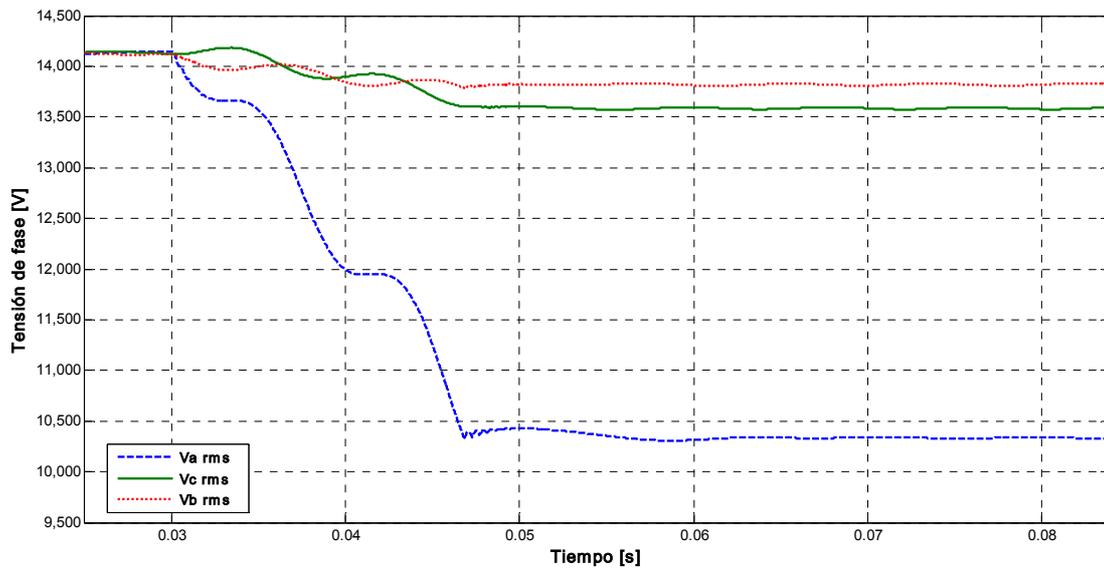


Figura 4.4. Señales de tensión para el caso de una falta monofásica A-T, en la barra 11, con una resistencia de falta de  $4\Omega$ .

### 4.3.2 Variación de la magnitud de corriente ( $dI$ )

Al igual que el descriptor anterior, la variación de la magnitud de la corriente está definida como a la diferencia del valor eficaz, entre los estados estables de falta y pre-falta. En la figura 4.5 se presentan de forma gráfica los descriptores para cada una de las fases.

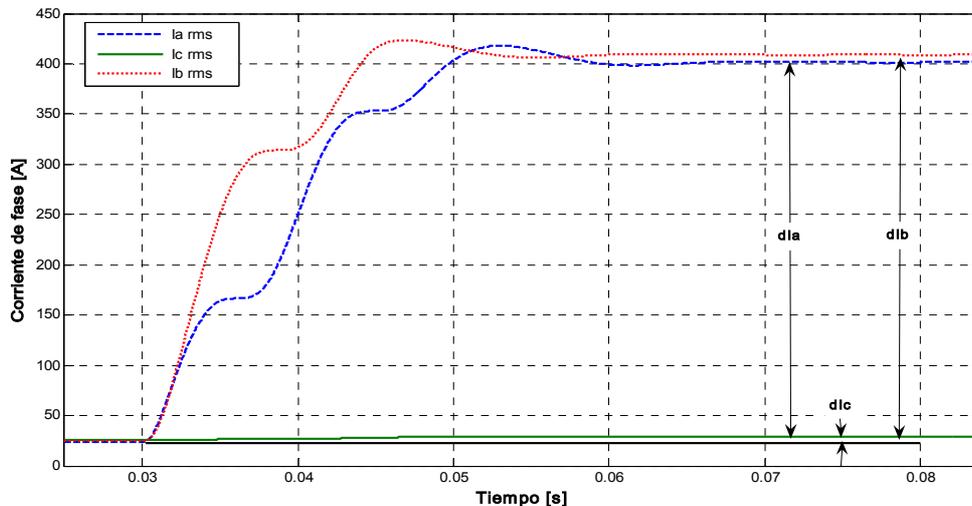


Figura 4.5. Descriptores de la variación del valor EFICAZ de la corriente.

Al igual que para los huecos de tensión, se utilizan los tres valores, uno para cada fase. A partir de aquí, y en este documento, siempre que se utilice  $dI$  se está haciendo referencia a la terna conformada por  $dIa$ ,  $dIb$  y  $dIc$ .

Los descriptores asociados a la variación del valor eficaz de la corriente se pueden obtener para señales de fase ( $dI$ ) y de línea ( $dIL$ ).

### 4.3.3 Variación de la potencia del sistema ( $dS$ )

La variación de la potencia del sistema está asociada a la variación de la carga. La presencia de la falta hace que cambie el circuito visto desde la subestación y por tanto su nivel de cargado. Para la representación de la variación de la carga, se presentan aquí cuatro modelos diferentes propuestos en varios estándares referenciados en [IEEE93].

#### 4.3.3.1 Modelo de carga polinomial

En este modelo, se representa la potencia activa y la reactiva como una función polinomial de la tensión ( $V$ ), tal como se presenta en la ecuación (4.1).

$$\begin{aligned} P &= a_0 + a_1|V| + a_2|V|^2 + \dots \\ Q &= b_0 + b_1|V| + b_2|V|^2 + \dots \end{aligned} \quad (4.1)$$

El anterior tipo de representación es válida para cargas individuales o para cargas agregadas (compuestas). De las ecuaciones anteriores si  $a_0$  y  $b_0$  son diferentes de cero y los demás coeficientes nulos, se tiene un modelo de potencia constante como el que se presenta en (4.2).

$$P = a_0 \quad Q = b_0 \quad (4.2)$$

Otro caso es cuando se toman sólo los valores de los coeficientes  $a_1$  y  $b_1$ , se obtiene el modelo de corriente constante tal como se presenta en (4.3).

$$P = a_1|V| \quad Q = b_1|V|. \quad (4.3)$$

Un último caso de gran utilización se tiene cuando todos los coeficientes son cero, exceptuando  $a_2$  y  $b_2$ , se consigue el modelo de impedancia constante, presentado en (4.4).

$$P = a_2|V|^2 \quad Q = b_2|V|^2 \quad (4.4)$$

#### 4.3.3.2 Modelo exponencial

Este modelo considera cargas compuestas que se asumen como combinaciones de las cargas mencionadas, se representan mediante la ecuación (4.5).

$$S = P_n \left| \frac{|V|}{|V_0|} \right|^k + j Q_n \left| \frac{|V|}{|V_0|} \right|^l \quad (4.5)$$

Donde  $k$  y  $l$  varían entre 0 y 3 [IEEE93].

#### 4.3.3.3 Modelo exponencial dependiente de la frecuencia

Las cargas raramente son modeladas para incluir los efectos de la frecuencia. Si las cargas son sensibles a los efectos de la frecuencia, entonces estos efectos deben ser incluidos, tal como se presenta en la ecuación (4.6).

$$S(f, |V|) = P_n \left[ \frac{\omega}{\omega_n} \right]^\alpha \left[ \frac{|V|}{|V_0|} \right]^k + j Q_n \left[ \frac{\omega}{\omega_n} \right]^\beta \left[ \frac{|V|}{|V_0|} \right]^l \quad (4.6)$$

Donde  $\omega = 2\pi f$  y  $\omega_n = 2\pi f_n$ , con  $f_n$  como frecuencia nominal (60 Hz para el caso Colombiano y 50 Hz para el caso Español). Además  $\alpha$  y  $\beta$  son exponentes constantes.

#### 4.3.3.4 Modelo estadístico

Este modelo es particularmente sensible para modelar cargas agregadas, cuando se conocen las curvas de carga diarias, mensuales, estacionales o anuales, y se asume una distribución Gaussiana como se muestra en (4.7).

$$S = (P_n + k_p \delta) + j (Q_n + k_p \delta) \quad (4.7)$$

Donde  $P_n$  es la potencia nominal,  $p$  es la probabilidad de exceder una potencia,  $k_p$  es un coeficiente relacionado con  $p$  y  $\delta$  es la desviación estándar.

### 4.3.3.5 Modelo agregado de carga en la subestación

Usando los modelos de carga propuestos, es posible determinar un modelo agregado de nodo único, visto desde la subestación, tal como el que se presenta en la figura 4.6.

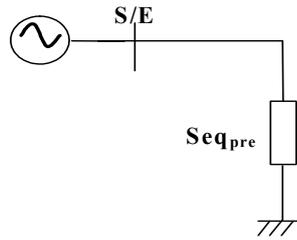


Figura 4.6. Modelo agregado de carga en la subestación

Con los modelos exponenciales, la potencia equivalente del modelo agregado de la subestación antes de la falta está dada por la ecuación (4.8).

$$Seq_{pre} = P_n \left| \frac{V_{pre}}{V_0} \right|^k + j Q_n \left| \frac{V_{pre}}{V_0} \right|^l \quad (4.8)$$

Durante la falta, el modelo presentado en la figura 4.6 sufre una alteración debida a la resistencia de fallo que hace que cambien las condiciones de potencia, tal como se presenta en la figura 4.7, para un modelo de nodo único.

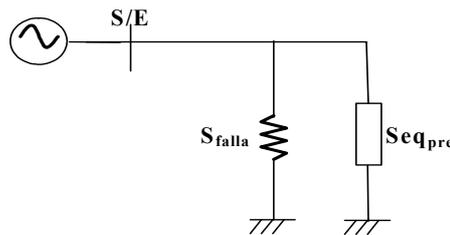


Figura 4.7. Modelo agregado de carga y falta en la subestación

Con este modelo se puede determinar una variación en la potencia aparente, tal como se presenta en la ecuación (4.9).

$$Seq_{falta} = P_n \left| \frac{V_{falta}}{V_0} \right|^k + j Q_n \left| \frac{V_{falta}}{V_0} \right|^l + P_f \left| \frac{V_{falta}}{V_0} \right|^{kf} + j Q_f \left| \frac{V_{falta}}{V_0} \right|^{lf} \quad (4.9)$$

Utilizando como descriptor la variación de la potencia aparente ( $dS$ ), se tiene la ecuación (4.10)

$$dS = Seq_{falta} - Seq_{pre} \quad (4.10)$$

Para el estado estable, la variación de potencia puede ser utilizada para conocer la carga que ha sido desconectada. Adicionalmente, el cambio en el factor de potencia da una indicación adicional de la característica de la carga [SEPA96]. Estos dos descriptores complementan la localización geográfica de la falta, mediante la identificación de la carga retirada del sistema de distribución.

Al igual que los descriptores anteriores, se utilizan los tres valores, uno para cada fase. A partir de aquí, y en este documento, siempre que se utilice  $dS$  se está haciendo referencia a la terna conformada por  $dSa$ ,  $dSb$  y  $dSc$ . Los descriptores asociados a la variación de la potencia aparente se pueden obtener para señales de fase ( $dS$ ) y de línea ( $dSL$ ).

### 4.3.4 Reactancia de falta ( $X_f$ )

Existen varios modelos para representar las admitancias de la carga en estudios de sistemas de

transmisión [DAS98]. La técnica aquí presentada para el análisis usa modelos de carga estáticos, tal como se presenta en la ecuación (4.11), para el nodo  $R$ .

$$Y_R = G_R |V_R|^{np-2} + j B_R |V_R|^{nq-2} \quad (4.11)$$

Donde  $V_R$  es la tensión en un nodo cualquiera  $R$ ;  $Y_R$  es la admitancia de carga;  $G_R$  y  $B_R$  son las constantes proporcionales a la conductancia y susceptancia estimadas de los valores de prefalta; y  $np$  y  $nq$  son las constantes para las componentes activa y reactiva de la carga.

Al igual que en el numeral anterior, una alternativa para seleccionar los descriptores es la de determinar el cambio en los exponentes  $np$  y  $nq$  del modelo agregado. Este cambio ocurre a partir de una admitancia de pre falta equivalente del sistema  $Y_{eq-pre}$ , y la adición de una admitancia  $1/R_f$ , en el momento de la falta, tal como se presenta en la figura 4.8.

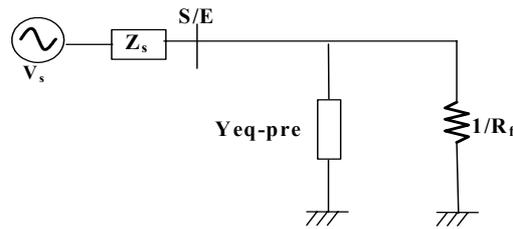


Figura 4.8: Modelo agregado de impedancias en la subestación

Debido que la falta es típicamente de carácter resistivo, se puede encontrar un modelo agregado, visto desde la subestación, para determinar la reactancia de falta, tal como se presenta en (4.12).

$$X_f = \text{Im} \left( \frac{1}{Y_{eq-pre} + 1/R_f} \right) \quad (4.12)$$

El cambio de la reactancia se puede expresar tal como se presenta en (4.13).

$$dX_f = \text{Im} \left( \frac{1}{\Delta Y_{S/E}} \right) = \text{Im} \left( \frac{1}{Y_{eq-pre} - Y_{eq-fault}} \right) \quad (4.13)$$

Para la reactancia, también se utilizan los tres valores, uno para cada fase. A partir de aquí, y en este documento, siempre que se utilice  $X_f$  se está haciendo referencia a la terna conformada por  $X_{fa}$ ,  $X_{fb}$  y  $X_{fc}$ . Éstos pueden ser fase ( $X_f$ ) y de línea ( $X_{fL}$ ).

#### 4.3.5 Frecuencia del transitorio ocasionado por la falta ( $f$ ).

En un sistema de potencia existe una relación entre los parámetros  $R$ ,  $L$  y  $C$ , y el tiempo de amortiguamiento y la frecuencia de la señal transitoria. Estas características están directamente relacionadas con la distancia desde la subestación hasta al sitio de falta. Los parámetros  $L$  y  $C$  definen la frecuencia natural de la oscilación transitoria, mientras que los parámetros  $R$ ,  $L$  y  $C$  definen el tiempo de amortiguamiento de la señal [GREE91]. Debido que la frecuencia sólo depende de  $L$  y  $C$ , éstos se pueden usar para encontrar la distancia al punto de falta, sin importar el valor de la resistencia de falta.

La propuesta se fundamenta entonces en el análisis de la teoría generalizada de transitorios en sistemas eléctricos [GREE91] y considera que la frecuencia del transitorio causado por la falta está relacionada con los parámetros  $L$  y  $C$ , del circuito desde la subestación o punto de medida hasta el sitio de falta. Para obtener el descriptor frecuencia, se emplean dos técnicas diferentes de análisis que son las transformadas de Wavelet y Fourier [GAIN04][MART02][GARC04].

Las señales transitorias de tensión son tratadas mediante la Transformada Wavelet Discreta (DWT) [GARC04]. Posteriormente, el detalle de máxima energía es transformado en el dominio del

tiempo, para ser analizado mediante la transformada rápida de Fourier FFT [OPPE97]. Como resultado, se obtienen la frecuencia del transitorio. El procedimiento se presenta en la figura 4.9 y se describe a continuación:

- La señal de tensión se mide en la subestación cuando ocurre una falta. El medidor debe estar configurado para registrar la señal desde antes de la falta, de tal manera que se tome todo el transitorio.
- La señal se analiza mediante el uso de la DWT para obtener un cierto número de detalles. La señal original de tensión se descompone en los detalles en los que se supone que existe alta probabilidad de encontrar la frecuencia del transitorio.
- El detalle con más alto índice Wavelet es el seleccionado.
- El detalle seleccionado, se convierte al dominio del tiempo mediante la utilización de la transformada inversa Wavelet IDWT.
- La frecuencia de la señal en el dominio del tiempo se obtiene mediante la aplicación de la transformada FFT, a los primeros ciclos después de la falta.

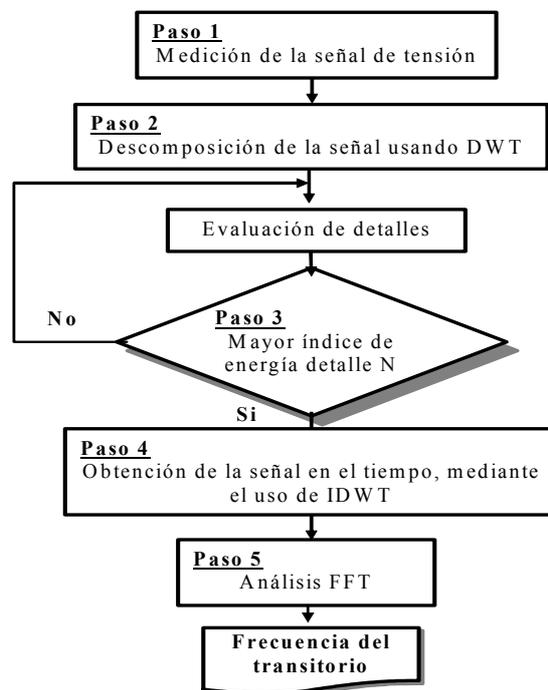


Figura 4.9. Procedimiento para obtener los descriptores asociados a la señal transitoria.

Para mostrar la aplicación del procedimiento expuesto, a continuación se presenta un ejemplo de la obtención de la frecuencia y la duración. Una falta monofásica en la fase B con una resistencia de falta de  $10\Omega$ , se simula en el nodo 5 del sistema prototipo presentado en la figura 4.1. La figura 4.10 muestra la señal de tensión transitoria en la fase B medida en la subestación.

En este caso, la señal fue muestreada con una frecuencia de 125 kHz. Como consecuencia, el detalle 1 contiene frecuencias entre 31,25 y 62,5 kHz aprox.; el detalle 2 contiene las señales con frecuencia entre 15,625 y 31,25 kHz aprox.; el detalle numero 3 corresponde a las frecuencias entre 7,812 y 15,625 kHz aprox.; el detalle 4 tiene las frecuencias entre 3,906 y 7,812 kHz aprox. y el detalle 5 tiene las frecuencias entre 1,953 y 3,906 kHz aprox.

Mediante simulaciones desarrolladas en el sistema de prueba, se determinó que los rangos de frecuencia se encuentran entre 2 y 35 kHz, por lo tanto la señal original se descompuso en cinco detalles usando la “deubichies 4” como Wavelet madre. En la figura 4.11 sólo se presentan los cuatro últimos detalles.

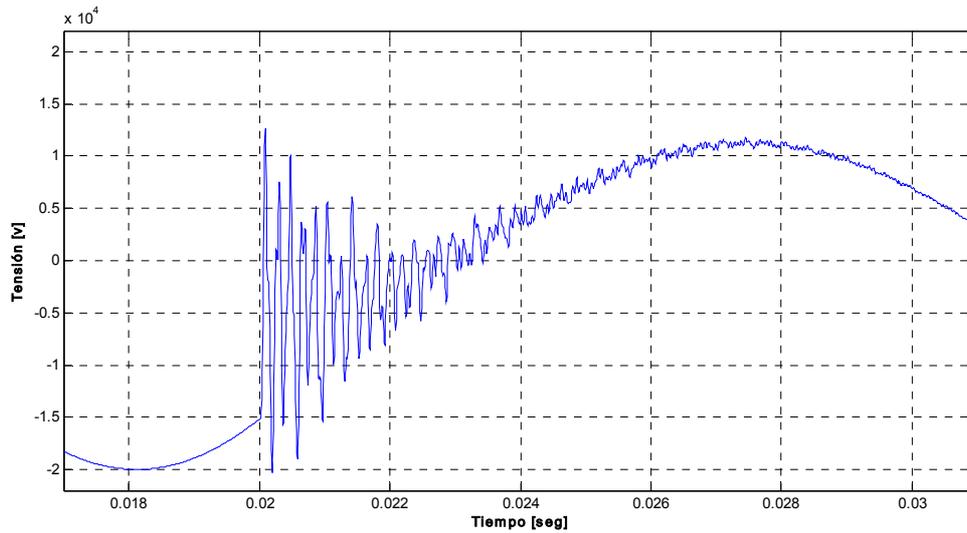


Figura 4.10. Señal transitoria de tensión en la fase B

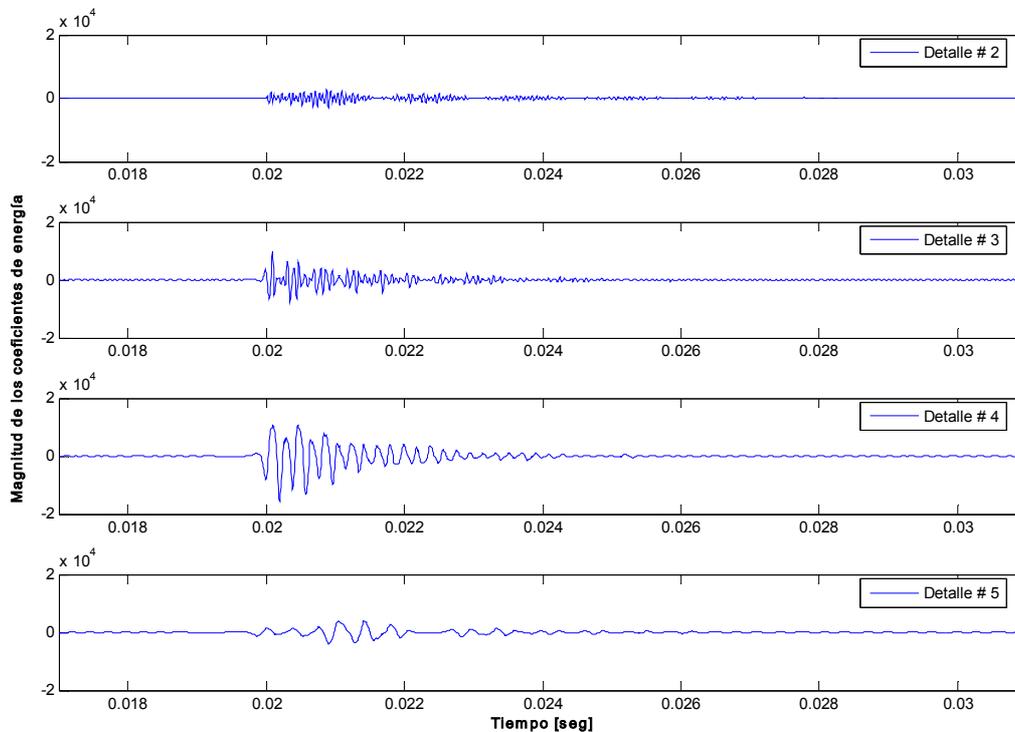


Figura 4.11. Detalles wavelet para la tensión de la fase C

De la figura 4.11 es posible observar que el detalle de más alto nivel de energía es el número 4 que corresponde al rango de las frecuencias entre 3,906 y 7,812 kHz aprox.

Esta señal se pasa nuevamente al dominio del tiempo mediante la aplicación de la transformada wavelet inversa IDWT. La señal así obtenida se analiza mediante la FFT, y los resultados se presentan en la figura 4.12. La frecuencia de la señal transitoria es 5354,6 Hz.

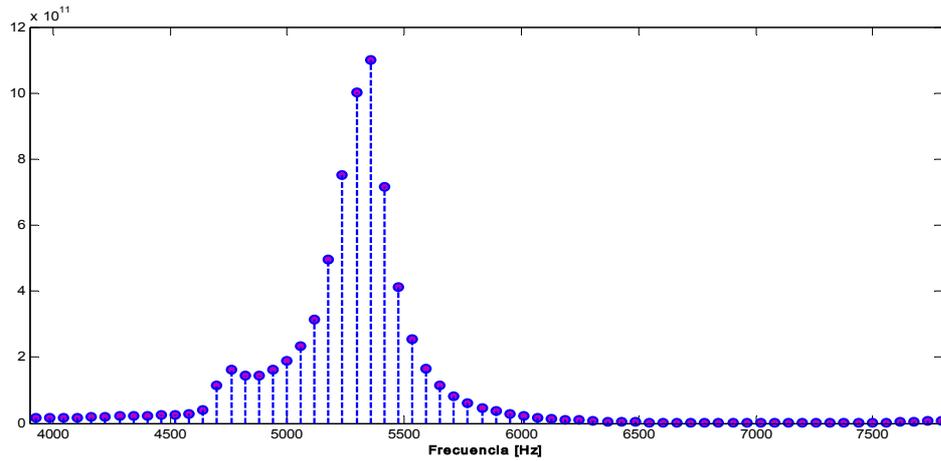


Figura 4.12. Análisis del detalle 4 mediante la FFT

Para la frecuencia se utilizan los tres valores, uno para cada fase, para prevenir errores por la presencia de transitorios con magnitud cero, debidos al ángulo de incidencia de la falta. A partir de aquí, y en este documento, siempre que se utilice  $f$  se está haciendo referencia a la terna conformada por  $f_a, f_b$  y  $f_c$ .

#### 4.3.6 Valor propio máximo de la matriz de correlación de corrientes ( $S_v$ )

Esta metodología se fundamenta en la transformación de las muestras de las tres señales de corriente de falta, utilizando la transformada de Clark. Las componentes se conocen como “Alfa” y “Beta” y “cero”, pero ésta última no será usada en esta aproximación [SOUS02] [SOUS03]. Para el caso específico de localización de faltas se utilizará la transformación de Clark-Concordia, tal como se presenta en la ecuación (4.14).

$$[Tc] = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (4.14)$$

Para la caracterización de fallos no es necesaria la utilización de la componente cero, por lo tanto únicamente se usan las primeras dos filas de la matriz  $Tc$ , tal como se presenta en (4.15).

$$\begin{bmatrix} I_\alpha \\ I_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (4.15)$$

Para interpretar el conjunto de señales de corriente  $\alpha$  y  $\beta$ , luego de aplicar la transformada  $Tc$ , se usa una aproximación basada en los valores propios. Esta técnica permite representar las corrientes de línea usando los valores propios de las muestras de datos de una matriz de correlación  $B$ . De esta forma, a partir de las corrientes de línea, transformadas como se presenta en (4.16) se obtiene la matriz  $A$ . El número de muestras significativas corresponde al número de filas de  $A$ , donde  $t_0$  es el tiempo inicial de la muestra de datos y  $\Delta t$  es el intervalo de muestreo.

$$A = \begin{bmatrix} i_\alpha(t_0) & i_\beta(t_0) \\ i_\alpha(t_0 + \Delta t) & i_\beta(t_0 + \Delta t) \\ \dots & \dots \\ i_\alpha(t_0 + (n-1)\Delta t) & i_\beta(t_0 + (n-1)\Delta t) \end{bmatrix} \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (4.16)$$

La matriz de correlación se obtiene aplicando la ecuación (4.17)

$$B = A^T \cdot A \quad (4.17)$$

Los valores propios se obtienen mediante la aplicación de Matlab® “*eig(B)*”, a partir del cual se obtienen las matrices  $V$  y  $D$  como se presenta en (4.18).

$$[V, D] = \text{eig}(B) \quad (4.18)$$

La matriz  $D$  contiene los valores propios de las dos componentes de corriente, tal como se presenta en (4.19).

$$D = \begin{bmatrix} \lambda_\alpha & 0 \\ 0 & \lambda_\beta \end{bmatrix} \quad (4.19)$$

De estos dos valores, únicamente el mayor valor de  $\lambda$  se utiliza para la definición de la distancia de la falta. A partir del valor propio máximo ( $Sv$ ), es posible encontrar una relación con la distancia a la falta  $m$ . En [SOUS02] se plantea una relación lineal tal como se presenta en la ecuación (4.20)

$$m = a(Sv)^k + b \quad (4.20)$$

Los términos  $a$  y  $b$  dependen del tipo de carga de prefalta. La constante  $k$  es conocida como un parámetro de falta, el cual según el tipo de falta tiene los siguientes valores ( $k=1$  para fase-tierra,  $k=-0,55$  para fase-fase-tierra,  $k=-0,552$  y  $k=0,551$  para falta trifásica) [SOUS02][SOUS03].

Según lo expuesto, es posible relacionar el valor propio máximo ( $Sv$ ), con la distancia a la cual ocurre la falta, por lo tanto este es otro descriptor útil para ser usado en un método basado en el conocimiento.

Este descriptor es único ya que para su obtención se requieren las señales de corriente de las tres fases.

#### 4.3.7 Firma del sistema

Lo que en este documento se define como firma del sistema, está asociada al registro de tensión y especialmente de corriente, donde se registran las características de activación de los elementos de protección del sistema de distribución. Fundamentalmente, un sistema de distribución utiliza reconectadores, seccionalizadores, relés de sobrecorriente de relación inversa tiempo corriente (51) y relés instantáneos (50), éstos dos últimos asociados al interruptor principal del alimentador [ZHU97][AGGA97][PHAD95]. El desempeño de estos elementos de protección depende de la configuración de los mismos y es algo característico de cada sistema de distribución, y por tanto es lo que se conoce como la firma o huella del sistema.

Para el circuito de la figura 4.1 un reconectador automático ( $R$ ), está configurado para abrir y recerrar el circuito ante faltas aguas abajo del circuito, de forma coordinada con los fusibles ( $F$ ). La secuencia puede contener uno o más disparos rápidos y recierres para despejar faltas transitorias y restaurar el servicio. En caso que la falta persista, los fusibles retiran las faltas permanentes que ocurren aguas debajo de su ubicación [PADK95]. Este tipo de coordinación se denomina “coordinación en salvamento de fusibles”, y sus curvas básicas de relación inversa tiempo-corriente, se presentan en la figura 4.13.

En la figura 4.14 se presenta la corriente de línea tal como es registrada en la subestación de distribución. La grafica corresponde a la corriente de la fase C debida a la falta de  $30\Omega$  en el nodo 25 del sistema de prueba de la figura 4.1. En este caso el reconectador ( $Rec$ ), tiene dos curvas rápidas y una curva lenta. La corriente de falta se interrumpe dos veces por las curvas rápidas del reconectador y finalmente como consecuencia de la falta permanente, el fusible opera antes que el reconectador abra nuevamente el circuito en la primera operación asociada a su curva lenta.

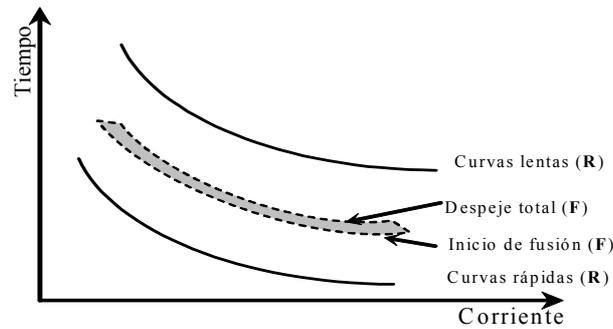


Figura 4.13. Curvas de coordinación de tiempo típicas del esquema de salvamento de fusibles.

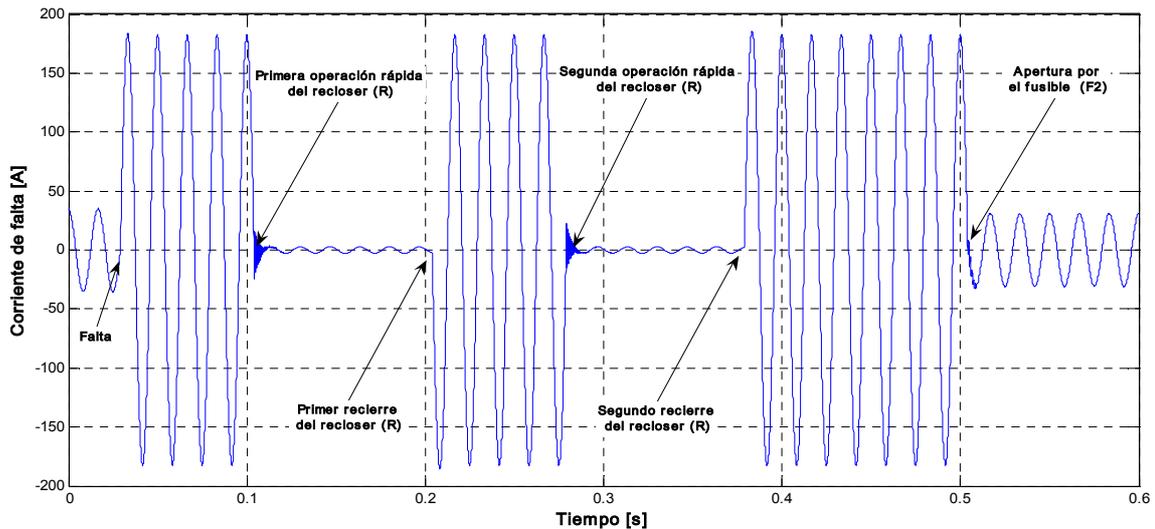


Figura 4.14. Corrientes de línea medidas en la subestación durante una falta en el nodo 25 de la figura 4.1

Hay varias opciones para obtener los descriptores de la firma del sistema y posteriormente relacionarlos con la localización de la falta. En este caso, los usados para determinar la zona en falta son los tiempos de reconexión después de la acción de las curvas rápidas (100 ms aprox.) y retrasadas (200 ms aprox). El cambio de la magnitud entre la corriente de prefalta y después que el reconectador se ha activado (45-7 [A] para la figura 4.14), también se puede emplear en la localización. Además también se puede obtener el cambio de la magnitud de corriente después de la acción del fusible (45-37 [A] para la figura 4.14). La magnitud de la corriente y la duración de la falta también pueden ser usadas si se conocen las curvas de relación inversa tiempo corriente.

Finalmente, es muy importante resaltar que el tiempo de recierre de cada reconectador puede conocerse por los parámetros de configuración del mismo y puede permitir la localización geográfica de la falta. Una adecuada configuración de este parámetro, puede ayudar a los equipos de mantenimiento de la subestación a decidir donde está la falta, reduciendo de esta manera el problema de la múltiple estimación.

Los descriptores asociados a la firma del sistema, permiten seccionar el sistema para identificar la zona probable de falta, mediante una serie de reglas *if-then*. Los sitios de seccionamiento corresponderían a la ubicación de los elementos de protección. Este descriptor no está aplicado y desarrollado en esta tesis, y sólo se presenta como una referencia ya que es citado por otros autores como [KRIS89] [ZHU97]. Como parte del trabajo exploratorio, el autor presenta una aplicación en este aspecto en particular, en [MORA06-b] y otra en [MORA06-e].

#### 4.4 Variación de los descriptores propuestos como entrada a los MCBC

Para la obtención de la base de faltas se realizaron simulaciones en cada uno de los nodos, usando 21 resistencias de falta con valores entre  $0,05\Omega$  y  $40\Omega$  [DAGE00]. Se simularon 10 tipos de faltas, tres monofásicas, tres bifásicas, tres bifásicas a tierra y una trifásica. En total para el sistema de prueba de la figura 4.2, se realizaron 2982 faltas.

Para este caso particular se realizó una simulación extensiva de faltas usando Matlab® y “Alternative Transients Program –ATP” [EMTP87]. La base de faltas, se obtuvo mediante la simulación automatizada de las mismas, tal como se propone en [MORA06-c].

Para el análisis se presentan las gráficas del comportamiento de los descriptores significativos a lo largo de un alimentador, de tal forma que se muestra su relación con la localización y la resistencia de falta. En este caso, y a manera de ejemplo se presentan los valores de fase, para falta monofásica de la fase A, considerando 21 valores diferentes de resistencia de falta desde  $0,05\Omega$  hasta  $40\Omega$ , para un alimentador, que en este caso se selecciona entre los nodos 1 y 12. En los otros alimentadores y para los otros tipos de falta el comportamiento es similar. En las figuras 4.15 a 4.20, cada punto corresponde a una falta, de las  $21 \times 12 = 256$  faltas.

La variación de la tensión ( $dVa$ ), tiene un comportamiento fuertemente influenciado por la resistencia de falta y la distancia a la falta (desde el punto de medida), tal como se presenta en la figura 4.15.

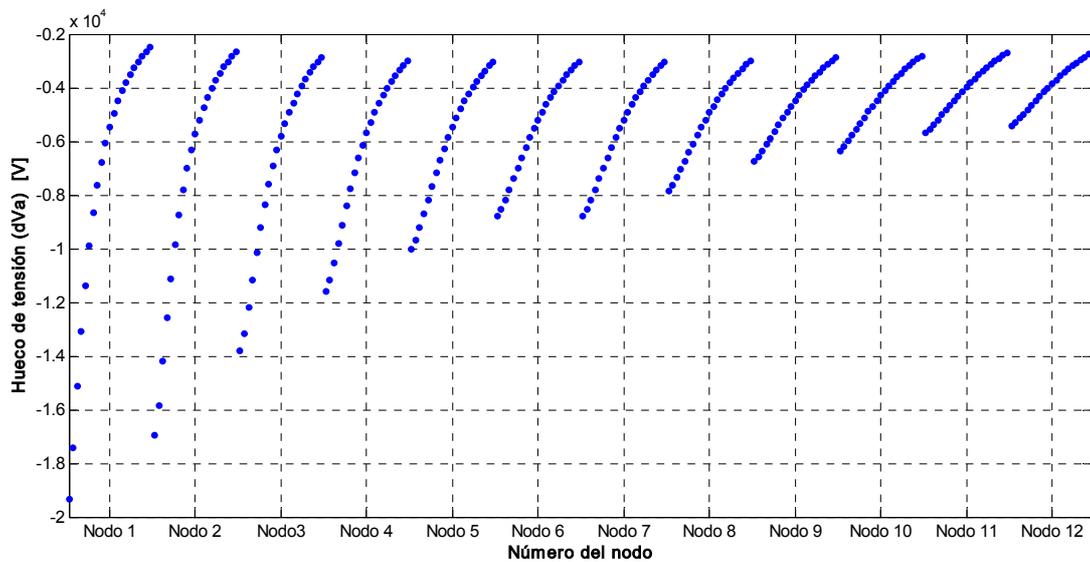


Figura 4.15. Variación del valor eficaz de la tensión entre estado estable de falta y prefalta. Hueco de tensión

Al igual que la variación de la tensión, la variación de la corriente ( $dI$ ) y la del valor propio máximo ( $Sv$ ), muestran una alta influencia de la resistencia de falta y la distancia al punto de medida. Este comportamiento se muestra en las figuras 4.16 y 4.17, para  $dI$  y  $Sv$ , respectivamente.

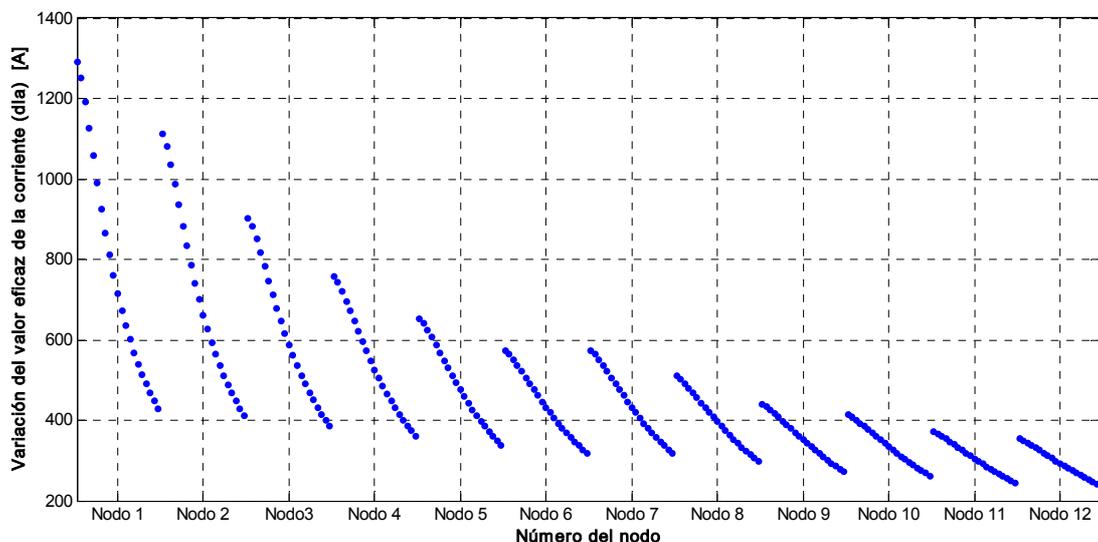


Figura 4.16. Variación del valor eficaz de la corriente entre los estados estables de falta y prefalta

Los valores máximos de variación de  $S_v$  pueden alcanzar la magnitud de  $3.5 \times 10^7$ , tal como se presenta en la figura 4.17, para el caso de faltas en el nodo 1, con valores de resistencia de  $0,05\Omega$  y  $40\Omega$ .

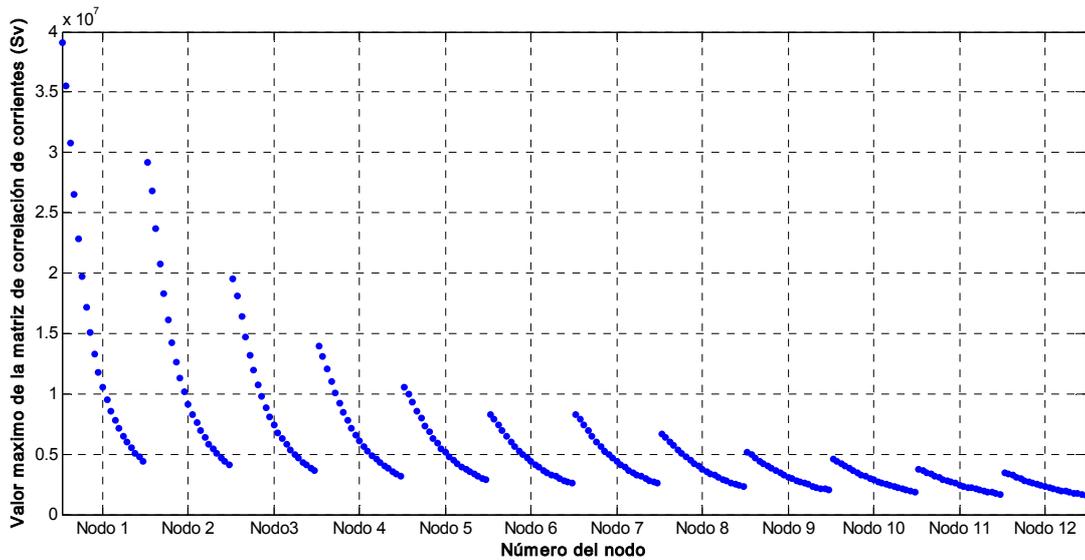


Figura 4.17. Valor propio máximo de la transformada Clark-Concordia de las señales de corriente durante la falta

La variación de la potencia aparente ( $dS$ ), no está tan influenciada por la variación de la resistencia de falta, dado que se produce un efecto de amortiguamiento por el filtro introducido por el producto entre la tensión y la corriente, tal como se presenta en la figura 4.18. Este comportamiento del descriptor hace que sea considerado como una buena entrada que le permite al localizador diferenciar entre una falta en un nodo u otro.

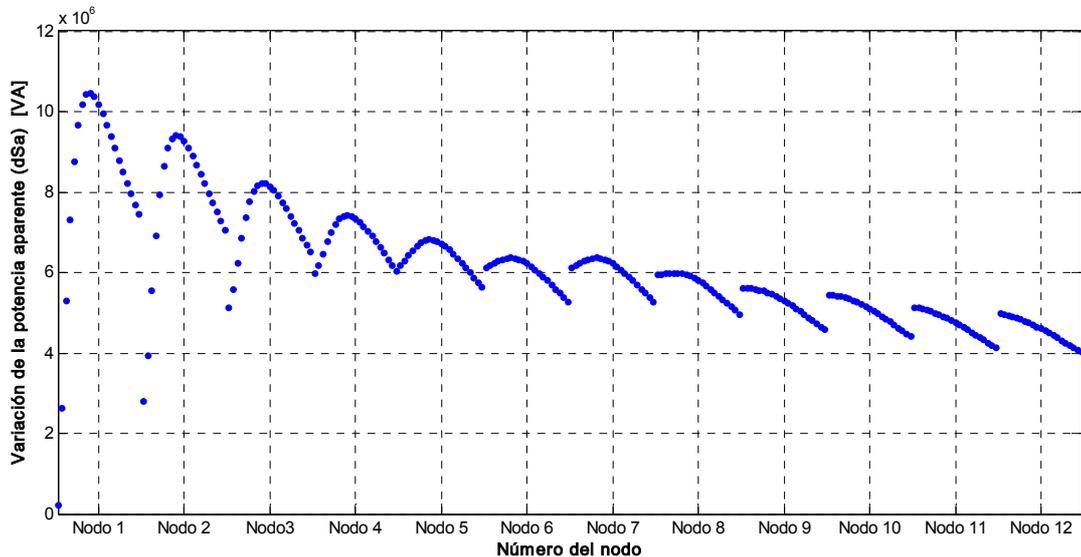


Figura 4.18. Variación de la potencia aparente entre los estados estables de falta y prefalta

Finalmente, la frecuencia del transitorio oscilante causado durante la falta ( $f$ ), y la reactancia de fase ( $X_f$ ) para el caso falta monofásica, muestran un comportamiento relativamente constante con la variación del valor de la resistencia de falta, lo cual es muy bueno para propósitos de localización. Con respecto a la distancia, se muestra una variación que permite diferenciar los diferentes nodos en los que ocurrió la falta, tal como se muestra en las figuras 4.19 y 4.20 para reactancia y frecuencia, respectivamente.

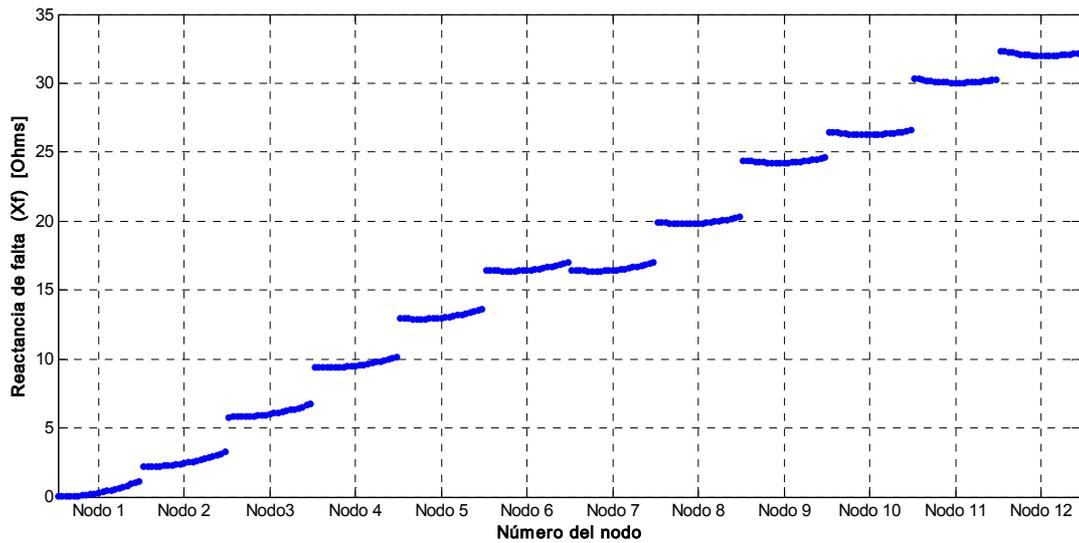


Figura 4.19. Reactancia de la fase en falta

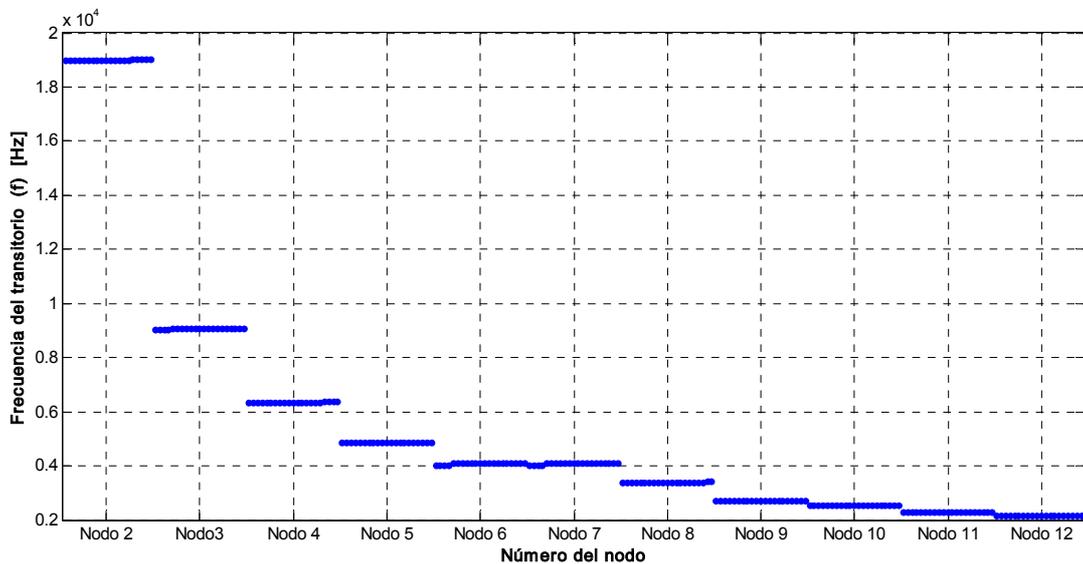


Figura 4.20. Frecuencia del transitorio medida en la fase en falta

#### 4.5 Conclusiones

En este capítulo se presentó la selección de descriptores para ser utilizados como información relevante para alimentar métodos de clasificación útiles en el problema de localización de faltas.

Algunos de los descriptores utilizados como la frecuencia del transitorio, el análisis de matriz de correlación de la corriente y la variación de la admitancia están directamente asociados con la distancia eléctrica desde la subestación hasta la falta. Un clasificador entrenado con estos descriptores, brinda soporte a los métodos basados en el modelo, que son clásicos para localización de faltas.

Otros descriptores como el análisis de la firma del sistema, la variación de carga de pre-falta, falta y post operación del sistema de protecciones, ayudan a localizar geográficamente la zona de falta, mediante la identificación de la carga que salió del servicio o la configuración del elemento de protección activado.

Según el análisis gráfico de los descriptores, se aprecia que éstos cambian de magnitud a medida que se varía la distancia desde la subestación hasta el nodo en falta. A partir de este comportamiento, se puede predecir que las técnicas de clasificación ofrecerán buenos resultados para identificar adecuadamente la zona en falta.

## 5 Aplicación de los MBM y MCBC para localización de faltas en sistemas de distribución

### 5.1 Introducción

La aplicación de los métodos basados en el modelo (MBM), como se presentó en el capítulo dos, requiere del conocimiento de los parámetros del sistema. De otra parte, la aplicación de los métodos de clasificación basados en el conocimiento (MCBC) presentados en el capítulo tres, requiere de la caracterización de las señales de tensión y corriente, tal como se presenta en el capítulo cuatro.

Mientras que los MBM localizan la distancia eléctrica desde el punto de medida hasta el sitio de falta, los MCBC localizan una zona en la cual se encuentra la falta. Mediante la combinación de estas dos metodologías, se obtienen los algoritmos híbridos que se presentan en el siguiente capítulo.

En este capítulo, se presentan inicialmente las pruebas de los métodos basados en el modelo presentados en el capítulo dos. Posteriormente, se presentan algunas pruebas comparativas realizadas con el MCBC LAMDA y SVM, haciendo énfasis en los aportes sobre los MBM. En la sección siguiente se presenta una aplicación más detallada de las SVM, al problema de localización y finalmente, se presentan las conclusiones del capítulo.

### 5.2 Localización de faltas mediante los Métodos Basados en el Modelo - MBM

Inicialmente, en este capítulo, se presentan las pruebas realizadas a los métodos que se basan en el modelo del sistema de distribución, para la localización de faltas. Estos métodos se fundamentan en la estimación de la impedancia a partir de las medidas de tensión y corriente registradas en la subestación.

A manera de resumen, el uso de tensiones y corrientes para propósitos de localización de faltas fue descrito inicialmente como el método de la componente reactiva [WARR68]. Un modelo simplificado del alimentador, y equivalente a un divisor de tensión es usado en [BOLL00] para explicar la dependencia entre la magnitud del hueco de tensión y la distancia a la falta. Adicionalmente, nueve diferentes métodos son también analizados para considerar la complejidad introducida por la topología radial del sistema, la presencia de múltiples calibres de conductor, la presencia de cargas intermedias y ramales laterales monofásicos. Como factor adicional, el problema de localización de faltas también debe tener en cuenta la incertidumbre asociada al valor de la resistencia de falta.

#### 5.2.1 Sistema de prueba y escenarios

El sistema de prueba, es un circuito de 25kV de Saskatoon Power and Light de la ciudad de Saskatoon, Canadá presentado en [DAS98], y frecuentemente citado y utilizado por otros investigadores del problema de la localización de faltas. En la figura 5.1 se presenta el diagrama unifilar del sistema prototipo.

Los escenarios utilizados para las pruebas, incluyen la simulación de faltas en 10 nodos, desde el 2 hasta el 12 (considerando que desde el punto de vista circuital, los nodos 6 y 7 corresponden al mismo valor de impedancia). Las resistencias de falta utilizadas corresponden a valores de 0,5  $\Omega$ , 5  $\Omega$ , y 25  $\Omega$ , valores que son representativos de los comúnmente usados en este tipo de pruebas [DAS98].

Para las situaciones descritas, se realizaron las pruebas que permiten determinar el desempeño de cada uno de los métodos analizados, para faltas en el alimentador principal y ante tres valores de resistencia de falta. En aquellos métodos donde se considera el acoplamiento mutuo, éste fue despreciado. Estas pruebas están orientadas a la comparación del desempeño de cada método, en cuanto a los resultados obtenidos en la distancia de falta.

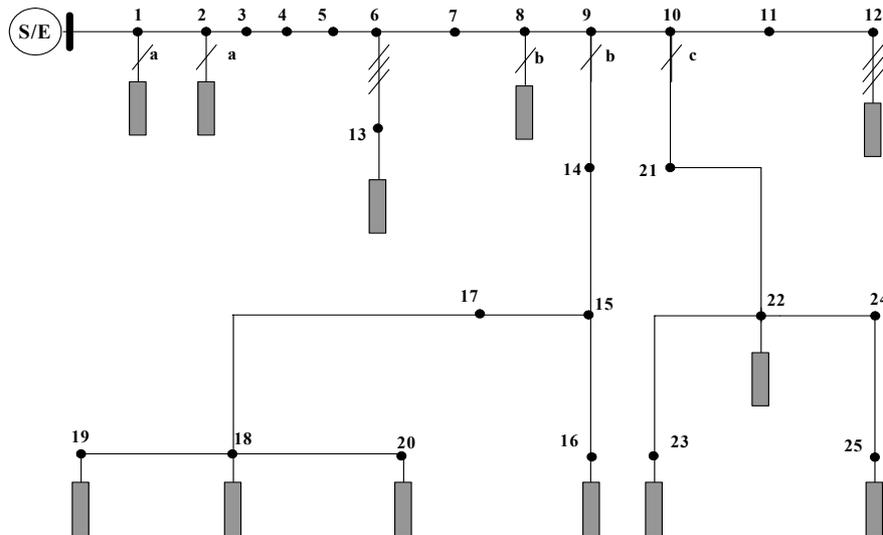


Figura 5.1. Sistema prototipo de prueba. Circuito de 25kV de Saskatoon Power and Light de la ciudad de Saskatoon, Canadá.

### 5.2.2 Resultados en la estimación de la distancia de falta

Los métodos que se presentan en el capítulo dos fueron probados para faltas entre fases y fase tierra, que involucran una, dos y tres fases. A partir de estos resultados de pruebas y de la tabla de comparaciones cualitativas presentada en el capítulo dos, se selecciona un método que servirá para las implementaciones híbridas, del capítulo 6.

En las tablas 5.1, 5.2 y 5.3 se presentan las distancias a la falta obtenidas por cada uno de los métodos en caso de faltas monofásicas, y en el anexo C se presentan las tablas para las demás faltas.

Los resultados se presentan para los nodos desde el 2 hasta el 12 del circuito de la figura 5.1, para observar al comportamiento de cada uno de los métodos en un mismo alimentador. Lo que se pretende con estas pruebas es observar la variación en la estimación de la localización de la falta, a medida que aumenta la distancia y ante tres valores de resistencia de falta  $0,05\Omega$ ,  $5\Omega$  y  $25\Omega$ . El análisis del desempeño de los métodos para las los cuatro tipos de faltas, se presentan de manera gráfica en la siguiente sección.

En la parte superior de cada tabla se presenta la distancia real a la cual ocurrió la falta.

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
A. Warrington [WARR68].	2,3984	6,367	10,296	14,185	18,035	21,847	26,666	28,884	33,000	35,16
K. Srinivasan et al. [SRIN89]	2,4138	6,438	10,462	14,486	18,508	22,53	27,684	30,096	34,596	36,998
A. Girgins et al [GIR93]	2,4138	6,4384	10,462	14,486	18,51	22,531	27,684	30,092	34,616	36,992
J. Zhu et al [ZHU97]	2,4203	6,4513	10,482	14,513	18,539	22,547	27,701	30,09	34,576	36,585
R. Aggarwal et al [AGGA97]	2,4453	6,4467	10,485	14,487	19,192	23,119	28,158	30,085	37,013	37,013
R. Das [DAS98]	2,410	6,430	10,451	14,474	18,501	22,525	27,675	30,090	34,566	37,006
D. Novosel [NOVO98]	2,3921	6,3781	10,36	14,337	18,311	22,281	27,481	29,906	34,443	36,824
L. Yang [YANG98]	2,3909	6,3732	10,35	14,322	18,288	22,251	27,463	29,895	34,449	36,826
Saha et al [SAHA02]	2,4513	6,5555	10,657	14,755	18,85	22,934	27,919	30,265	34,73	37,013
M. Choi et al [CHOI04]	2,4131	6,4372	10,461	14,484	18,507	22,528	27,683	30,096	34,595	37,003

Tabla 5.1. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta  $R_f=0,05\Omega$

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
<b>A. Warrington</b> [WARR68].	2,410	6,317	9,719	14,138	17,970	21,480	26,552	28,760	32,858	34,998
<b>K. Srinivasan et al.</b> [SRIN89]	2,473	6,473	10,007	14,600	18,632	22,400	28,034	30,229	34,755	37,162
<b>A. Girgins et al</b> [GIRG93]	2,491	6,520	10,088	14,703	18,775	22,556	27,996	30,374	34,904	37,317
<b>J. Zhu et al</b> [ZHU97]	2,626	6,812	10,571	15,483	19,781	23,904	29,880	32,673	37,029	37,270
<b>R. Aggarwal et al</b> [AGGA97]	2,445	6,447	10,485	14,487	19,155	23,119	28,121	30,085	37,013	37,013
<b>R. Das</b> [DAS98]	2,434	6,447	10,462	14,447	18,539	22,562	27,704	30,135	34,611	37,047
<b>D. Novosel</b> [NOVO98]	2,395	6,334	9,799	14,327	18,296	21,975	27,435	29,853	34,382	36,754
<b>L. Yang</b> [YANG98]	2,447	6,395	9,872	14,397	18,371	22,063	27,542	29,975	34,534	36,908
<b>Saha et al</b> [SAHA02]	1,344	5,415	8,862	13,711	17,830	21,566	26,996	29,367	33,901	36,461
<b>M. Choi et al</b> [CHOI04]	2,411	6,392	9,899	14,489	18,517	22,256	27,706	30,124	34,642	37,045

Tabla 5.2. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta  $R_f=5\Omega$ 

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
<b>A. Warrington</b> [WARR68].	2,850	6,725	10,558	14,354	18,113	21,830	26,515	28,683	32,695	34,793
<b>K. Srinivasan et al.</b> [SRIN89]	2,736	6,820	10,942	15,015	19,069	23,375	29,634	30,707	35,289	37,759
<b>A. Girgins et al</b> [GIR93]	2,816	7,088	11,346	15,607	19,873	23,975	29,278	31,389	36,144	38,649
<b>J. Zhu et al</b> [ZHU97]	1,521	6,112	10,714	15,514	20,078	24,875	30,991	34,126	35,564	38,044
<b>R. Aggarwal et al</b> [AGGA97]	1,260	6,447	10,485	14,487	19,155	23,082	28,158	30,085	37,013	37,013
<b>R. Das</b> [DAS98]	2,417	6,543	10,471	14,426	18,381	22,479	27,707	30,075	34,622	37,074
<b>D. Novosel</b> [NOVO98]	2,423	6,369	10,328	14,280	18,227	22,169	27,269	29,662	34,135	36,497
<b>L. Yang</b> [YANG98]	2,636	6,670	10,678	14,681	18,677	22,668	27,857	30,293	34,800	37,244
<b>Saha et al</b> [SAHA02]	-2,630	1,560	5,766	9,980	14,195	18,402	23,729	26,168	30,870	33,544
<b>M. Choi et al</b> [CHOI04]	2,320	6,353	10,410	14,480	18,601	22,689	27,941	30,405	34,984	37,421

Tabla 5.3. Resultados de las pruebas –Falta monofásica a través de una resistencia de falta  $R_f=25\Omega$ 

### 5.2.3 Análisis comparativo de resultados

Como parámetros de comparación de los métodos se han utilizado la estimación del error en la distancia a la falta, su dependencia respecto a la localización de la misma y la resistencia de falta. El error en la estimación de la distancia se calcula usando la ecuación (5.1)

$$e [\%] = \left| \frac{\text{localización estimada} - \text{localización real}}{\text{Longitud total del alimentador}} \right| \times 100 \quad (5.1)$$

En caso de faltas monofásicas y según las tablas 5.1, 5.2 y 5.3, los límites para la estimación de los errores son -13,63% y 10,89%. Para una adecuada visualización en las gráficas, únicamente las cuatro técnicas de localización con el error más bajo en la estimación, se presentan en las figuras 5.2 y 5.3, para valores de resistencia de falta de  $0,05\Omega$  y  $25\Omega$ , respectivamente.

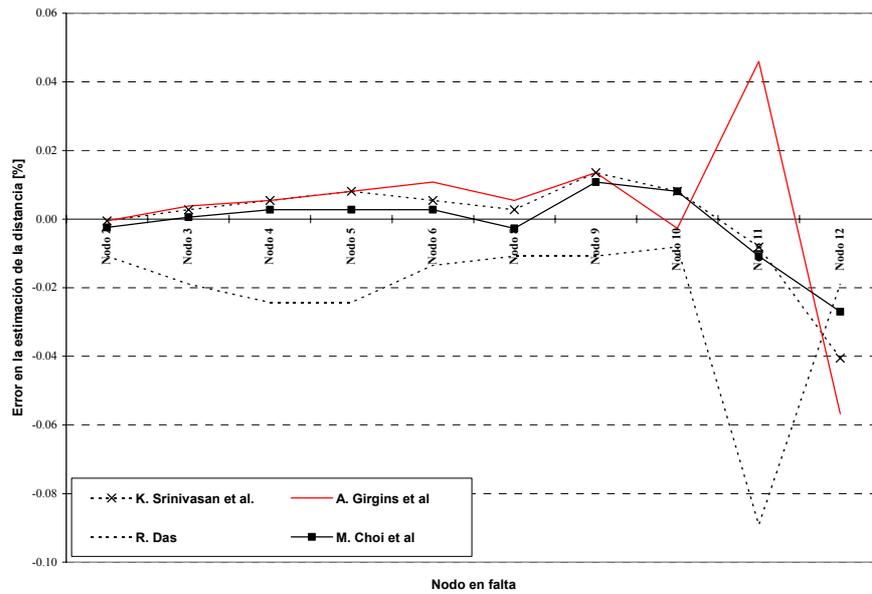


Figura 5.2: Mejores resultados en la estimación de la distancia a la falta. Falta monofásica con RF= 0,05 Ω

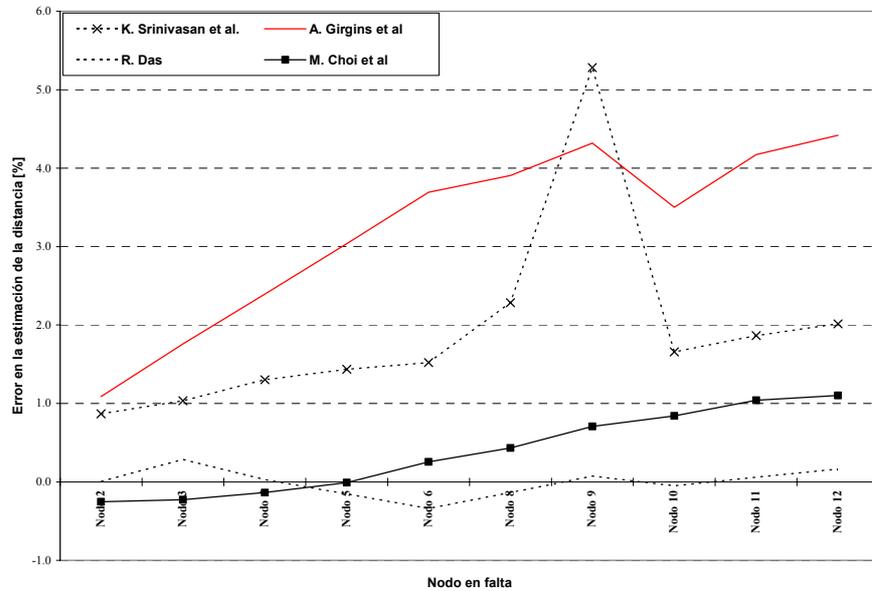


Figura 5.3. Mejores resultados en la estimación de la distancia a la falta. Falta monofásica con RF= 25 Ω

En la prueba de los métodos ante faltas de fase, se observa que el propuesto por R. Aggarwal [AGGA97] presenta la peor estimación (error del 89%), debido a que se fundamenta en la componente superimpuesta de secuencia cero, la cual no está presente en esta tipo de falta. Los errores de estimación de los otros métodos analizados oscilan entre -3,61% y 6,45%. En las figuras 5.4 y 5.5 se presentan los resultados comparativos de las mejores cuatro técnicas para valores de resistencia de falta de 0,05Ω y 25Ω, respectivamente.

De otra parte, los métodos propuestos por K. Srinivasan et al. [SRIN89], J. Zhu et al [ZHU97] y M. Choi et al [CHOI04], no están planteados para determinar la distancia a la falta en caso de faltas bifásicas a tierra. Para los demás métodos, el error en la estimación está entre -1,11% y 6,25%. En las figuras 5.6 y 5.7 se presentan los resultados comparativos para las cuatro técnicas que presentan mejores resultados cuando se simulan faltas bifásicas a tierra con valores de resistencia de falta de 0,05Ω y 25Ω, respectivamente.

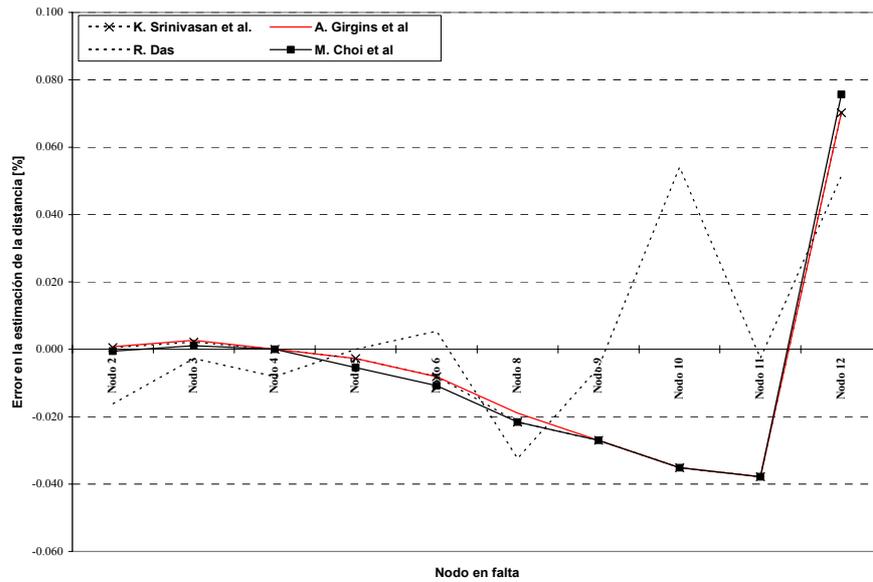


Figura 5.4. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica con  $RF=0,05 \Omega$

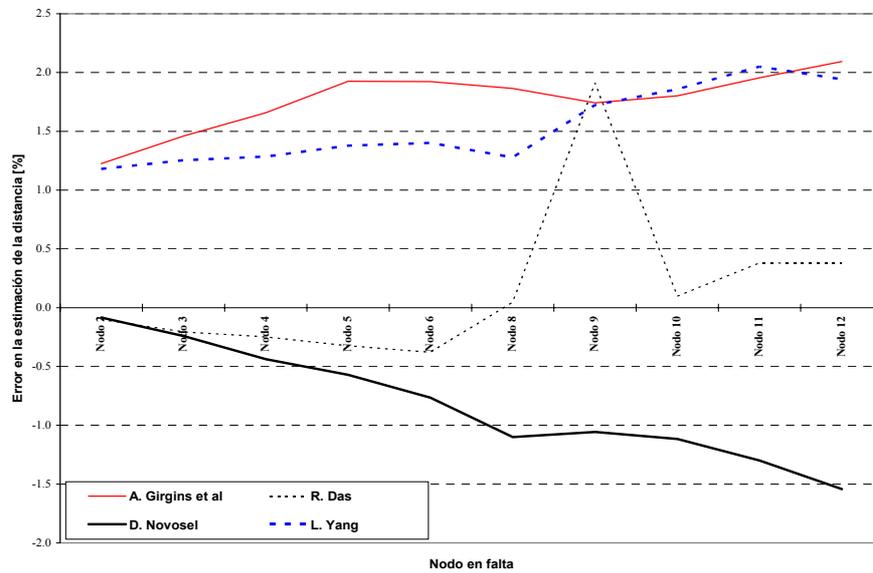


Figura 5.5. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica con  $RF=25 \Omega$

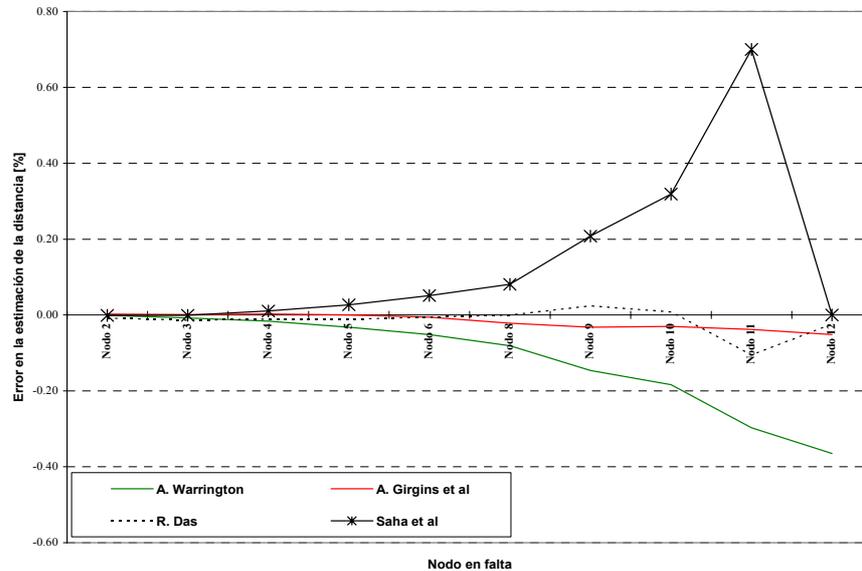


Figura 5.6. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica a tierra con  $RF=0,5 \Omega$

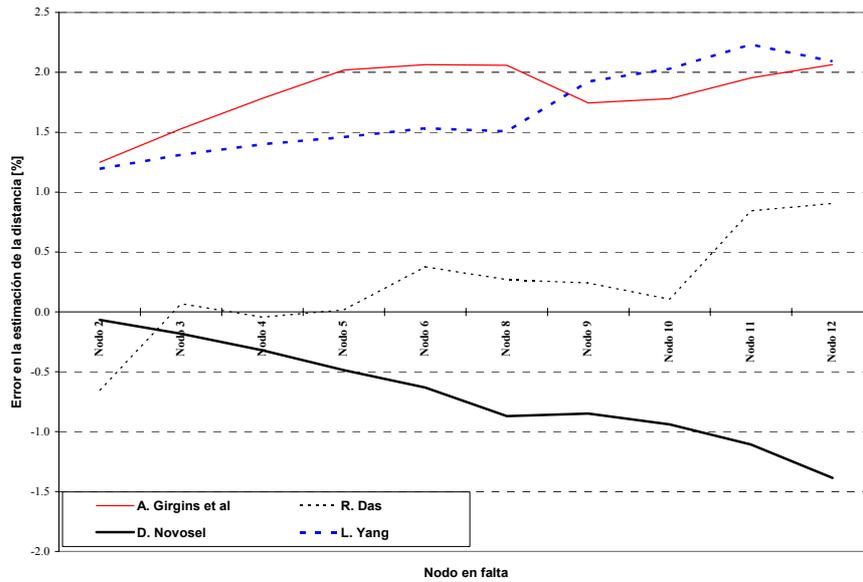


Figura 5.7. Mejores resultados en la estimación de la distancia a la falta. Falta bifásica a tierra con  $RF=25 \Omega$ . Finalmente, en el caso de las faltas trifásicas, los métodos propuestos por A. Girgins et al [GIR93], J. Zhu et al [ZHU97], R. Aggarwal et al [AGGA97] y M. Choi et al [CHOI04], no están deducidos para este tipo de faltas.

Las pruebas desarrolladas usando los restantes métodos presentan errores de estimación que oscilan entre -1,48% y 1,97%. En las figuras 5.8 y 5.9 se presentan los resultados comparativos para las cuatro técnicas que mejor estiman la distancia a la falta, para valores de resistencia de falta de  $0,05\Omega$  y  $25\Omega$ , respectivamente.

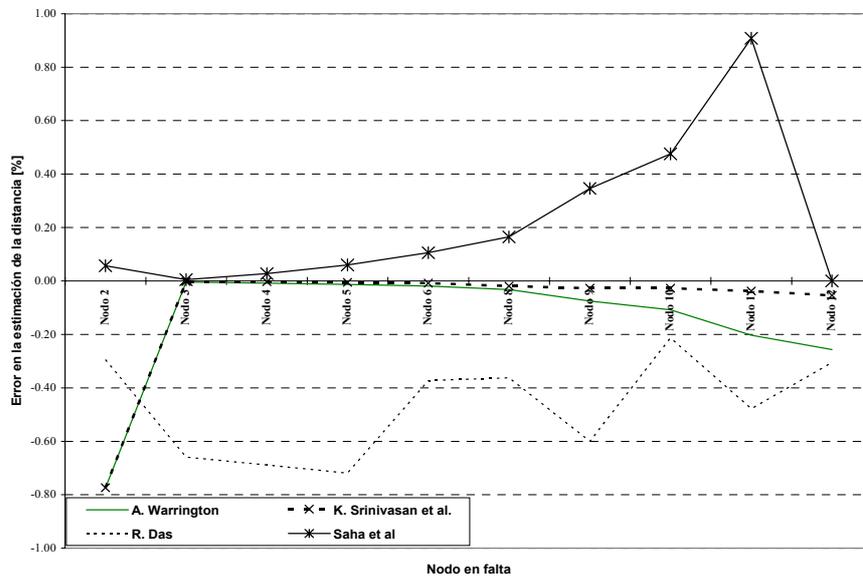


Figura 5.8. Mejores resultados en la estimación de la distancia a la falta. Falta trifásica con  $RF=0,5 \Omega$

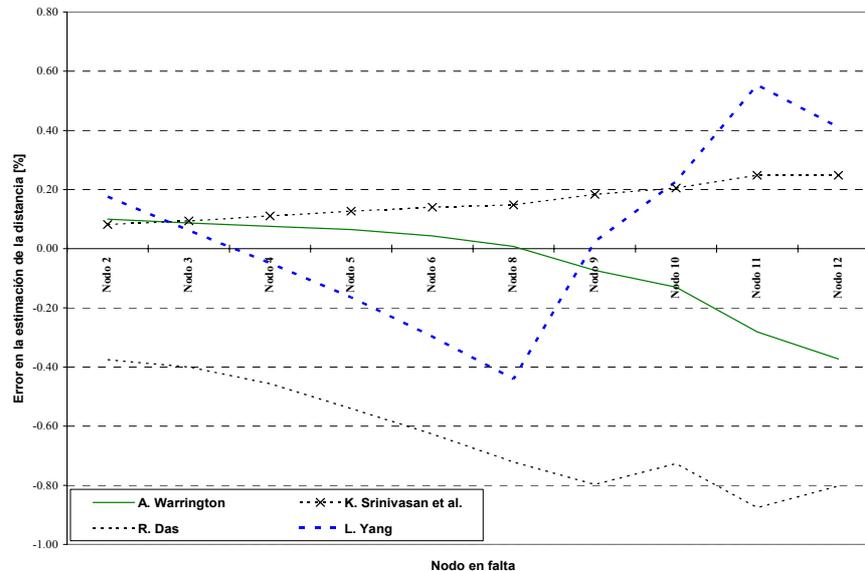


Figura 5.9. Mejores resultados en la estimación de la distancia a la falta. Falta trifásica con  $RF=25\ \Omega$

Dos aspectos importantes asociados con la localización de faltas son: la resistencia de la falta y la distancia desde el punto de medición hasta la falta. A partir de los resultados comparativos presentados, se comprueba que a medida que crece el valor de la resistencia de falta o la distancia al punto de medición, mayor es el error en la estimación, esto es mayor es la dificultad de localizar precisamente la falta ocurrida.

Los mayores errores se obtuvieron en las pruebas de los métodos ante faltas monofásicas. Para este tipo de faltas, los métodos que presentan mejores resultados son los propuestos por R. Das [DAS98] y M. Choi *et al* [CHOI04].

En el caso de faltas que involucran dos fases, los métodos propuestos por A. Girgins *et al* [GIRG93] y R. Das [DAS98] muestran el mejor desempeño.

Para el caso de las faltas trifásicas, los métodos propuestos por R. Das [DAS98], Warrington [WARR68] y K. Srinivasan *et al* [SRIN89], están incluidos dentro de los métodos que presentan mejores resultados tanto para faltas de baja como de alta resistencia.

Finalmente, como resultado de las pruebas se encuentra que los errores obtenidos para el caso de faltas monofásicas son siempre mayores que los obtenidos para otros tipos de faltas. El método con el mejor desempeño global es el propuesto por R Das [DAS98], sin embargo métodos simples como los propuestos por Warrington [DAS98] y D. Novosel [NOVO98] son también adecuados para dar una buena aproximación al problema de localización de falta, en el sistema prototipo utilizado en las pruebas.

### 5.3 Localización de faltas mediante los Métodos de Clasificación basados en el Conocimiento - MCBC

La localización de faltas mediante MCBC, se fundamenta en la utilización de métodos de clasificación que aprovechan la información relevante, obtenida de las medidas disponibles y que esté relacionada directamente con la localización de la falta.

En este numeral, inicialmente se presenta la metodología general utilizada y luego las pruebas preliminares para determinar el desempeño de cada técnica (LAMDA y SVM), usando el sistema prototipo. Las respuestas de esta prueba inicial sirven para descartar una técnica y profundizar en el análisis de la otra.

#### 5.3.1 Metodología para localizar la zona bajo falta usando MCBC

En este numeral se presenta la metodología utilizada para localizar la región probable de falta y la posterior aplicación de los MCBC. El proceso parte de la zonificación de la red, la adquisición de

la base de datos de falta, el pre-procesamiento de la señal, el entrenamiento del MCBC y la prueba de precisión del MBC. Estas cinco etapas generales que se explican a continuación.

### 5.3.1.1 Zonificación de la red

Como se está utilizando un método de clasificación, el problema es más sencillo si el sistema de potencia se divide en zonas (en esta tesis, las clases corresponden a zonas del alimentador), a las cuales se asigna un conjunto de descriptores de entrada.

El criterio de zonificación debe estar fundamentado en la división del sistema de potencia, de tal manera que la identificación de una zona como la más probable de falta, permita al equipo de mantenimiento de la red, reducir el tiempo de localización de la falta. Como recomendación básica, una zona no debe contener más de un lateral con las mismas fases, para evitar el problema de la múltiple estimación presentada por los MBM.

Los criterios de zonificación por tanto son la topología de la red, la localización de las protecciones, la longitud de los alimentadores, la importancia del ramal, la disponibilidad de datos de faltas para entrenar el MCBC y finalmente las prácticas de la empresa.

Respecto al último aspecto, asociado a los procedimientos bajo falta, algunas veces se requiere localizar una pequeña zona a la cual se deben enviar el equipo que repare la falta lo antes posible. En otros casos, únicamente con la definición de una zona grande, limitada por los equipos de seccionamiento, es suficiente para aislar la zona y alimentar el circuito que no está en falta, a través de operaciones de cierre y apertura de seccionadores adyacentes.

### 5.3.1.2 Definición de la estructura de los MCBC

La estructura conformada para el desarrollo del localizador contempla cinco módulos básicos. Cada módulo contiene ya sea una red LAMDA o red SVM entrenada para un propósito particular según se explica a continuación y de acuerdo con la figura 5.10.

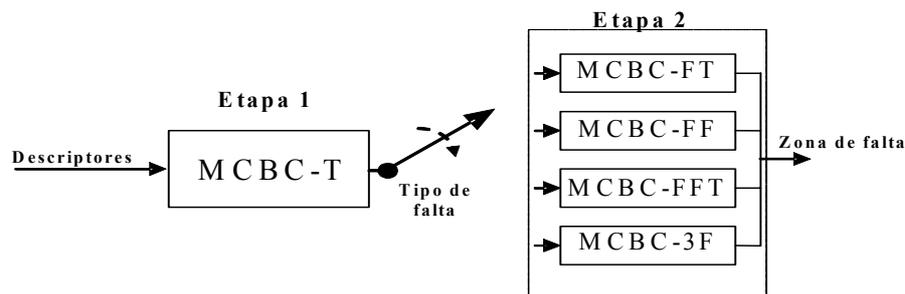


Figura 5.10. Estructura general de los MCBC para localización de la falta.

En la etapa uno, el MCBC-T se entrena para determinar el tipo de falta. En la etapa dos, cada MCBC está entrenado para reconocer la zona en la cual ocurre cada tipo de falta así: MCBC-FT esta encargado de encontrar la zona en la cual ocurrió una falta monofásica, en cualquier fase; MCBC-FF se entrena para encontrar la zona en la cual se produjo una falta bifásica, MCBC-FFT, para una falta bifásica que incluye tierra y finalmente, MCBC-3F para el caso de faltas trifásicas.

### 5.3.1.3 Adquisición de la base de datos de falta

Cuando un evento de falta ocurre en el sistema de potencia, se debe tener la posibilidad de registrar tanto las señales de tensión como de corriente, medidas en la subestación. A partir del registro de las señales de falta, se debe crear una base de datos, en la cual, cada registro esté identificado con el tipo de falta y la localización de la misma.

Teniendo en cuenta que en un sistema las faltas son excepcionales, el número de datos de falta puede ser insuficiente para estructurar una base de datos confiable. Para compensar este problema, la base de datos puede ser obtenida con una combinación de registros de falta reales, o con el uso de un software especializado para realizar simulaciones con un circuito base, modelado con los parámetros del circuito real.

### 5.3.1.4 Pre-procesamiento de la señal

La etapa de pre-procesamiento de la señal está asociada a la obtención del conjunto de descriptores, tal como se presentó en la sección anterior. Cada conjunto de descriptores está asociado con una

etiqueta que corresponde a la zona en la cual ocurrió la falta.

En esta tesis se consideraron únicamente descriptores que fueran de fácil obtención para que la propuesta aquí presentada no esté condicionada por las limitaciones de equipos. Los descriptores aquí presentados, se calculan de una manera simple, para evitar la incertidumbre asociada al proceso de estimación de las magnitudes eléctricas, tal como se presenta en [MORA96] [MORA05-f].

### 5.3.1.5 Entrenamiento del MCBC

La primera parte de esta etapa es la selección del conjunto de descriptores para el proceso de entrenamiento. En esta selección se analizan las posibles combinaciones de descriptores que conforman el conjunto de entrenamiento que posibilitan una clasificación adecuada y por ende la localización correcta de la zona en la cual ocurre la falta. A partir de las pruebas de cada posible combinación de descriptores, se selecciona el conjunto con el cual los resultados de precisión son más altos. La precisión se calcula como se muestra en la ecuación (5.2)

En este trabajo, se desarrolló el entrenamiento del tipo supervisado, con el cual se obtuvieron resultados adecuados para resolver el problema planteado.

### 5.3.1.6 Pruebas de precisión del MCBC

Las pruebas de precisión se realizan para estimar el comportamiento de los MCBC ante datos que no fueron usados para el proceso de entrenamiento. Los datos de prueba tienen que corresponder a los mismos descriptores y deben recibir el mismo procesamiento que el realizado para los datos usados en el proceso de entrenamiento.

La precisión del método se calcula como se presenta en (5.2).

$$\text{Precisión} = \frac{\text{Número de registros clasificados correctamente}}{\text{Número total de registros}} \quad (5.2)$$

La precisión en el reconocimiento de las zonas en falta, es un buen indicador que permite predecir el comportamiento del MCBC ante nuevas situaciones. Este parámetro se tomará como indicador para seleccionar la alternativa que mejor permita abordar el problema de la localización de faltas.

## 5.3.2 Aplicación de LAMDA y SVM a la localización de faltas – Pruebas comparativas

### 5.3.2.1 Zonificación de la red

El sistema de distribución seleccionado para las pruebas de los MCBC es el presentado en el capítulo 1, sección 1.5, el cual se había definido previamente como sistema inicial de pruebas. Este circuito ha sido zonificado con cinco zonas tal como se presenta en la figura 5.11.

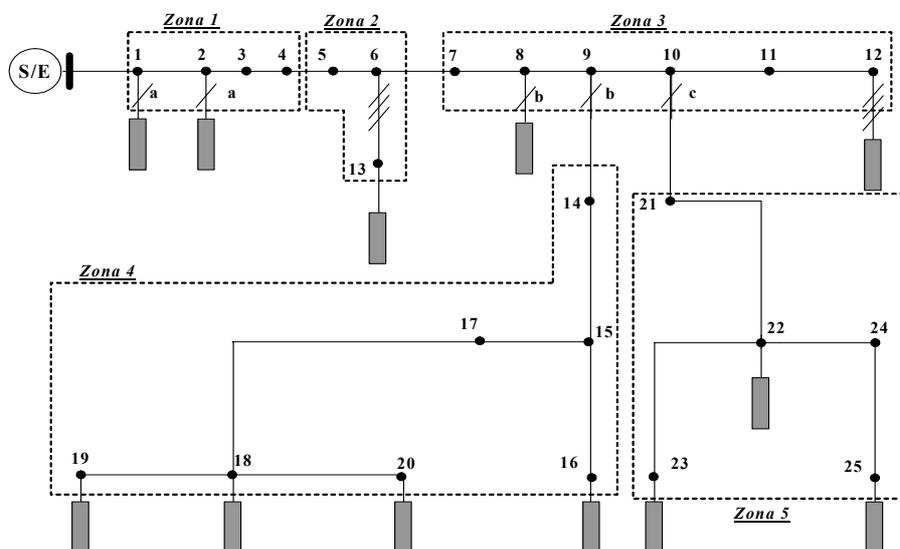


Figura 5.11. Zonificación del sistema de potencia para localización de faltas usando MCBC – Cinco zonas

La zonificación propuesta ayuda a reducir el problema de la múltiple estimación y es un primer escenario de prueba para comparación de los métodos LAMDA y SVM. A partir de este escenario y de las pruebas comparativas, se debe encontrar la opción más adecuada para resolver el problema.

### 5.3.2.2 Adquisición de la base de datos de falta

Se realizaron simulaciones en cada uno de los nodos, usando 21 resistencias de falta con valores entre  $0,05\Omega$  y  $40\Omega$  [DAGE00]. Se simularon 10 tipos de faltas, tres monofásicas, tres bifásicas, tres bifásicas a tierra y una trifásica. En total para el sistema de prueba se realizaron 2982 faltas.

Para este caso particular se realizó una simulación extensiva de faltas usando Matlab® y “Alternative Transients Program –ATP” [EMTP87]. La base de faltas, se obtuvo mediante la simulación automatizada de las mismas, tal como se propone en [MORA06-c].

### 5.3.2.3 Pre-procesamiento de la señal

A partir de los datos de falta de la base de faltas obtenida tal como se describe en el numeral anterior, se le aplican funciones de pre-procesamiento para obtener los descriptores que se presentan en el capítulo 4.

A partir de las señales de tensión y corriente se obtienen descriptores tales como la variación del valor eficaz entre los estados estables de prefalta y falta, ( $dV$  y  $dI$  respectivamente). Adicionalmente, también se estima la variación del valor de la potencia aparente ( $dS$ ), la reactancia de falta ( $X$ ) y la frecuencia de la señal transitoria de tensión ( $f$ ). Estos descriptores se definen para realizar las diferentes combinaciones que permiten hacer una prueba preliminar comparativa.

### 5.3.2.4 Entrenamiento y pruebas de precisión de los localizadores MCBC

Los datos utilizados tanto para el entrenamiento como para la prueba de precisión (ver ecuación 5.2), corresponden a cada conjunto de descriptores obtenidos de los registros de falta. El número de registros utilizados para prueba y entrenamiento se presentan en la tabla 5.4

Etapa	MCBC por tipo de falta	Cantidad de registros de entrenamiento	Cantidad de registros de prueba
Uno	T	1745	1237
	FT	625	446
Dos	FF	480	339
	FFT	480	339
	3F	160	113

Tabla 5.4. Cantidad de registros de falta usados para el entrenamiento y en las pruebas de precisión.

### 5.3.2.5 Pruebas de precisión de los localizadores MCBC

En la tabla 5.5 se presentan los resultados para las pruebas con seis diferentes combinaciones de descriptores seleccionados para esta prueba comparativa inicial. Esta prueba se realiza con los datos que no se utilizaron en el entrenamiento, tal como se presenta en la tabla 5.4. El cálculo de la precisión se realiza tal como se define en la ecuación (5.2).

### 5.3.2.6 Análisis de resultados

A partir de las pruebas realizadas para el circuito prototipo, con un tamaño de las zonas relativamente grande, se aprecia la ventaja evidente de las máquinas de soporte vectorial, razón por la cual se seleccionan como MCBC más adecuado para el tratamiento de este problema.

Sin embargo, la técnica LAMDA se probó posteriormente con una estrategia semi-supervisada, encontrándole que los resultados son un poco mejores a los presentados en la tabla 5.5, tal como se reporta en [MORA06-g] y [BARR06]. Esta estrategia se presenta en el anexo D, y muestra resultados de clasificación del 0.88 para el caso de faltas monofásicas, 0.90 para el caso de faltas bifásicas, 0.88 para faltas bifásicas a tierra y 0.84 para faltas trifásicas. Para obtener estos resultados, los descriptores de entrada fueron preprocesados mediante dos herramientas ampliamente conocidas: la búsqueda de la proyección (PPEDA) [MART03], y el análisis de componentes principales (PCA) [RENC95], tal como se muestra en el anexo C.

Sin embargo, y pese al tratamiento adicional de los descriptores, la precisión de LAMDA no es comparable con la obtenida usando la SVM como MCBC para localización de faltas. De acuerdo con los resultados previos, a continuación se presenta un análisis más detallado de las SVM y su

potencialidad para resolver el problema de la localización de faltas en sistemas de distribución.

MCBC por Tipo de falta	Combinación de descriptores	Precisión de LAMDA en la prueba	Precisión de la SVM en la prueba
Tipo de falta MCBC-T	<i>dV, dI</i>	0.96	1.00
	<i>dV, dS</i>	0.95	1.00
	<i>dI, dS</i>	0.95	1.00
	<i>dV, dI, f</i>	0.94	1.00
	<i>dV, dS, f</i>	0.94	1.00
	<i>dS, dV, dI, f</i>	0.94	1.00
Línea – tierra MCBC-F	<i>dV, dI</i>	0.63	0.99
	<i>dV, dS</i>	0.72	1.00
	<i>dI, dS</i>	0.78	1.00
	<i>dV, dI, f</i>	0.75	1.00
	<i>dV, dS, f</i>	0.77	1.00
	<i>dS, dV, dI, f</i>	0.72	1.00
Fase – fase MCBC-FF	<i>dV, dI</i>	0.75	1.00
	<i>dV, dS</i>	0.76	1.00
	<i>dI, dS</i>	0.71	1.00
	<i>dV, dI, f</i>	0.73	1.00
	<i>dV, dS, f</i>	0.77	1.00
	<i>dS, dV, dI, f</i>	0.78	1.00
Fase – fase- tierra MCBC-FFT	<i>dV, dI</i>	0.76	1.00
	<i>dV, dS</i>	0.77	1.00
	<i>dI, dS</i>	0.78	1.00
	<i>dV, dI, f</i>	0.77	1.00
	<i>dV, dS, f</i>	0.75	1.00
	<i>dS, dV, dI, f</i>	0.81	1.00
Fase – fase – tierra MCBC-3F	<i>dV, dI</i>	0.78	1.00
	<i>dV, dS</i>	0.77	0.98
	<i>dI, dS</i>	0.79	1.00
	<i>dV, dI, f</i>	0.74	1.00
	<i>dV, dS, f</i>	0.80	0.99
	<i>dS, dV, dI, f</i>	0.82	1.00

Tabla 5.5. Resultados de precisión para cada uno de los grupos de descriptores de la prueba preliminar de los MCBC.

Previamente, la técnica SVM fue probada con otro grupo de descriptores extraídos sólo de la señal de tensión [GOME05]. Sin embargo, el desempeño fue inferior al presentado en este documento.

### 5.3.3 Perfeccionamiento de la estrategia de localización de faltas usando SVM

A partir de los resultados comparativos de LAMDA y SVM se determina que la capacidad de estas últimas es comparativamente superior para resolver el problema de localización de faltas. Por tanto, se considera que se requieren pruebas más exigentes y exhaustivas que permitan determinar la capacidad real de esta técnica.

En este numeral, se presenta nuevamente el problema de localización, con un sistema zonificado de manera más exigente (Más zonas, de menor tamaño), para el cual se verifica el funcionamiento de las SVM en condición nominal y ante la variación de la carga del sistema en  $\pm 40\%$  a partir del valor nominal. A partir de este análisis, se puede observar el comportamiento del localizador basado en SVM, ante circunstancias de prueba diferentes a las de entrenamiento.

#### 5.3.3.1 Zonificación de la red

Este circuito de prueba se dividió en siete zonas tal como se presenta en la figura 5.12, considerando que se debe reducir al mínimo el problema de la múltiple estimación, característico de los MBM. Para este caso, únicamente se presenta el problema de la múltiple estimación en caso de faltas en las secciones de línea 18-19, y 18-20, las cuales por falta de datos de faltas, se dejaron en una misma zona (zona 4).

La definición de las siete zonas de falta, hace que el problema sea de mayor complejidad que en la prueba inicial, pero permite una mejor discriminación de la región en falta. Esto lo hace claramente complementario con los métodos MBM que estiman la distancia a la falta desde el sitio de medida.

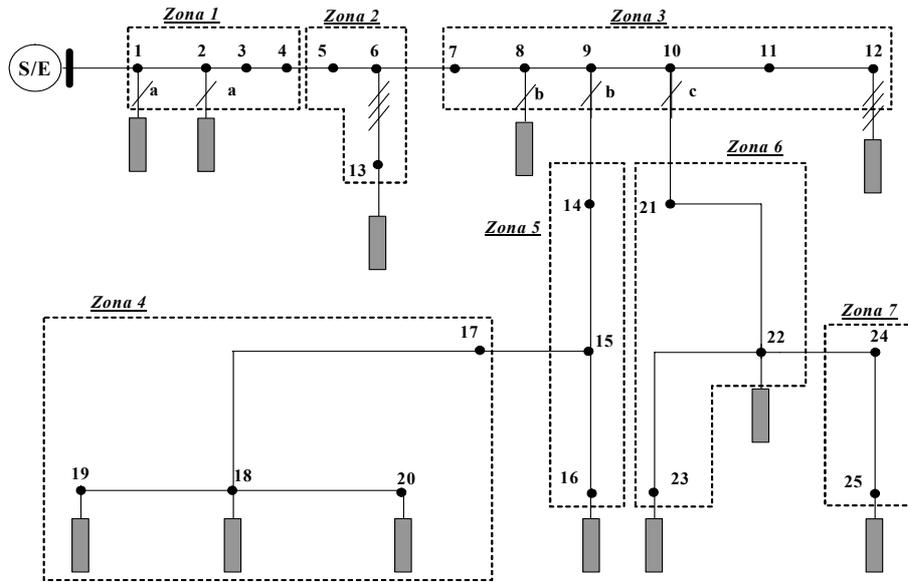


Figura 5.12: Zonificación del sistema de potencia para localización de faltas usando SVM

### 5.3.3.2 Definición de las estructuras del localizador basado en SVM

Para la conformación de las redes de SVM usadas para localización de faltas, se definieron dos estructuras, que se diferencian en la necesidad de conocer o no el tipo de falta a localizar.

La primera estructura es la misma usada para las pruebas iniciales y se denomina estructura DTF (dependiente del tipo de falta). Ésta consta de un clasificador general (SVM-T) entrenado para localizar cada tipo de falta, en serie con un arreglo de clasificadores por tipo de falta, tal como se presenta en la figura 5.13. La SVM-FT corresponde a la máquina entrenada para reconocer faltas monofásicas y se entrenó con 225 registros. Las SVM-FF y SVM-FFT corresponden a las máquinas encargadas de reconocer las faltas bifásicas y bifásicas a tierra, para cada una de las cuales se usaron 180 datos. Para el caso de la SVM-3F, que corresponde a las faltas trifásicas, se usaron 60 datos en el entrenamiento. Todo esto se presenta en la tabla 5.6.

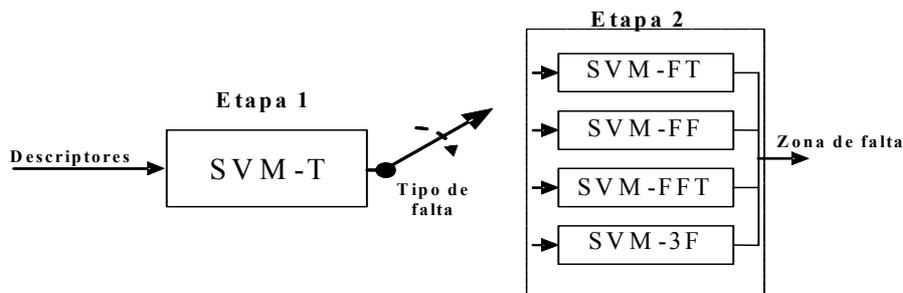


Figura 5.13. Estructura de localización de zona dependiente del tipo de falta - DTF

La segunda estructura se denomina no dependiente del tipo de falta (NDTF) y está conformada por una única máquina que sin importar el tipo de falta, obtiene la zona más probable. La estructura se presenta en la figura 5.14 y para este caso se usaron 645 registros de entrenamiento. Esta estructura se supone más exigente para la clasificación, pero se propone de igual manera, para mostrar la capacidad de las SVM para abordar el problema de localización.



Figura 5.14. Estructura de localización de zona No dependiente del tipo de falta - NDTF

### 5.3.3.3 Adquisición de datos de falta

Los datos de falta son los mismos utilizados en la sección anterior y se organizaron como se

presenta en la tabla 5.6

Localizador	Cantidad de datos de entrenamiento	Cantidad de datos de prueba
SVM-FT	225	846
SVM-FF	180	639
SVM-FFT	180	639
SVM-3F	60	213
SVM-T	645	2337
SVM G	645	2337

Tabla 5.6. Cantidad de datos de entrenamiento y prueba del localizador basado en SVM

Los localizadores SVM-T, SVM-FT, SVM-FF, SVM-FFT, SVM-3F y SVM G se presentan en las figuras 5.12 y 5.13. Para este caso, se usa un porcentaje cercano al 20% del total de los datos para el entrenamiento, mientras que para la prueba se utiliza el restante 80 % aproximadamente lo cual hace aún más exigente el problema para las SVM, si se quiere encontrar una baja cota de error [GUYO94].

#### 5.3.3.4 Pre-procesamiento de la señal

Como pre-procesamiento de la señal, se consideró únicamente la extracción de descriptores enunciados en la definición presentada en el capítulo 4.

Los descriptores que se usan son la profundidad de los huecos de tensión ( $dV$ ), la variación de el valor eficaz de la corriente ( $dI$ ), la variación de la potencia aparente ( $dS$ ), la frecuencia del transitorio ( $f$ ), la reactancia de falta ( $Xf$ ) y el valor máximo de la matriz de correlación de las muestras de corriente ( $Sv$ ).

#### 5.3.3.5 Estrategia de entrenamiento de las SVM

Las SVM se definen mediante la selección de la constante de penalización  $C$  y el valor del parámetro  $\sigma$  para el caso del “kernel” RBF seleccionado. En esta sección, como complemento de la estructura planteada y de los descriptores seleccionados, se plantea una estrategia de selección de los dos valores que definen la SVM ( $C$  y  $\sigma$ ) con el fin de tener un mejor desempeño. Esta estrategia es una alternativa básica que permite hacer la exploración en un espacio de búsqueda, para hallar la mejor configuración de la SVM para cada caso en particular.

En el entrenamiento de cada una de las máquinas de soporte vectorial, se utilizó la validación cruzada y la búsqueda en malla [LIN04] que se definen a continuación, para encontrar el mejor valor de  $C$  y del parámetro  $\sigma$  del kernel RBF.

##### a. Validación cruzada.

La validación cruzada es entendida como la división del conjunto de entrenamiento en  $n$  subconjuntos. La SVM se entrena con  $(n-1)$  subconjuntos y se valida en el entrenamiento con el restante subconjunto. Lo anterior se repite  $n$  veces, hasta que cada subconjunto haya sido utilizado como conjunto de validación. El promedio de los  $n$  resultados de la validación, es el valor que determina el comportamiento global de la SVM. Para esta aplicación y considerando la gran cantidad de datos y pruebas, el conjunto de entrenamiento se subdividió en cinco subconjuntos ( $n=5$ ).

##### b. búsqueda en malla.

La búsqueda en malla se define como la variación de los dos parámetros ( $C$  y  $\sigma$ ), en un intervalo de solución factible ( $C_{min} < C < C_{max}$  y  $\sigma_{min} < \sigma < \sigma_{max}$ ), dentro del cual se entrena cada SVM, y se valida durante el entrenamiento mediante el método de validación cruzada. De esta manera se puede escoger el mejor par de valores, para un determinado conjunto de descriptores de entrada. Mediante pruebas realizadas con diferentes intervalos, se determinó empíricamente que los valores de  $C_{min}$  y  $C_{max}$  son  $2^4$  y  $2^{32}$ , respectivamente y los pasos son de  $2^4$ . De otra parte, los valores de  $\sigma_{min}$  y  $\sigma_{max}$  son  $2^{-6}$  y  $2^6$ , respectivamente con pasos de 2.

En la figura 5.15 se presenta una sección de la malla en la cual se evalúa la función para encontrar el máximo margen de separación de las clases, mediante los métodos mencionados.

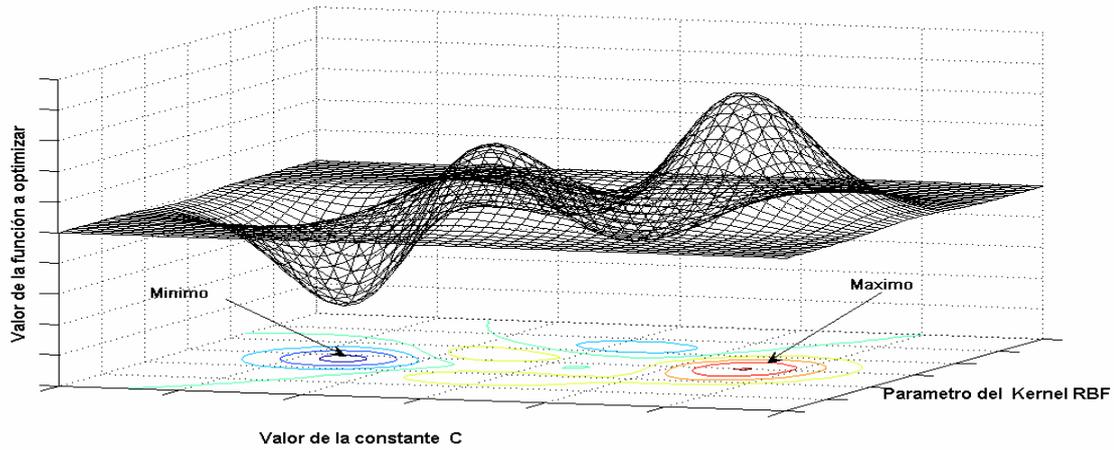


Figura 5.15. Ejemplo gráfico de la selección de  $C$  y de  $\sigma$  mediante búsqueda de malla

Luego de la selección de parámetros, la máquina se entrena con todos los datos del conjunto de prueba, para así obtener los vectores de soporte, que definen la separación de las clases.

En la tabla 5.7 se presentan los parámetros  $C$  y  $\sigma$  seleccionados para cada conjunto de descriptores, como resultado de las pruebas de validación cruzada y búsqueda en malla con los datos del conjunto de entrenamiento. En la tabla 5.7, el valor de A corresponde a  $\text{Log}_2 \sigma$ , mientras que B es  $\text{Log}_2 C$ . Esta representación se eligió para colocar resultados con números pequeños y enteros, según se explicó para la malla elegida.

Grupo	Conjunto de descriptores	SVM-FT		SVM-FF		SVM-FFT		SVM-3F		SVM-G		Grupo	Conjunto de descriptores	SVM-FT		SVM-FF		SVM-FFT		SVM-3F		SVM-G	
		A	B	A	B	A	B	A	B	A	B			A	B	A	B	A	B	A	B	A	B
1	<i>dS</i>	0	24	0	4	-1	16	-1	24	-2	24	33	<i>dV, dl, f</i>	0	28	5	24	-1	12	6	28	0	28
2	<i>dV</i>	-1	28	-1	16	-1	16	-2	28	-1	28	34	<i>dV, dl, Sv</i>	0	28	1	16	1	20	3	24	-1	20
3	<i>dl</i>	0	20	-1	24	0	20	-1	24	0	20	35	<i>dV, Xf, f</i>	-1	24	-2	8	0	16	6	24	-2	8
4	<i>Xf</i>	-3	20	1	20	2	24	4	28	-3	20	36	<i>dV, Xf, Sv</i>	-2	24	1	16	3	28	6	24	-1	24
5	<i>f</i>	-4	28	-4	8	-4	12	3	28	-4	28	37	<i>dV, f, Sv</i>	0	28	-2	8	-1	12	6	28	-1	12
6	<i>Sv</i>	2	24	5	16	5	28	-6	28	2	24	38	<i>dl, Xf, f</i>	-1	28	-2	12	-2	8	6	24	-1	28
7	<i>dS, dV</i>	-1	28	-2	12	0	12	2	24	-1	28	39	<i>dl, Xf, Sv</i>	-2	24	-1	16	3	24	6	24	-1	28
8	<i>dS, dl</i>	0	20	1	20	-2	8	1	20	-1	28	40	<i>dl, f, Sv</i>	-2	12	-3	8	-2	12	3	24	-3	8
9	<i>dS, Xf</i>	-1	24	4	24	1	20	5	28	-1	24	41	<i>Xf, f, Sv</i>	-3	20	-1	12	-2	12	6	24	-2	28
10	<i>dS, f</i>	-3	28	-2	12	-2	8	3	24	-3	8	42	<i>dS, dV, dl, Xf</i>	-1	24	4	24	1	12	6	24	-1	28
11	<i>dS, Sv</i>	0	24	-1	12	0	20	1	28	-2	24	43	<i>dS, dV, dl, f</i>	-2	24	3	16	-1	12	5	24	0	28
12	<i>dV, dl</i>	-1	28	2	20	1	20	3	28	-1	28	44	<i>dS, dV, dl, Sv</i>	1	28	3	20	0	16	4	24	1	24
13	<i>dV, Xf</i>	-2	24	3	20	2	24	6	24	0	28	45	<i>dS, dV, Xf, f</i>	-2	24	-2	8	-2	8	6	24	-1	24
14	<i>dV, f</i>	-4	28	-2	12	-1	12	3	28	-1	16	46	<i>dS, dV, Xf, Sv</i>	-1	28	-1	12	1	12	6	24	-1	24
15	<i>dV, Sv</i>	-1	16	1	24	2	20	2	28	-1	28	47	<i>dS, dV, f, Sv</i>	-2	24	-2	8	-2	8	6	28	-1	28
16	<i>dl, Xf</i>	-2	28	3	24	-1	12	6	24	-2	24	48	<i>dS, dl, Xf, f</i>	-1	24	-2	12	-2	8	6	24	-3	20
17	<i>dl, f</i>	-2	12	-3	8	-3	8	3	28	-2	12	49	<i>dS, dl, Xf, Sv</i>	-2	24	1	20	-1	8	6	24	-2	24
18	<i>dl, Sv</i>	-1	28	-1	24	0	28	0	28	-2	24	50	<i>dS, dl, f, Sv</i>	-2	28	-2	12	-2	8	5	24	-3	28
19	<i>Xf, f</i>	-2	24	5	24	3	24	6	28	-1	28	51	<i>dS, Xf, f, Sv</i>	-2	24	-3	8	-2	8	6	24	-3	24
20	<i>Xf, Sv</i>	-3	20	2	24	1	24	6	24	-2	28	52	<i>dV, dl, Xf, f</i>	-1	28	3	16	-1	12	6	24	0	28
21	<i>f, Sv</i>	-4	20	-3	12	-4	4	3	28	-4	28	53	<i>dV, dl, Xf, Sv</i>	-1	28	2	16	2	16	6	24	-1	24
22	<i>dS, dV, dl</i>	0	24	3	24	3	24	4	28	-1	28	54	<i>dV, dl, f, Sv</i>	3	28	4	24	-1	12	6	28	0	24
23	<i>dS, dV, Xf</i>	-1	24	-2	8	0	16	6	24	-1	28	55	<i>dV, Xf, f, Sv</i>	-1	28	-2	8	-1	12	6	24	-1	28
24	<i>dS, dV, f</i>	-1	28	-2	8	-2	8	4	24	-1	28	56	<i>dl, Xf, f, Sv</i>	-2	24	-2	12	0	16	6	24	-1	28
25	<i>dS, dV, Sv</i>	1	28	5	24	0	16	3	28	0	28	57	<i>dS, dV, dl, Xf, f</i>	0	28	5	24	-1	12	6	24	0	28
26	<i>dS, dl, Xf</i>	-2	24	-2	12	-1	8	6	24	-2	24	58	<i>dS, dV, dl, Xf, Sv</i>	-1	24	4	20	1	12	6	24	-1	24
27	<i>dS, dl, f</i>	-3	28	-2	12	-3	8	4	28	-3	28	59	<i>dS, dV, dl, f, Sv</i>	-1	28	5	24	-1	12	6	28	2	28
28	<i>dS, dl, Sv</i>	2	24	1	20	-1	12	2	28	-2	28	60	<i>dS, dV, Xf, f, Sv</i>	-2	28	-2	8	-1	8	6	20	-1	24
29	<i>dS, Xf, f</i>	-2	24	-1	12	0	12	6	24	-2	24	61	<i>dS, dl, Xf, f, Sv</i>	-1	28	-2	12	-2	8	6	24	-3	20
30	<i>dS, Xf, Sv</i>	-2	24	-1	12	-4	8	6	24	-1	24	62	<i>dV, dl, Xf, f, Sv</i>	-2	28	4	24	0	12	6	24	0	28
31	<i>dS, f, Sv</i>	-3	28	-3	8	-3	8	3	24	-3	24	63	<i>dS, dV, dl, Xf, f, Sv</i>	-1	28	6	24	-1	12	6	24	0	12
32	<i>dV, dl, Xf</i>	-2	28	4	20	3	16	6	24	-1	28												

Tabla 5.7., Valores de la constante  $C$  y el parámetro  $\sigma$  del kernel RBF, obtenidos como resultado de la validación cruzada y la búsqueda en malla

### 5.3.3.6 Pruebas de precisión del localizador basado en SVM

La siguiente parte del proceso de obtención del modelo en los MCBC, es la prueba de precisión, que determina si el entrenamiento realizado en la etapa anterior es válido o se requiere de uno adicional, con nuevos y más datos o con descriptores diferentes.

Los datos utilizados para las pruebas de precisión no se utilizaron en el entrenamiento (ver tabla 5.6). En la tabla 5.8 se presentan los resultados para las combinaciones de descriptores en la etapa de prueba. Para el caso de la SVM-T, por ser un problema de fácil resolución, solo se usaron cuatro combinaciones de dos descriptores, con todos los cuales se obtiene una precisión igual a la unidad.

Grupo	Conjunto de descriptores	SVM-FT	SVM-FF	SVM-FFT	SVM-3F	SVM-G	Grupo	Conjunto de descriptores	SVM-FT	SVM-FF	SVM-FFT	SVM-3F	SVM-G
1	<i>dS</i>	0,924	0,914	1,000	0,906	0,940	33	<i>dV, dl, f</i>	0,979	1,000	1,000	1,000	0,993
2	<i>dV</i>	0,898	1,000	1,000	0,690	0,868	34	<i>dV, dl, Sv</i>	0,981	1,000	1,000	1,000	0,994
3	<i>dl</i>	0,879	1,000	0,995	1,000	0,869	35	<i>dV, Xf, f</i>	0,993	1,000	1,000	1,000	0,989
4	<i>Xf</i>	1,000	1,000	1,000	1,000	0,965	36	<i>dV, Xf, Sv</i>	1,000	1,000	0,989	1,000	0,997
5	<i>f</i>	0,981	1,000	1,000	1,000	0,981	37	<i>dV, f, Sv</i>	0,970	1,000	1,000	1,000	0,979
6	<i>Sv</i>	0,546	0,653	0,671	0,347	0,382	38	<i>dl, Xf, f</i>	0,993	1,000	1,000	1,000	0,998
7	<i>dS, dV</i>	0,998	1,000	1,000	1,000	0,997	39	<i>dl, Xf, Sv</i>	1,000	1,000	1,000	1,000	0,999
8	<i>dS, dl</i>	0,950	1,000	1,000	0,986	0,994	40	<i>dl, f, Sv</i>	0,957	1,000	1,000	1,000	0,980
9	<i>dS, Xf</i>	1,000	0,992	1,000	1,000	0,998	41	<i>Xf, f, Sv</i>	1,000	1,000	1,000	1,000	0,998
10	<i>dS, f</i>	0,980	1,000	0,998	0,986	0,983	42	<i>dS, dV, dl, Xf</i>	1,000	1,000	1,000	1,000	1,000
11	<i>dS, Sv</i>	0,927	0,991	0,997	0,981	0,972	43	<i>dS, dV, dl, f</i>	0,996	1,000	1,000	1,000	0,991
12	<i>dV, dl</i>	0,994	1,000	1,000	1,000	0,996	44	<i>dS, dV, dl, Sv</i>	0,975	1,000	1,000	0,995	0,994
13	<i>dV, Xf</i>	1,000	1,000	1,000	1,000	0,995	45	<i>dS, dV, Xf, f</i>	0,996	1,000	1,000	1,000	0,997
14	<i>dV, f</i>	0,991	1,000	1,000	1,000	0,984	46	<i>dS, dV, Xf, Sv</i>	1,000	1,000	1,000	1,000	0,999
15	<i>dV, Sv</i>	0,908	1,000	0,995	0,991	0,975	47	<i>dS, dV, f, Sv</i>	0,998	1,000	1,000	1,000	0,996
16	<i>dl, Xf</i>	1,000	1,000	1,000	1,000	1,000	48	<i>dS, dl, Xf, f</i>	0,996	1,000	1,000	1,000	1,000
17	<i>dl, f</i>	0,961	1,000	1,000	0,995	0,985	49	<i>dS, dl, Xf, Sv</i>	1,000	1,000	1,000	1,000	1,000
18	<i>dl, Sv</i>	0,955	1,000	0,981	1,000	0,971	50	<i>dS, dl, f, Sv</i>	0,980	1,000	1,000	0,991	0,993
19	<i>Xf, f</i>	0,996	1,000	1,000	1,000	0,995	51	<i>dS, Xf, f, Sv</i>	0,995	1,000	1,000	1,000	0,999
20	<i>Xf, Sv</i>	1,000	1,000	1,000	1,000	1,000	52	<i>dV, dl, Xf, f</i>	0,994	1,000	1,000	1,000	0,996
21	<i>f, Sv</i>	0,979	0,997	1,000	0,991	0,989	53	<i>dV, dl, Xf, Sv</i>	1,000	1,000	1,000	1,000	0,999
22	<i>dS, dV, dl</i>	0,983	1,000	1,000	1,000	0,994	54	<i>dV, dl, f, Sv</i>	0,948	1,000	1,000	1,000	0,993
23	<i>dS, dV, Xf</i>	1,000	1,000	1,000	1,000	0,999	55	<i>dV, Xf, f, Sv</i>	0,993	1,000	1,000	1,000	0,996
24	<i>dS, dV, f</i>	0,991	1,000	1,000	0,995	0,995	56	<i>dl, Xf, f, Sv</i>	0,995	1,000	1,000	1,000	0,995
25	<i>dS, dV, Sv</i>	0,981	0,995	1,000	1,000	0,995	57	<i>dS, dV, dl, Xf, f</i>	0,983	1,000	1,000	1,000	0,995
26	<i>dS, dl, Xf</i>	1,000	1,000	1,000	1,000	1,000	58	<i>dS, dV, dl, Xf, Sv</i>	0,995	1,000	1,000	1,000	1,000
27	<i>dS, dl, f</i>	0,978	1,000	1,000	1,000	0,995	59	<i>dS, dV, dl, f, Sv</i>	0,981	1,000	1,000	1,000	0,992
28	<i>dS, dl, Sv</i>	0,956	1,000	1,000	0,977	0,996	60	<i>dS, dV, Xf, f, Sv</i>	0,996	1,000	1,000	1,000	0,998
29	<i>dS, Xf, f</i>	0,994	1,000	1,000	1,000	0,997	61	<i>dS, dl, Xf, f, Sv</i>	0,986	1,000	1,000	1,000	0,998
30	<i>dS, Xf, Sv</i>	1,000	1,000	0,997	1,000	1,000	62	<i>dV, dl, Xf, f, Sv</i>	0,998	1,000	1,000	1,000	0,993
31	<i>dS, f, Sv</i>	0,981	1,000	1,000	0,991	0,991	63	<i>dS, dV, dl, Xf, f, Sv</i>	0,995	1,000	1,000	1,000	0,991
32	<i>dV, dl, Xf</i>	1,000	1,000	1,000	1,000	0,998							

Tabla 5.8. Resultados de la prueba de los localizadores basados en SVM para una zonificación de siete zonas del sistema de prueba.

### 5.3.3.7 Análisis de resultados a partir del comportamiento de los descriptores

Los resultados obtenidos con una precisión perfecta, usando los descriptores en la tabla 5.8 muestran las máquinas de soporte vectorial como una excelente alternativa para el desarrollo de localizadores de la zona de falta.,

Una ventaja importante de la metodología basada en SVM está asociada con la simplicidad del pre-tratamiento que reciben las medidas de tensión y corriente, lo cual le confiere una potencialidad enorme.

La variación de la tensión (*dV*), tiene un comportamiento fuertemente influenciado por la resistencia de falta y la distancia a la falta. Sin embargo, y según los resultados de la tabla 5.7, este descriptor puede dar excelentes resultados para la localización de faltas fase-fase-tierra., En combinación con otros descriptores, se muestra que es útil para localizar cualquier tipo de falta.

Al igual que la variación de la tensión, la variación de la corriente ( $dI$ ) y la del valor propio, muestran una alta influencia de la resistencia de falta y la distancia al punto de medida. Sin embargo, el descriptor asociado a la variación de la corriente presenta buenos resultados en combinación con otros descriptores, según se muestra en la tabla 5.8.

En el caso del valor propio máximo de la matriz de correlación de la corriente ( $S_v$ ), los resultados no tienen el desempeño comparativo como los demás descriptores. La precisión cuando éste es el único descriptor en la entrada es muy baja, tal como se presenta en la tabla 5.8. Este fenómeno es debido a la fuerte variación de su valor, como consecuencia de la variación del valor de la resistencia de falta.

La variación de la potencia aparente ( $dS$ ), no está tan influenciada por la variación de la resistencia de falta., Este descriptor presenta buen comportamiento para la localización de faltas, y para alcanzar un comportamiento óptimo, debe estar acompañado de otro descriptor, tal como se evidencia en la tabla 5.8.

Finalmente, la frecuencia del transitorio causado al ocurrir la falta ( $f$ ) y la reactancia de fase ( $X_f$ ) presentan un muy buen comportamiento para propósitos de localización, para todos los localizadores evaluados, según se muestra en la tabla 5.8.

#### **5.3.4 Análisis del desempeño del localizador de faltas basado en SVM ante la variación de carga en el sistema de distribución**

Debido a que la carga de un sistema varía considerablemente durante la operación, el desempeño del método MCBC basado en SVM se probó para evaluar la sensibilidad de la localización ante la variabilidad de la carga., En la prueba del método, se entrena y valida tal como se presenta en la tabla 5.8 para una condición de carga nominal. Posteriormente, se verifica su desempeño utilizando para las pruebas, datos obtenidos de faltas en el sistema, pero ahora sometido a variaciones  $\pm 40\%$  en la carga del sistema.

Se realizaron pruebas para 19530 faltas (todos los nodos y los distintos tipos de faltas), las diferentes condiciones de carga (60%, 80%, 100%, 125% y 140% de la carga nominal), y 21 resistencias de falta que varían entre  $0,05\Omega$  y  $40\Omega$  [DAGE00][MORA-c], como se presenta a continuación.

##### **5.3.4.1 Definición de las pruebas**

Para las pruebas se definen diferentes escenarios que permiten ver la variación de los resultados del localizador basado en SVM entrenado con la carga nominal, y la realización de pruebas con datos de faltas correspondientes a otros niveles de carga.,

El primer escenario está asociado al entrenamiento del localizador de faltas basado en SVM para la condición de carga nominal, el cual se presentó en la tabla 5.8., Los otros escenarios están asociados a la prueba del localizador con datos de otras condiciones de carga, tales como 60%, 80%, 125% y 140% de la carga nominal, que se presentan a continuación.

Para estas pruebas, además de los descriptores de fase, se toman también los obtenidos de las señales de línea ( $L$ ), ya que a partir de éstos últimos se ofrece indirectamente información del ángulo, no considerada en los descriptores si se presentan sólo los de las señales de fase.

##### **5.3.4.2 Resultados de las pruebas**

A continuación se presentan los resultados para las pruebas realizadas al localizador de faltas basado en SVM, ante las circunstancias expuestas en el numeral anterior.

En el caso de todas las situaciones diferentes a la carga nominal, se presentan los mejores resultados de precisión obtenidos en las pruebas con los datos de faltas del sistema bajo la nueva condición de carga. Los resultados para 60%, 80%, 125% y 140 %, de la carga nominal se presentan en la tabla 5.9.

Tipo de falta	Sistema al 60 % de la carga nominal		Sistema al 80 % de la carga nominal		Sistema al 125 % de la carga nominal		Sistema al 140 % de la carga nominal	
	Descriptores seleccionados	Precisión [%]	Descriptores seleccionados	Precisión [%]	Descriptores seleccionados	Precisión [%]	Descriptores seleccionados	Precisión [%]
Línea – tierra SVM-FT	<i>dV, f</i>	93,756614	<i>dS, dV, f</i>	95,8730159	<i>f</i>	94,8148148	<i>f</i>	94,6031746
	<i>f</i>	90,158731	<i>dV, f</i>	95,7671958	<i>dV, f</i>	94,2857143	<i>dl, f</i>	93,1216931
	<i>dl, f</i>	86,984127	<i>dS, dV, dl, f</i>	94,8148148	<i>dl, f</i>	94,1798942	<i>dV, f</i>	92,3809524
	<i>dV, dl, f</i>	84,761905	<i>dV, dl, f</i>	94,7089947	<i>dV, dl, f</i>	93,7566138	<i>dS, f</i>	90,5820106
	<i>dV</i>	83,597884	<i>f</i>	93,3333333	<i>dS, f</i>	92,6984127	<i>dV, dl, f</i>	90,2645503
	<i>dVL, dIL</i>	82,222222	<i>dS</i>	93,1216931	<i>dS, dV, dl, f</i>	91,6402116		
	<i>dS</i>	82,010582	<i>dS, dl, f</i>	93,1216931	<i>dS, dV, f</i>	90,8994709		
	<i>dS, dV, f</i>	80,529101	<i>dl, f</i>	92,3809524	<i>dS</i>	90,7936508		
		<i>dVL, dIL</i>	90,8994709					
Fase – fase SVM-FF	<i>f</i>	98,9417989	<i>f</i>	100	<i>f</i>	99,7354497	<i>dV, f</i>	99,4708995
	<i>dl, f</i>	98,4126984	<i>dV, dVL</i>	100	<i>dV, f</i>	99,7354497	<i>dl, f</i>	99,3386243
	<i>dV, f</i>	98,2804233	<i>dV, f</i>	99,8677249	<i>dl, f</i>	99,7354497	<i>f</i>	99,0740741
	<i>dS, dl, f</i>	97,6190476	<i>dl, f</i>	99,7354497	<i>dS, dl, f</i>	99,7354497	<i>dS, dl, f</i>	99,0740741
	<i>dS, f</i>	95,1058201	<i>dS, dl, f</i>	99,7354497	<i>dS, f</i>	99,2063492	<i>dS, f</i>	97,8835979
	<i>dV, dl, f</i>	91,9312169	<i>dS, f</i>	98,5449735	<i>dV</i>	98,8095238	<i>dV, dl, f</i>	95,6349206
	<i>dV</i>	91,9312169	<i>dVL, dl</i>	97,3544974	<i>dV, dl, f</i>	97,8835979		
	<i>dVL, dl</i>	91,7989418	<i>dV, dl, f</i>	96,8253968	<i>dV, dVL</i>	97,6190476		
	<i>dS, dV, dl, f</i>	91,7989418	<i>dV</i>	95,8994709	<i>dS, dV, f</i>	96,1640212		
	<i>dV, dVL</i>	91,6666667	<i>dV, dl</i>	95,2380952				
<i>dV, dl</i>	91,2698413							
<i>dS, dV, f</i>	91,2698413							
Fase - fase – tierra SVM-FFT	<i>dl, f</i>	98,2804233	<i>dV, dl, f</i>	100	<i>dS, f</i>	97,3544974	<i>dS, f</i>	96,8253968
	<i>dV, f</i>	97,0899471	<i>dV, dl</i>	99,7354497	<i>dl, f</i>	96,957672	<i>dl, f</i>	96,8253968
	<i>dS, f</i>	96,4285714	<i>dVL, dl</i>	99,6031746	<i>dV, dl, f</i>	96,6931217	<i>dV, f</i>	95,6349206
	<i>dl</i>	93,2539683	<i>dV, dVL</i>	99,6031746	<i>dV, f</i>	96,031746	<i>dS, dl, f</i>	93,7830688
	<i>dV, dl, f</i>	92,9894181	<i>dV</i>	98,9417989	<i>dV, dVL</i>	95,6349206	<i>dl</i>	91,7989418
	<i>dVL, dl</i>	92,1957672	<i>dl</i>	98,4126984	<i>dV</i>	95,6349206	<i>dS, dV, dl, f</i>	91,7989418
	<i>dV, dl</i>	91,9312169	<i>dl, f</i>	98,2804233	<i>dl</i>	95,5026455	<i>dV, dl, f</i>	91,6666667
	<i>dV</i>	91,6666667	<i>dS, dl, f</i>	98,2804233	<i>dV, dl</i>	95,2380952	<i>dV, dl</i>	91,4021164
	<i>dV, dVL</i>	91,6666667	<i>dV, f</i>	97,4867725			<i>dVL, dl</i>	91,4021164
	<i>dV, dIL</i>	91,6666667	<i>dS, dV, dl, f</i>	97,2222222			<i>f</i>	91,4021164
	<i>dVL, dIL</i>	91,6666667						
	<i>dS, dV, dl</i>	91,5343915						
	<i>dS, dl, f</i>	91,5343915						
<i>f</i>	91,4021164							
Trifásica SVM-3F	<i>dS, dV, dl</i>	100	<i>dS, dV, dl</i>	100	<i>dS, dV, dl</i>	99,2063492	<i>dS, dV, dl</i>	98,4126984
	<i>dV, dl</i>	91,6666667	<i>dV, dl</i>	97,6190476	<i>dV, dl</i>	91,6666667	<i>dV, dl</i>	90,0793651
	<i>dV, dIL</i>	91,6666667	<i>dV, dIL</i>	97,6190476	<i>dV, dIL</i>	91,6666667	<i>dV, dIL</i>	90,0793651
	<i>dVL, dl</i>	91,6666667	<i>dVL, dl</i>	97,6190476	<i>dVL, dl</i>	91,6666667	<i>dVL, dl</i>	90,0793651
	<i>dVL, dIL</i>	91,6666667	<i>dVL, dIL</i>	97,6190476	<i>dVL, dIL</i>	91,6666667	<i>dVL, dIL</i>	90,0793651
			<i>dS, dl</i>	90,8730159				

Tabla 5.9. Resultados de precisión para el caso de carga diferente a la carga nominal

### 5.3.4.3 Análisis de resultados

En la tabla 5.10 se muestran los mejores resultados promedios por conjunto de descriptores obtenidos para las diferentes condiciones de carga (60%, 80%, 125% y 140% de la carga nominal). A partir de los resultados promedios, se puede determinar el comportamiento de las mejores combinaciones de descriptores.

En la tabla 5.10 se observa que la frecuencia es el descriptor que más aporta información de la localización de la falta (excepto del tipo de falta trifásica). Éste es el descriptor más robusto ante cambios de carga tal como se presenta en la figura 5.16 para los cinco valores de carga del sistema., Allí se observa que los puntos correspondientes a diferentes cargas se traslapan, lo cual indica que la frecuencia del transitorio no varía significativamente en función de la variación de la carga.

Tipo de falta	Descriptores seleccionados	Precisión de la SVM en la prueba [%]
Línea – tierra SVM 1	$dV, f$	94,0476191
	$f$	93,2275132
	$dI, f$	91,6666667
	$dV, dI, f$	90,8730159
	$dS, dV, f$	88,6772487
	$dS, dV, dI, f$	88,3068783
	$dS$	88,1481481
	$dS, f$	87,7248677
	$dS, dI, f$	87,2486772
	$dVL, dIL$	85,6349206
Fase – fase SVM 2	$dV$	84,6825397
	$f$	99,4378307
	$dV, f$	99,3386243
	$dI, f$	99,3055556
	$dS, dI, f$	99,0410053
	$dS, f$	97,6851852
	$dV, dI, f$	95,5687831
	$dV$	95,2380952
	$dV, dVL$	95,2380952
	Fase - fase – tierra SVM 3	$dI, f$
$dS, f$		96,8584656
$dV, f$		96,5608466
$dV, dI, f$		95,3373016
$dI$		94,7420635
$dS, dI, f$		94,5767196
$dV, dI$		94,5767196
$dV, dVL$		94,5436508
$dVL, dI$		94,510582
$dV$		94,1137566
Trifásica SVM 4	$dS, dV, dI$	99,4047619
	$dV, dI$	92,7579365
	$dV, dIL$	92,7579365
	$dVL, dI$	92,7579365
	$dVL, dIL$	92,7579365

Tabla 5.10. Resultados de precisión promedio para distintos tipos de carga

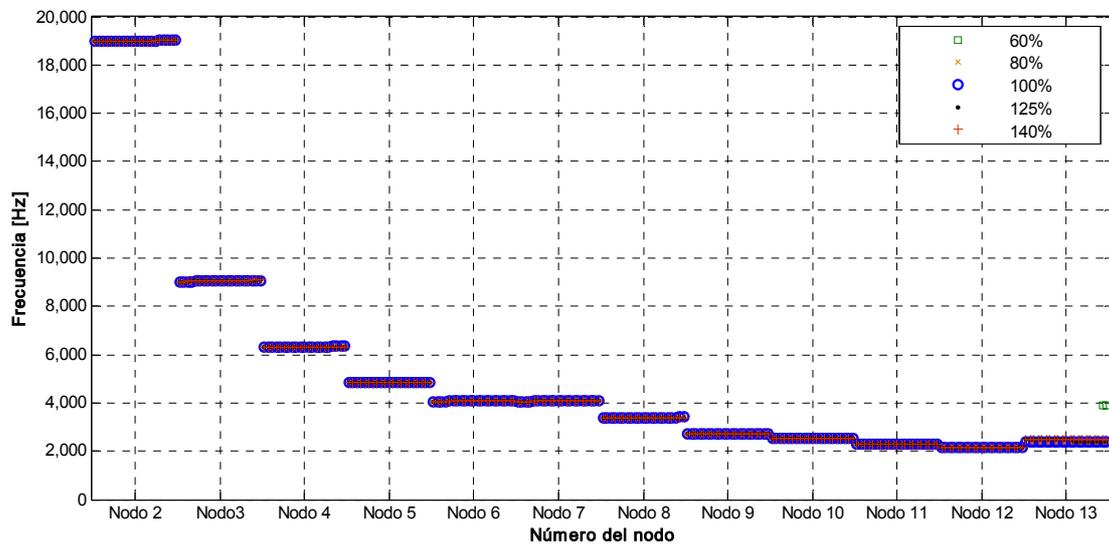


Figura 5.16. Comportamiento del descriptor de frecuencia en la fase en falta monofásica para diferentes condiciones de carga.

De otra parte, también se observa que el descriptor de reactancia no aparece en ninguno de los resultados presentados en la tabla 5.8, se puede decir que este descriptor es poco robusto ante cambios de carga. En las figuras 5.17 y 5.18 se presentan tanto el valor de la reactancia de la fase en falta como la de las fases que no están en falta, respectivamente, en el caso de faltas monofásicas.

En la figura 5.17 se muestra un buen comportamiento de la reactancia en la fase en falta para diferentes condiciones de carga. Como se presentó en el capítulo 4, es necesario incluir la

información de las otras fases para así poder atacar el problema de la múltiple estimación., En la figura 5.18 se muestra el comportamiento de una fase que no está en falta (ante una falta monofásica) para las condiciones diferentes de carga (60%, 80%, 100%, 125% y 140%), y se observa que este descriptor varía considerablemente con el cambio de carga.

La reactancia es un descriptor poco robusto ante variaciones de carga del sistema de distribución. Este comportamiento lo hace poco adecuado para localización de faltas, tal como se corrobora en los resultados de la tabla 5.8, en la cual no aparece dentro de los mejores resultados.

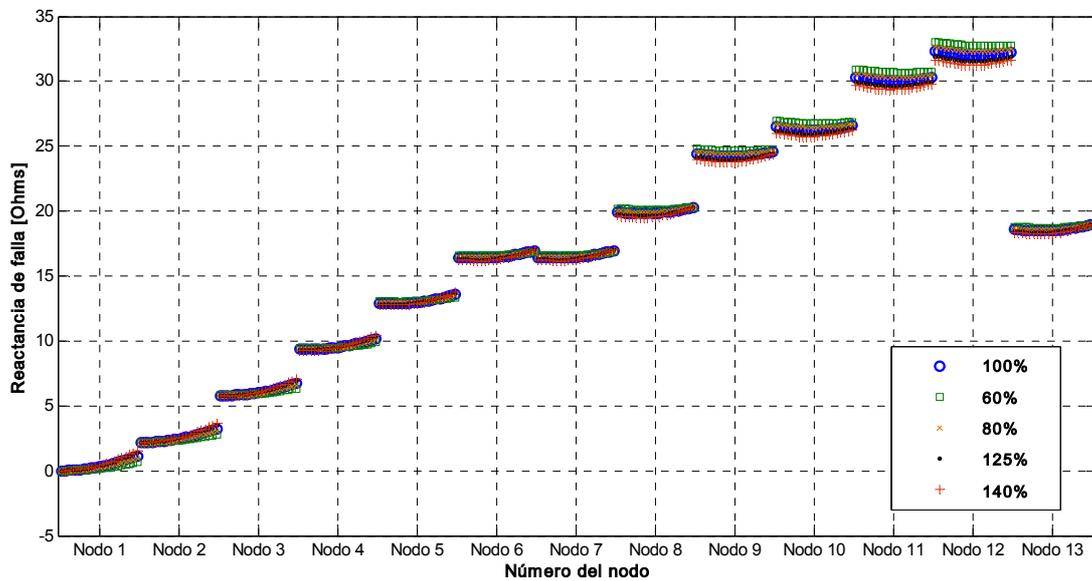


Figura 5.17. Comportamiento del descriptor de reactancia en la fase en falta monofásica para diferentes condiciones de carga

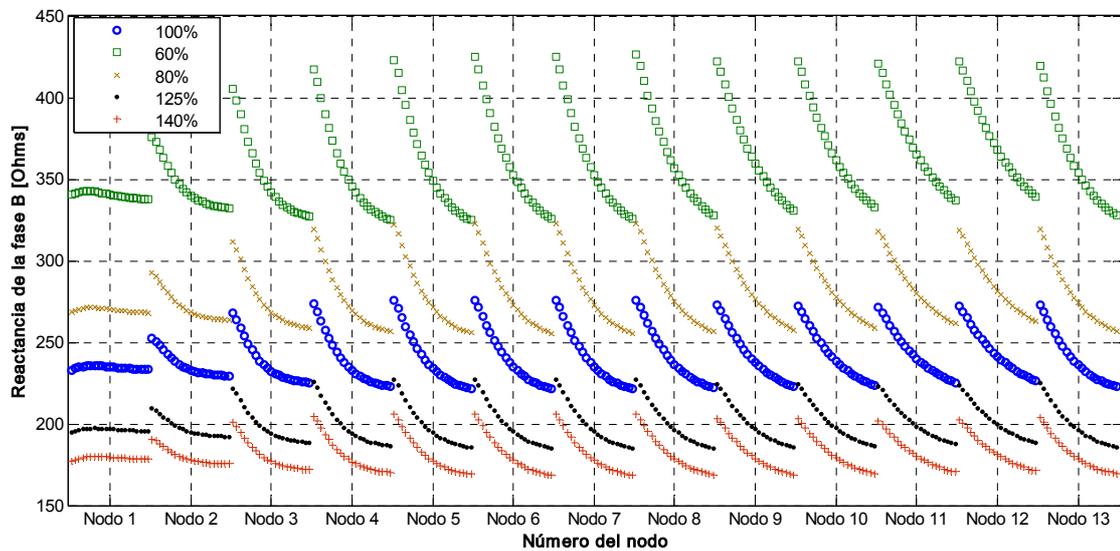


Figura 5.18. Comportamiento del descriptor de reactancia una fase que no esta en falta, ante una falta monofásica para diferentes condiciones de carga

Finalmente se puede observar de la tabla 5.10, que los descriptores asociados al cambio de la magnitud de tensión, corriente y potencia aparente, conservan cierta independencia con el cambio en la carga, de tal forma que presentan buenos resultados en la localización de la zona bajo falta.

## 5.4 Conclusiones

En este capítulo se presentaron las pruebas de los MBM, conocidos también como métodos basados en la estimación de la impedancia, aplicados a la localización de faltas en sistemas de

distribución de energía eléctrica. Estas técnicas se basan en el cálculo de la impedancia a partir de la componente fundamental de la tensión y la corriente, medidas en un terminal de la línea.,

La estimación de la distancia a la falta, obtenida a partir de los métodos basados en la impedancia, permite obtener una buena aproximación de la distancia entre el punto de medida (generalmente la subestación), y el nodo bajo falta., El principal problema de estos métodos está asociado al conocimiento del sistema y a la múltiple estimación de la distancia debido a la presencia de ramales que cumplen con el valor de impedancia o reactancia estimado.

Las técnicas LAMDA y SVM son una buena opción de solución del problema de localizar faltas en los sistemas de distribución. Sin embargo las SVM presentan mejor comportamiento por sus altos índices de precisión. La metodología se desarrolla usando como descriptores los obtenidos a partir de las señales de tensión y corriente medidas en la subestación de cabecera de los circuitos.

Los resultados obtenidos muestran que mediante la utilización de los MCBC es posible hallar relaciones entre los conjuntos de descriptores y las zonas en las cuales se presenta la falta. A partir de estos resultados se hace evidente que con la combinación de un MCBC con un MBM, se puede reducir el problema de la múltiple estimación del sitio de falta, debido que el MCBC determina la zona de falta, mientras que el MBM encuentra el sitio exacto en kilómetros, desde el punto de medida hasta la falta.

Los resultados y análisis desarrollados en esta sección muestran el primer paso para la formulación de un algoritmo híbrido de localización de faltas, que se tratará en el capítulo siguiente. Comparativamente las SVM han mostrado su versatilidad y mejores resultados, por lo que serán utilizadas en la integración con los MBM.

Los descriptores asociados a la reactancia de falta ( $X_f$ ) y al valor propio máximo de la matriz de correlación ( $S_v$ ), se descartan para su utilización en otras pruebas en este documento, debido a su bajo desempeño y robustez.

## 6 Estructuras híbridas para localización de faltas

### 6.1 Introducción

Como se presentó en el capítulo 2, existen varios MBM que se han desarrollado para su aplicación en la localización del sitio de falta en sistemas de distribución. Uno de sus principales inconvenientes, es la múltiple estimación del sitio de falta

Adicionalmente y como se describe en el capítulo 3, numerosos estudios sobre el análisis de la información y su tratamiento como conocimiento aplicado al problema del diagnóstico de faltas muestran la potencial aplicabilidad de los MCBC a la resolución del problema de localización de faltas en sistemas de distribución.

En este documento, lo que se conoce como híbrido, es la combinación de los MBM y los MCBC de una forma complementaria. Para resolver el problema de la múltiple estimación de los MBM, los MCBC identifican la zona en falta. En la zona en falta identificada, se debe buscar la estimación de los MBM para localizar con mayor precisión el sitio de falta.

A manera de contenido, este capítulo presenta inicialmente una comparación básica de los MBM y los MBC. Posteriormente se presenta la discusión de una arquitectura, citada en referencias bibliográficas sobre estudios de técnicas híbridas, a partir de la cual se presenta el modelo de integración híbrida propuesta en esta tesis. A continuación, se presentan las aplicaciones de la metodología híbrida al sistema de pruebas utilizado, así como de un sistema de prueba adicional tomado de los “test feeders” del “Distribution System Analysis Subcomitte” de la IEEE [IEEE00].

### 6.2 Comparación entre los MCBC y MBM

En la tabla 6.1 se hace una comparación generalizada entre los MCBC y los MBM. A partir de este análisis pueden plantearse métodos híbridos que suplen las deficiencias de cada uno de los métodos aplicados individualmente [MORA05-e].

Métodos Basados en el Modelo - MBM	Métodos de Clasificación basados en el Conocimiento - MCBC
La introducción de conocimientos teóricos sobre el problema es simple y directa y basta con depurarlo y convertirlo al formalismo de representación utilizado ( <i>ventaja</i> ).	Generalmente no aprovechan los conocimientos teóricos disponibles sobre el problema. Para adquirir conocimientos son necesarios los ejemplos ( <i>desventaja</i> ).
El tratamiento del problema es secuencial y el tiempo de respuesta del sistema puede ser demasiado largo ( <i>desventaja</i> ).	Los métodos están formados por redes con unidades de tratamiento de información que operan en paralelo, con tiempo de respuesta rápido ( <i>ventaja</i> ).
El aprendizaje no es un proceso de este tipo de sistemas. La adquisición de conocimientos se hace más bien por clarificación; por tanto se tiene el problema del estrangulamiento de la adquisición de conocimientos ( <i>desventaja</i> ).	El aprendizaje y la generalización de conocimientos a partir de un conjunto de ejemplos son los puntos fuertes y básicos de este tipo de métodos ( <i>ventaja</i> ).
Permiten obtener explicaciones sobre las respuestas otorgadas por el sistema con base en el proceso de razonamiento empleado y en los conocimientos cifrados en su base ( <i>ventaja</i> ).	Normalmente son cajas negras donde conocimientos se cifran en pesos e interconexiones. No se tiene acceso a una forma comprensible del conocimiento, interpretable directamente por un usuario final para explicar las respuestas obtenidas ( <i>desventaja</i> ).

Métodos Basados en el Modelo - MBM	Métodos de Clasificación basados en el Conocimiento - MCBC
Para que el sistema funcione bien se requiere que los conocimientos teóricos sean correctos y completos. Estos métodos no se adaptan al tratamiento de información aproximada o incompleta. Las variables numéricas plantea problemas ( <i>desventaja</i> ).	Las respuestas del sistema se deterioran progresivamente en presencia de una entrada distorsionada. Las redes (clasificadores) se adaptan al tratamiento de información aproximada e incompleta. Las variables continuas no plantean problemas para los métodos conexionistas ( <i>ventaja</i> ).
Los conocimientos están representados por formulas y estructuras de datos. Son conocimientos de nivel superior ( <i>ventaja</i> ).	Los conocimientos incomprensibles en las redes (clasificadores), representan bien las relaciones existentes entre sus variables de entrada y de salida ( <i>ventaja</i> ).

Tabla 6.1. Comparación entre los MBM y los MCMC. Ventajas y desventajas

### 6.3 Sistemas híbridos

Para la solución de problemas reales, generalmente complejos, se procura aprovechar todo el conocimiento disponible. La mayoría de investigaciones han demostrado que se obtienen mejores respuestas ante un problema específico, mediante la unión diferentes técnicas en lo que se conoce como “sistemas híbridos” [GIAC92].

Uno de los aspectos más importantes en el diseño de un sistema híbrido es la complementación de los dos tipos de conocimiento (teórico y empírico), de los que normalmente se dispone cuando se aborda un problema [SANT98]. Existen varias combinaciones entre métodos; la utilización de este tipo de sistemas puede tener muchas ventajas, entre ellas:

- La integración de dos técnicas complementarias permite que una supla las deficiencias de la otra, de forma que pueda obtenerse un mejor desempeño.
- El uso de diferentes técnicas de representación del conocimiento amplía la capacidad del sistema para recibir nueva información; un solo método aumenta las imposiciones y restringe el sistema, dificultando la solución del problema.
- Algunos sistemas complejos que no pueden ser tratados empleando una sola técnica pueden ser divididos en subproblemas para su solución, y luego tratarlos como un todo para obtener una solución global óptima del problema analizado.
- El procesamiento en paralelo de información proveniente de diferentes módulos, permite diseñar un sistema con mayor desempeño, más robusto y con una mayor tolerancia a fallos.

En conclusión, con la utilización de sistemas híbridos, esto es, con la integración de varios módulos que cooperan entre sí, cada uno con especificaciones particulares, se busca mejorar el tratamiento para alcanzar un objetivo común o solucionar un problema [INVA03].

#### 6.3.1 Formas de integración de métodos

La arquitectura de los sistemas puede estar determinada por las siguientes cuatro categorías: a) de componente simple; b) fusión-base; c) jerárquicos; d) híbridos. Estas categorías se presentan en [MORA05-e] [INVA03].

##### 6.3.1.1 Sistemas de componente simple

Emplean una sola técnica y buscan la forma de adaptarla para obtener la mejor solución al problema propuesto.

##### 6.3.1.2 Sistemas basados en fusión

Incluyen sistemas que combinan diferentes técnicas en un modelo único de computación.

Las ventajas de cada técnica empleada son aprovechadas de una forma complementaria, en la cual las salidas de una técnica son entradas de la otra. Esta estrategia permite realizar una fusión para alcanzar un mejor desempeño del conjunto en la resolución de problemas. Un ejemplo esquemático de estos sistemas se presenta en la figura 6.1.

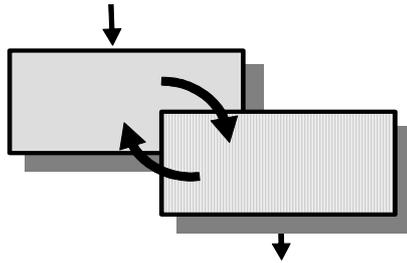


Figura 6.1. Esquema del sistema basado en fusión

### 6.3.1.3 Sistemas jerárquicos

Se conforman mediante varios módulos, cada uno con una función por desarrollar. El sistema se diseña de tal forma que cada función sea ejecutada por la técnica que brinde mejores resultados para el subproblema particular. Su correcto funcionamiento depende de la operación de las partes que lo conforman; así, un posible error en una de las partes tiende a propagarse afectando el desempeño general. El sistema no tiene lazos de realimentación o de atenuación de errores y cada técnica trabaja de manera aislada, suministrando salidas a partir de las entradas que recibe, tal como se representa en la figura 6.2.

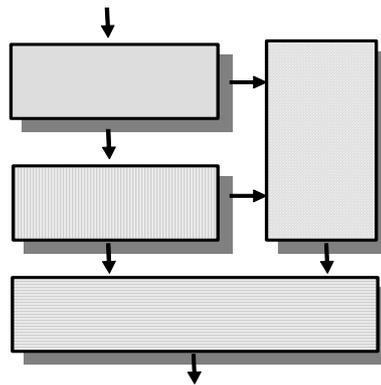


Figura 6.2. Esquema del sistema jerárquico

### 6.3.1.4 Sistemas híbridos

Tienen una arquitectura en la cual interactúan todas o algunas de las técnicas empleadas. La interacción permite explotar la integración de técnicas y su complementariedad. De esta forma se aprovechan las ventajas de diferentes técnicas para la solución de tareas específicas. Un esquema se presenta en la figura 6.3.

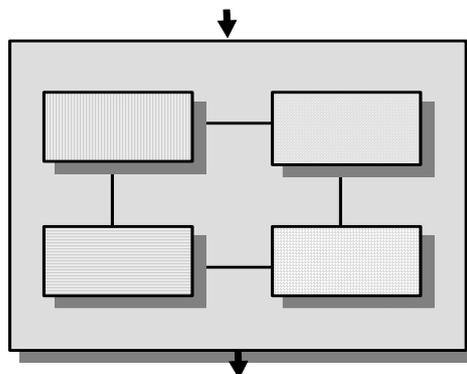


Figura 6.3. Esquema del sistema híbrido

## 6.4 Arquitectura híbrida para localización de faltas

La arquitectura propuesta para la localización de faltas, considera la unión de los métodos de clasificación basados en el conocimiento MCBC y los métodos basados en el modelo MBM. La solución propuesta pretende resolver al problema de la múltiple estimación, aun no resuelto por los métodos basados en el modelo.

### 6.4.1 Estructura básica de integración híbrida

La propuesta de arquitectura híbrida aquí presentada se fundamenta en la estructura propuesta en [RUSS02]. Esta estructura se utiliza, por considerar que estos son elementos que combinan diferentes técnicas que interactúan entre sí y con el medio externo, tal como se presenta en la figura 6.4

La arquitectura incluye la interacción de varios módulos con un sistema de realimentación que permite mejorar su desempeño interno, así como la medición y adaptación de las acciones de acuerdo con el efecto que tenga sobre el medio.

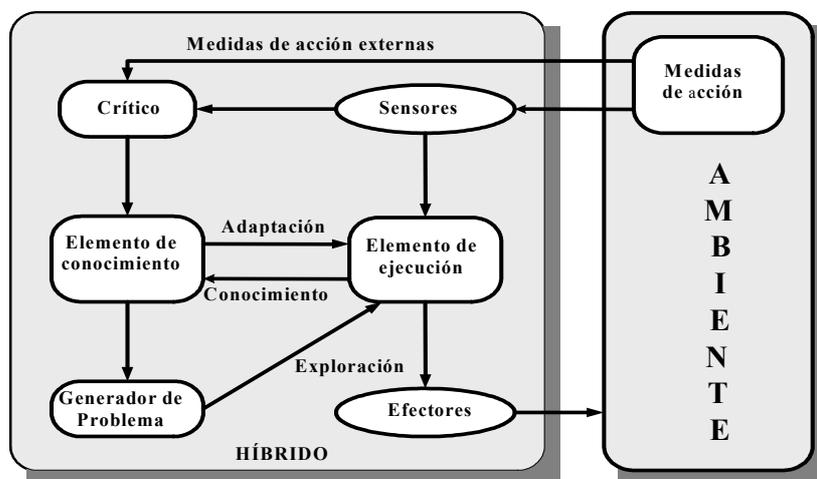


Figura 6.4. Arquitectura funcional de un híbrido.

Las partes que interactúan en la arquitectura son:

- Ambiente. En él se encuentra el proceso por controlar, el espacio de decisión por analizar o el problema de aprendizaje por resolver. Entra en contacto con el híbrido a través de los sensores. Asimismo, los efectores se constituyen en el elemento de acción.
- Elemento de desempeño. Tiene el conocimiento necesario para controlar los efectores y, por tanto, las diferentes acciones sobre el ambiente.
- Elemento de aprendizaje. Actualiza el conocimiento representado en el elemento de desempeño para optimizar las acciones del híbrido. Tiene acceso a los estados del ambiente, a las acciones anteriores y a una señal de refuerzo inmediato de las acciones, que indica la idoneidad de la última acción sobre el medio. Después de recibir la información, el elemento de desempeño se adapta para que las futuras acciones sean más pertinentes.
- Crítico. Se encarga de transformar una señal externa de refuerzo en una interna. El problema general es que la señal externa es muy leve o puede indicar algún tipo de fallo complejo, y debe traducirse en una señal de refuerzo útil como elemento de aprendizaje, con el propósito de resolver el problema en forma adecuada. La señal de refuerzo indica cuándo una acción es o no adecuada.
- Generador del problema. Su función es contribuir a la exploración del espacio del problema. En forma abstracta, propone diferentes acciones que pueden permitir descubrir nuevas y mejores soluciones. En muchos de los sistemas existentes, esto se logra mediante la adición de ruido a la acción de salida; si el sistema aún se desempeña bien puede decidirse que no se necesita una nueva solución; en caso contrario es necesario buscar mejores alternativas para la solución del problema.

### 6.4.2 Instanciación de la arquitectura híbrida al problema de localización de faltas

La estructura del sistema híbrido de localización de faltas en sistemas de distribución, se presenta en la figura 6.5. Este estructura es la adaptación particular o “instanciación” de la alternativa general propuesta en la figura 6.4

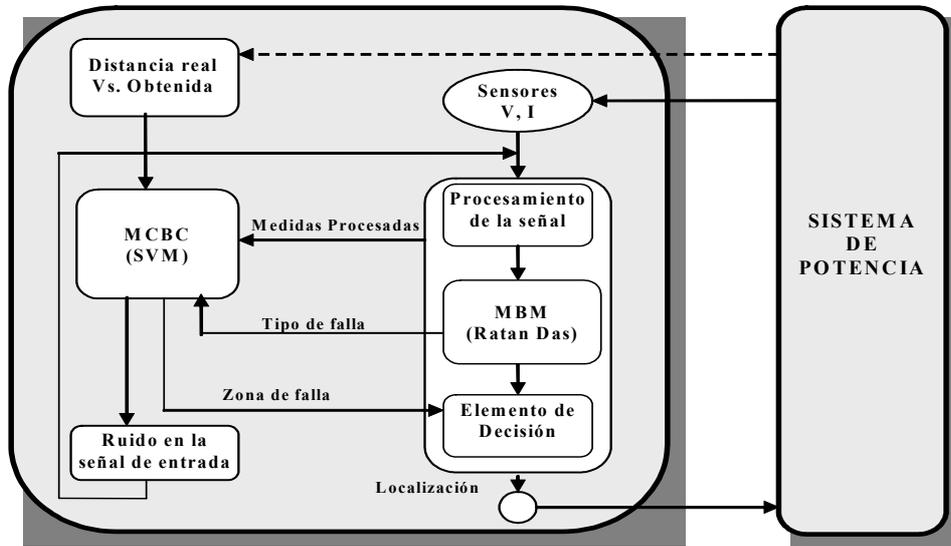


Figura 6.5. Sistema híbrido basado en MCBC y MBM para localización de faltas

En este caso de instanciación, el crítico es sólo un elemento de validación que determina la distancia real a la cual ocurrió la falta, y la distancia estimada. Esta herramienta sólo debe actuar en el periodo de entrenamiento, con el fin de incluir en la base de aprendizaje sólo los ejemplos que conducen a resultados aceptables dentro de un margen de tolerancia.

El elemento de aprendizaje está constituido por los métodos de clasificación basados en el conocimiento (MCBC) y tiene como entradas, las señales preprocesadas en forma de descriptores y la señal de refuerzo del aprendizaje definida por el crítico. El elemento de conocimiento del sistema híbrido de localización de faltas son las Máquinas de Soporte Vectorial [CORT95] debido a su desempeño, tal como se presentó en el capítulo 5. Éstas, mediante la caracterización de las señales de tensión y corriente, determinan la zona de ocurrencia de la falta.

El elemento de ejecución está conformado principalmente por el método basado en el modelo (MBM). Como MBM se implementó el algoritmo de localización de faltas de Ratan Das [DAS98], el cual, basado en las medidas de tensión y corriente, encuentra el tipo de falta y la distancia probable al punto de ocurrencia del evento.

Las otras partes que conforman el elemento de ejecución son el elemento de decisión y el de procesamiento de la señal; mediante una estrategia de resolución de conflictos (simple intersección o ajuste de pesos), el elemento de decisión obtiene la salida; el elemento generador de problema es el ruido que puede tener la señal introducida por los sensores, o que se desprende del procesamiento de la señal por la pérdida de información significativa.

Dado que no se efectúa ninguna acción sobre el sistema, no existe un elemento actuador propiamente dicho. Básicamente los sensores son los transformadores de tensión y de corriente, así como los registradores de eventos existentes en las subestaciones.

### 6.4.3 Metodología de reducción de la múltiple estimación del sitio de falta

En la estructura híbrida propuesta en el numeral anterior (figura 6.5), las señales preprocesadas en forma de descriptores constituyen las entradas del localizador de faltas basado en SVM. Los valores eficaces de tensión y corriente de prefalta y falta, son usados por el método propuesto por Ratan Das (MBM), para estimar las posibles localizaciones de la falta a partir de la estimación de una distancia. Adicionalmente, el localizador basado en SVM (MCBC), tiene como entrada, el tipo de la falta estimado por el método de Ratan Das. Posteriormente, el método basado en SVM estima la zona bajo falta, en la cual se busca la distancia entregada por el método de Ratan Das.

En la figura 6.6 se presenta gráficamente el principio de solución aplicado.

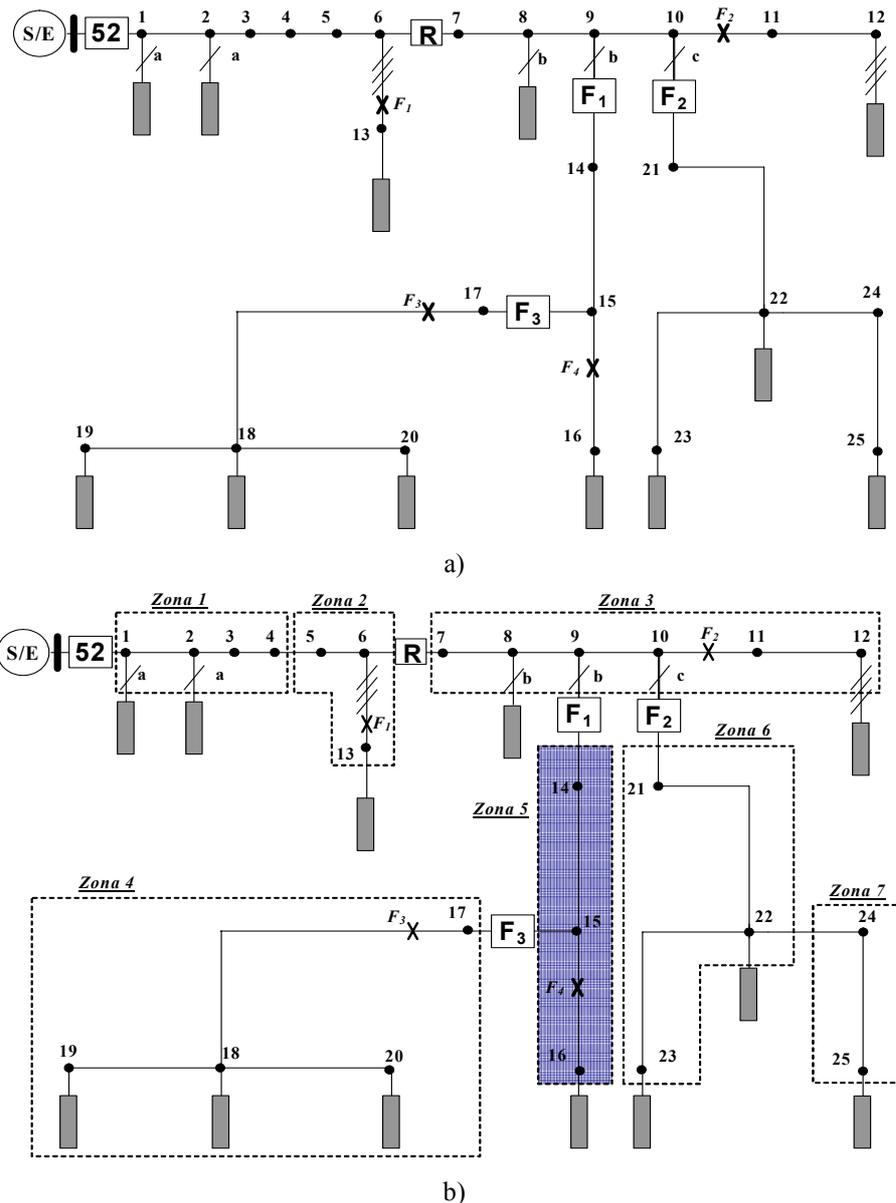


Figura 6.6. Aplicación del principio de intersección para resolver el problema de la múltiple estimación de la localización de la falta. a) Posibles localizaciones ( $F_1, F_2, F_3$  y  $F_4$ ), detectadas en diferentes secciones por un método basado en la estimación de la impedancia. b) Activación de la zona 5 como zona de falta.

Para una falta ( $F$ ), el MBM estima cuatro posibles localizaciones ( $F_1, F_2, F_3$  y  $F_4$ ), que corresponden todas a la misma reactancia vista desde la subestación y que luego se convierten a distancias equivalentes para cada una de las ramificaciones, tal como se presenta en la figura 6.6a.

En la figura 6.6b se presenta la división de zonas, las cuales deben ser identificadas con el MCBC. Para este caso, la zona determinada como la zona en falta es la zona 5.

La intersección con la zona bajo falta y los diferentes puntos identificados como bajo falta por el MBM elimina el problema de la múltiple estimación, tal como se presenta en la figura 6.6b.

Para este caso y por la presencia de laterales en el circuito, a partir de la activación de la zona 5, se reduce el problema de la múltiple localización del sitio de falta, reduciendo así el tiempo de restauración del circuito. Para el ejemplo gráfico aquí presentado, el equipo de mantenimiento, se debe buscar una distancia medida desde la subestación que está localizada entre los nodos 15 y 16.

Según el ejemplo, se muestra la importancia de una adecuada selección de las zonas, en el momento de entrenar las máquinas de soporte vectorial. Las zonas en lo posible, deben contener un solo circuito radial; sin embargo esto no siempre es posible por la necesidad de suficientes datos para entrenar adecuadamente la máquina de soporte. Siempre existe un compromiso entre el tamaño de la zona y la precisión deseada.

En caso que una falta se presente en la zona 4, aun existe la posibilidad de que haya múltiple estimación. Esto se debe a que esta zona contiene dos circuitos radiales, uno entre el nodo 18 y 19 y el otro entre los nodos 18 y 20 en donde pueden ocurrir faltas. Es por esto, que de forma general y para cualquier circuito, se propone la reducción del problema y no su eliminación completa.

Para este caso se utiliza la estructura dependiente del tipo de falta (DTF), para el localizador basado en SVM que se presenta en la figura 5.13.

## 6.5 Resultados de la metodología aplicada al sistema de prueba

A manera de prueba de la integración de los métodos, se presenta un caso de algunas faltas en diferentes zonas, como ejemplo de la localización. Estos resultados se obtienen a partir del uso de la estructura propuesta en la figura 6.5, donde se integra el método propuesto por R. Das [DAS98] como MBM y las máquinas de soporte vectorial, como MCBC.

Para este caso y a partir de los resultados obtenidos con el sistema de prueba y presentados en el capítulo 5, se presenta la integración de métodos. No se muestran resultados extensivos debido a que con anterioridad se ha mostrado la validez de cada uno de los métodos por separado, con un número de datos lo suficientemente válido para considerarse estable bajo las condiciones propuestas, tal como se presenta en las tablas 5.1, 5.2 y 5.3 para los MBM (R. Das) y en la tabla 5.7 para el MCBC (SVM). Los resultados de la integración de los métodos se presentan en la tabla 6.2.

Tipo de falta	Resistencia de falta [ $\Omega$ ]	Nodo de falta	Localización real de la falta		Resultado del híbrido	
			Distancia [km]	Zona	Distancia [km]	Zona
Monofásica	0,05	2	2,414	1	2,410	1
	10	2	2,414	1	2,423	1
	25	2	2,414	1	2,417	1
	0,05	6	18,507	2	18,501	2
	10	6	18,507	2	18,534	2
	25	6	18,507	2	18,381	2
	0,05	19	34,922	4	35,234	4
	10	19	34,922	4	35,266	4
	25	19	34,922	4	35,331	4
	0,05	25	38,946	7	39,519	7
10	25	38,946	7	39,684	7	
25	25	38,946	7	39,778	7	
Bifásica	0,05	3	6,437	1	6,436	1
	10	3	6,437	1	6,456	1
	25	3	6,437	1	6,361	1
	0,05	8	22,530	3	22,517	3
	10	8	22,530	3	22,609	3
	25	8	22,530	3	22,546	3
	0,05	12	37,014	3	37,032	3
10	12	37,014	3	37,088	3	
25	12	37,014	3	37,153	3	
Bifásica a tierra	0,05	2	2,414	1	2,411	1
	10	2	2,414	1	2,309	1
	25	2	2,414	1	2,171	1
	0,05	7	18,507	2	18,504	2
	10	7	18,507	2	18,568	2
	25	7	18,507	2	18,465	2
	0,05	11	34,600	3	34,560	3
	10	11	34,600	3	34,926	3
25	11	34,600	3	34,911	3	
Trifásica	0,05	4	10,461	1	10,205	1
	10	4	10,461	1	10,210	1
	25	4	10,461	1	10,291	1
	0,05	9	27,680	3	27,457	3
	10	9	27,680	3	27,296	3
	25	9	27,680	3	27,384	3
	0,05	12	37,014	3	36,900	3
	10	12	37,014	3	36,668	3
25	12	37,014	3	36,719	3	

Tabla 6.2: Resultados de la pruebas con el método híbrido integrado

Según el compendio de resultados presentados en la tabla 6.2, complementado con los resultados obtenidos para el MBM y el MCBC presentados en las tablas 5.1, 5.2, 5.3 y 5.7, se puede demostrar que la estructura híbrida presentada, permite la solución de problemas mediante la integración del conocimiento teórico y empírico. En este caso, se particulariza para uno de los problemas no resueltos del sector eléctrico, la localización de faltas en sistemas de distribución.

El problema de la múltiple estimación, para el ejemplo de aplicación es plenamente resuelto, ya que a partir de la adecuada definición de las zonas y el error del cero % en la validación permiten identificar correctamente la zona bajo falta.

## 6.6 Pruebas de la metodología híbrida aplicada al sistema de 34 nodos de la IEEE

Como complemento a las pruebas desarrolladas en el sistema definido en la sección 1.5, en esta sección se presentan la aplicación de la metodología híbrida a un sistema plenamente definido como prototipo de validación por el “Distribution System Analysis Subcommittee” del “Institute of Electrical and Electronics Engineers”.

### 6.6.1 Sistema de distribución IEEE de 34 nodos

El sistema de distribución utilizado como prueba, es el alimentador de 34 nodos que se presenta en la figura 6.7. Los parámetros del sistema “IEEE 34 Node Test Feeder” se presentan en el anexo E. Según [IEEE00], este alimentador está localizado en Arizona, y opera a una tensión nominal de 24.9 kV. Entre sus principales características están la presencia de ramales monofásicos y trifásicos, múltiples calibres de conductor, cargas desbalanceadas de naturaleza concentrada y distribuida.

Como condiciones particulares del circuito se tiene las siguientes: El sistema original tiene dos reguladores de tensión y bancos de condensadores, que no fueron implementados en esta aplicación. Las cargas se modelaron como impedancia constante y se conectaron en estrella. La fuente se conectó con una potencia de cortocircuito en la subestación de 20000 MVA [VIAW06].

### 6.6.2 Zonificación del sistema

El sistema de distribución se zonificó de acuerdo con los criterios presentados en la sección 5.3.1.1. Los criterios que mayor influencia tuvieron en la zonificación, fueron la reducción del problema de la múltiple estimación y la disponibilidad de datos para cada una de las zonas del sistema. El sistema completamente zonificado se presenta en la figura 6.7

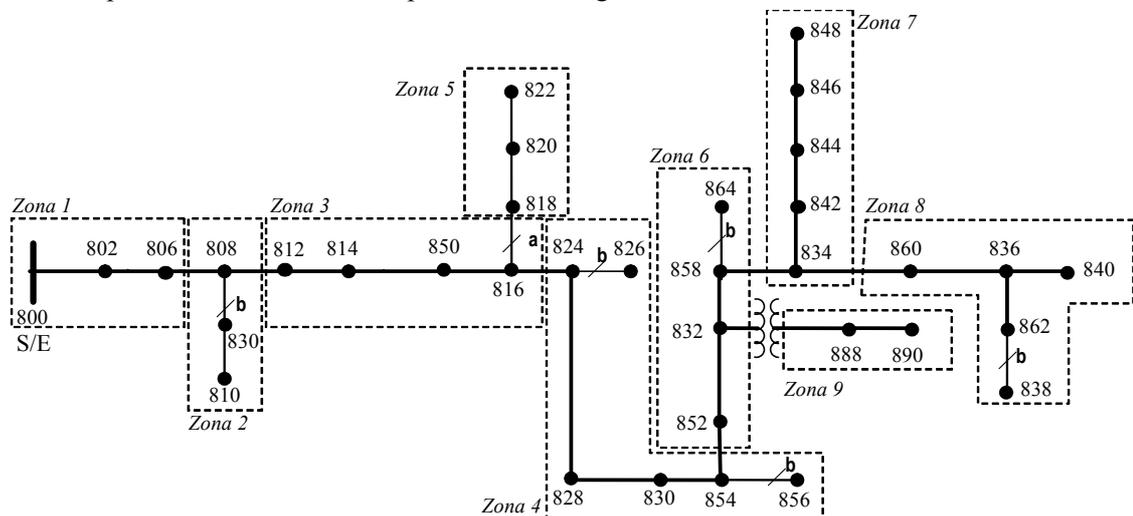


Figura 6.7. Zonificación del sistema prototipo de la IEEE de 34 nodos, de tensión nominal de 24.9 kV

### 6.6.3 Configuración del MCBC

De acuerdo a los resultados obtenidos en el capítulo anterior, la estructura del localizador es la DTF que se presenta en la figura 5.13. Adicionalmente, se utiliza el mismo criterio de validación cruzada con la subdivisión del conjunto de entrenamiento en cinco grupos ( $n=5$ ), y la búsqueda de malla, para la definición de los valores de  $C$  y  $\sigma$ , de las SVM con “kernel” RBF.

Los datos de entrenamiento y de prueba de cada una de los localizadores SVM se presentan en la tabla 6.3. Al igual que las pruebas presentadas en el capítulo 5, también se usan menos datos en el entrenamiento que en la prueba. Esto permite mostrar la robustez de la propuesta de localizador.

Etapa	MCBC por tipo de falta	Cantidad de datos de entrenamiento	Cantidad de datos de prueba
Uno	T	1345	4304
	FT	435	1392
Dos	FF	390	1248
	FFT	390	1248
	3F	130	416

Tabla 6.3. Cantidad de registros de falta usados para el entrenamiento y en las pruebas de precisión.

### 6.6.4 Resultados del MCBC

Para la red SVM-T de la etapa uno, se comprueba que se obtiene un resultado del 100% de precisión para la mayoría de las combinaciones de más de dos descriptores. En las tablas 6.4, 6.5, 6.6 y 6.7 se presentan los resultados de precisión, para el caso de los localizadores SVM-FT, SVM-FF, SVMFFT y SVM-3F, respectivamente.

Las tablas muestran los resultados para todas las combinaciones de uno a cuatro descriptores, de los preseleccionados en el capítulo anterior como los mejores para propósito de localización de falta. Los descriptores utilizados son la variación de la potencia aparente de falta con respecto a la de prefalta ( $dS$ ), la variación del valor eficaz de la tensión y corriente ( $dV$ ,  $dI$ ), y finalmente la frecuencia de la señal transitoria causada por la falta ( $f$ ). Se usaron tanto valores de fase, como de línea ( $L$ ), para un total de siete descriptores. La precisión en la prueba se calcula como se presenta en la ecuación (5.2).

Finalmente, debido a la alta cantidad de resultados con acierto perfecto en el reconocimiento de zonas, no se considera necesario intentar un mayor número de combinaciones de descriptores. De esta forma se presentan varias alternativas prácticas de implementación en circuitos reales, con un relativamente bajo número de entradas.

#### 6.6.4.1 Parámetros seleccionados en el entrenamiento

Como resultado del entrenamiento, se presentan en las tablas 6.4 a 6.7 con los parámetros seleccionados a partir del entrenamiento para cada una de las combinaciones de descriptores y para cada uno de los localizadores de la estructura DFT. En las tablas, se presenta el valor del logaritmo en base dos, de  $C$  y  $\sigma$ , debido que se usaron las potencias de dos como variación de la malla.

Conjunto de descriptores	$\text{Log}_2 \sigma$	$\text{Log}_2 C$	Conjunto de descriptores	$\text{Log}_2 \sigma$	$\text{Log}_2 C$	Conjunto de descriptores	$\text{Log}_2 \sigma$	$\text{Log}_2 C$	Conjunto de descriptores	$\text{Log}_2 \sigma$	$\text{Log}_2 C$
$dS$	-1	24	$dI, dIL$	-2	20	$dSL, dI, dIL$	-1	20	$dS, dV, dVL, dIL$	-1	20
$dSL$	-5	16	$dI, f$	-2	20	$dSL, dI, f$	-2	16	$dS, dV, dVL, f$	0	24
$dV$	-2	28	$dIL, f$	-3	16	$dSL, dIL, f$	-1	20	$dS, dV, dI, dIL$	-2	16
$dVL$	-3	24	$dS, dSL, dV$	-1	20	$dV, dVL, dI$	-1	20	$dS, dV, dI, f$	-1	20
$dI$	-2	20	$dS, dSL, dVL$	-1	20	$dV, dVL, dIL$	-1	24	$dS, dV, dIL, f$	-2	16
$dIL$	-3	24	$dS, dSL, dI$	-1	20	$dV, dVL, f$	-1	24	$dS, dVL, dI, dIL$	-1	20
$f$	-6	4	$dS, dSL, dIL$	-1	20	$dV, dI, dIL$	-2	24	$dS, dVL, dI, f$	-1	20
$dS, dSL$	-1	20	$dS, dSL, f$	-1	16	$dV, dI, f$	-2	16	$dS, dVL, dIL, f$	0	24
$dS, dV$	-1	24	$dS, dV, dVL$	-1	20	$dV, dIL, f$	0	20	$dS, dI, dIL, f$	-1	16
$dS, dVL$	-1	20	$dS, dV, dI$	-2	20	$dVL, dI, dIL$	-2	24	$dSL, dV, dVL, dI$	-1	24
$dS, dI$	-3	16	$dS, dV, dIL$	-2	20	$dVL, dI, f$	-1	16	$dSL, dV, dVL, dIL$	-1	28
$dS, dIL$	-2	24	$dS, dV, f$	-2	16	$dVL, dIL, f$	-1	20	$dSL, dV, dVL, f$	1	20
$dS, f$	-2	24	$dS, dVL, dI$	-1	20	$dI, dIL, f$	-2	20	$dSL, dV, dI, dIL$	-1	20
$dSL, dV$	-1	20	$dS, dVL, dIL$	-1	20	$dS, dSL, dV, dVL$	-1	20	$dSL, dV, dI, f$	-2	16
$dSL, dVL$	-2	24	$dS, dVL, f$	0	20	$dS, dSL, dV, dI$	-1	20	$dSL, dV, dIL, f$	-2	12
$dSL, dI$	-1	20	$dS, dI, dIL$	-2	20	$dS, dSL, dV, dIL$	-1	20	$dSL, dVL, dI, dIL$	-1	24
$dSL, dIL$	-2	24	$dS, dI, f$	-1	24	$dS, dSL, dV, f$	-2	12	$dSL, dVL, dI, f$	-2	12
$dSL, f$	-1	20	$dS, dIL, f$	-2	20	$dS, dSL, dVL, dI$	-1	20	$dSL, dVL, dIL, f$	0	24
$dV, dVL$	-2	24	$dSL, dV, dVL$	-1	20	$dS, dSL, dVL, dIL$	-1	20	$dSL, dI, dIL, f$	-2	12
$dV, dI$	-2	20	$dSL, dV, dI$	-1	20	$dS, dSL, dVL, f$	0	24	$dV, dVL, dI, dIL$	-1	20
$dV, dIL$	-2	24	$dSL, dV, dIL$	-1	28	$dS, dSL, dI, dIL$	-1	20	$dV, dVL, dI, f$	-1	16
$dV, f$	-2	20	$dSL, dV, f$	-2	16	$dS, dSL, dI, f$	-2	12	$dV, dVL, dIL, f$	0	20
$dVL, dI$	-2	24	$dSL, dVL, dI$	-1	24	$dS, dSL, dIL, f$	-1	20	$dV, dI, dIL, f$	-2	16
$dVL, dIL$	-2	24	$dSL, dVL, IL$	-2	20	$dS, dV, dVL, dI$	-1	20	$dVL, dI, dIL, f$	-1	16
$dVL, f$	-2	20	$dSL, dVL, f$	-2	12						

Tabla 6.4. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FT

Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C
<i>dS</i>	-2	28	<i>dl, dIL</i>	-3	20	<i>dSL, dl, dIL</i>	-2	24	<i>dS, dV, dVL, dIL</i>	-2	20
<i>dSL</i>	-3	24	<i>dl, f</i>	-3	20	<i>dSL, dl, f</i>	-3	20	<i>dS, dV, dVL, f</i>	-2	24
<i>dV</i>	-2	24	<i>dIL, f</i>	-3	20	<i>dSL, dIL, f</i>	-2	20	<i>dS, dV, dl, dIL</i>	-2	20
<i>dVL</i>	-3	24	<i>dS, dSL, dV</i>	-2	20	<i>dV, dVL, dl</i>	-2	24	<i>dS, dV, dl, f</i>	-2	24
<i>dl</i>	-3	24	<i>dS, dSL, dVL</i>	-2	20	<i>dV, dVL, dIL</i>	-2	24	<i>dS, dV, dIL, f</i>	-2	24
<i>dIL</i>	-3	28	<i>dS, dSL, dl</i>	-2	20	<i>dV, dVL, f</i>	-2	24	<i>dS, dVL, dl, dIL</i>	-2	20
<i>f</i>	-5	12	<i>dS, dSL, dIL</i>	-2	20	<i>dV, dl, dIL</i>	-2	28	<i>dS, dVL, dl, f</i>	-2	24
<i>dS, dSL</i>	-2	24	<i>dS, dSL, f</i>	-2	20	<i>dV, dl, f</i>	-1	24	<i>dS, dVL, dIL, f</i>	-2	24
<i>dS, dV</i>	-2	20	<i>dS, dV, dVL</i>	-2	20	<i>dV, dIL, f</i>	-2	20	<i>dS, dl, dIL, f</i>	-2	24
<i>dS, dVL</i>	-2	24	<i>dS, dV, dl</i>	-2	20	<i>dVL, dl, dIL</i>	-3	24	<i>dSL, dV, dVL, dl</i>	-2	28
<i>dS, dl</i>	-2	24	<i>dS, dV, dIL</i>	-2	20	<i>dVL, dl, f</i>	-3	20	<i>dSL, dV, dVL, dIL</i>	-2	24
<i>dS, dIL</i>	-2	20	<i>dS, dV, f</i>	-2	24	<i>dVL, dIL, f</i>	-3	20	<i>dSL, dV, dVL, f</i>	-2	24
<i>dS, f</i>	-2	28	<i>dS, dVL, dl</i>	-2	20	<i>dl, dIL, f</i>	-2	20	<i>dSL, dV, dl, dIL</i>	-2	20
<i>dSL, dV</i>	-3	20	<i>dS, dVL, dIL</i>	-2	20	<i>dS, dSL, dV, dVL</i>	-2	20	<i>dSL, dV, dl, f</i>	-2	24
<i>dSL, dVL</i>	-3	20	<i>dS, dVL, f</i>	-2	24	<i>dS, dSL, dV, dl</i>	-2	20	<i>dSL, dV, dIL, f</i>	-2	20
<i>dSL, dl</i>	-3	20	<i>dS, dl, dIL</i>	-2	20	<i>dS, dSL, dV, dIL</i>	-2	20	<i>dSL, dVL, dl, dIL</i>	-2	24
<i>dSL, dIL</i>	-3	28	<i>dS, dl, f</i>	-2	24	<i>dS, dSL, dV, f</i>	-2	24	<i>dSL, dVL, dl, f</i>	-2	24
<i>dSL, f</i>	-3	24	<i>dS, dIL, f</i>	-2	24	<i>dS, dSL, dVL, dl</i>	-2	20	<i>dSL, dVL, dIL, f</i>	-2	24
<i>dV, dVL</i>	-3	24	<i>dSL, dV, dVL</i>	-2	28	<i>dS, dSL, dVL, dIL</i>	-2	20	<i>dSL, dl, dIL, f</i>	-2	20
<i>dV, dl</i>	-3	24	<i>dSL, dV, dl</i>	-2	28	<i>dS, dSL, dVL, f</i>	-2	24	<i>dV, dVL, dl, dIL</i>	-2	24
<i>dV, dIL</i>	-3	20	<i>dSL, dV, dIL</i>	-2	20	<i>dS, dSL, dl, dIL</i>	-2	20	<i>dV, dVL, dl, f</i>	-2	24
<i>dV, f</i>	-2	28	<i>dSL, dV, f</i>	-2	24	<i>dS, dSL, dl, f</i>	-2	24	<i>dV, dVL, dIL, f</i>	-2	24
<i>dVL, dl</i>	-3	20	<i>dSL, dVL, dl</i>	-2	24	<i>dS, dSL, dIL, f</i>	-2	24	<i>dV, dl, dIL, f</i>	-2	24
<i>dVL, dIL</i>	-3	20	<i>dSL, dVL, IL</i>	-2	24	<i>dS, dV, dVL, dl</i>	-2	20	<i>dVL, dl, dIL, f</i>	-2	24
<i>dVL, f</i>	-3	24	<i>dSL, dVL, f</i>	-3	20						

Tabla 6.5. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FF

Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C
<i>dS</i>	-2	20	<i>dl, dIL</i>	-2	20	<i>dSL, dl, dIL</i>	-1	24	<i>dS, dV, dVL, dIL</i>	0	24
<i>dSL</i>	-3	24	<i>dl, f</i>	-2	24	<i>dSL, dl, f</i>	-2	20	<i>dS, dV, dVL, f</i>	-1	20
<i>dV</i>	-2	24	<i>dIL, f</i>	-3	20	<i>dSL, dIL, f</i>	-2	20	<i>dS, dV, dl, dIL</i>	-2	20
<i>dVL</i>	-2	28	<i>dS, dSL, dV</i>	-1	20	<i>dV, dVL, dl</i>	-1	24	<i>dS, dV, dl, f</i>	-1	20
<i>dl</i>	-2	16	<i>dS, dSL, dVL</i>	-1	20	<i>dV, dVL, dIL</i>	-1	20	<i>dS, dV, dIL, f</i>	-1	20
<i>dIL</i>	-2	24	<i>dS, dSL, dl</i>	-1	20	<i>dV, dVL, f</i>	-2	24	<i>dS, dVL, dl, dIL</i>	-2	20
<i>f</i>	-6	28	<i>dS, dSL, dIL</i>	-1	20	<i>dV, dl, dIL</i>	-1	24	<i>dS, dVL, dl, f</i>	0	28
<i>dS, dSL</i>	-3	20	<i>dS, dSL, f</i>	-1	20	<i>dV, dl, f</i>	-1	20	<i>dS, dVL, dIL, f</i>	-1	20
<i>dS, dV</i>	-1	24	<i>dS, dV, dVL</i>	-1	20	<i>dV, dIL, f</i>	-2	16	<i>dS, dl, dIL, f</i>	-1	20
<i>dS, dVL</i>	-1	24	<i>dS, dV, dl</i>	-2	20	<i>dVL, dl, dIL</i>	-1	24	<i>dSL, dV, dVL, dl</i>	-1	20
<i>dS, dl</i>	-2	16	<i>dS, dV, dIL</i>	-1	24	<i>dVL, dl, f</i>	-1	24	<i>dSL, dV, dVL, dIL</i>	-1	24
<i>dS, dIL</i>	-2	20	<i>dS, dV, f</i>	-1	24	<i>dVL, dIL, f</i>	-2	20	<i>dSL, dV, dVL, f</i>	-2	20
<i>dS, f</i>	-3	24	<i>dS, dVL, dl</i>	-2	20	<i>dl, dIL, f</i>	-2	20	<i>dSL, dV, dl, dIL</i>	-1	24
<i>dSL, dV</i>	-1	28	<i>dS, dVL, dIL</i>	-1	20	<i>dS, dSL, dV, dVL</i>	-1	24	<i>dSL, dV, dl, f</i>	-1	20
<i>dSL, dVL</i>	-2	20	<i>dS, dVL, f</i>	-1	24	<i>dS, dSL, dV, dl</i>	-1	24	<i>dSL, dV, dIL, f</i>	-2	20
<i>dSL, dl</i>	-2	20	<i>dS, dl, dIL</i>	-2	16	<i>dS, dSL, dV, dIL</i>	-1	24	<i>dSL, dVL, dl, dIL</i>	0	24
<i>dSL, dIL</i>	-1	24	<i>dS, dl, f</i>	-1	24	<i>dS, dSL, dV, f</i>	-1	20	<i>dSL, dVL, dl, f</i>	-1	24
<i>dSL, f</i>	-3	20	<i>dS, dIL, f</i>	-1	24	<i>dS, dSL, dVL, dl</i>	-1	24	<i>dSL, dVL, dIL, f</i>	-2	20
<i>dV, dVL</i>	-2	24	<i>dSL, dV, dVL</i>	-1	24	<i>dS, dSL, dVL, dIL</i>	-1	24	<i>dSL, dl, dIL, f</i>	-1	20
<i>dV, dl</i>	-2	20	<i>dSL, dV, dl</i>	-1	20	<i>dS, dSL, dVL, f</i>	-1	20	<i>dV, dVL, dl, dIL</i>	-1	20
<i>dV, dIL</i>	-1	28	<i>dSL, dV, dIL</i>	-1	24	<i>dS, dSL, dl, dIL</i>	-1	24	<i>dV, dVL, dl, f</i>	-1	24
<i>dV, f</i>	-2	28	<i>dSL, dV, f</i>	-1	20	<i>dS, dSL, dl, f</i>	-1	20	<i>dV, dVL, dIL, f</i>	-2	20
<i>dVL, dl</i>	-2	20	<i>dSL, dVL, dl</i>	-1	20	<i>dS, dSL, dIL, f</i>	-1	20	<i>dV, dl, dIL, f</i>	0	24
<i>dVL, dIL</i>	-1	24	<i>dSL, dVL, IL</i>	-1	20	<i>dS, dV, dVL, dl</i>	-2	20	<i>dVL, dl, dIL, f</i>	-1	20
<i>dVL, f</i>	-2	24	<i>dSL, dVL, f</i>	-1	20						

Tabla 6.6. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FFT

Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C	Conjunto de descriptores	Log <sub>2</sub> σ	Log <sub>2</sub> C
<i>dS</i>	-3	24	<i>dl, dIL</i>	-3	24	<i>dSL, dl, dIL</i>	-2	24	<i>dS, dV, dVL, dIL</i>	-1	24
<i>dSL</i>	-4	24	<i>dl, f</i>	-3	24	<i>dSL, dl, f</i>	0	28	<i>dS, dV, dVL, f</i>	-1	20
<i>dV</i>	-2	24	<i>dIL, f</i>	0	20	<i>dSL, dIL, f</i>	-2	24	<i>dS, dV, dl, dIL</i>	-1	24
<i>dVL</i>	-2	24	<i>dS, dSL, dV</i>	0	28	<i>dV, dVL, dl</i>	0	24	<i>dS, dV, dl, f</i>	-1	28
<i>dl</i>	-4	24	<i>dS, dSL, dVL</i>	-1	24	<i>dV, dVL, dIL</i>	0	24	<i>dS, dV, dIL, f</i>	-2	24
<i>dIL</i>	-4	24	<i>dS, dSL, dl</i>	-1	24	<i>dV, dVL, f</i>	0	28	<i>dS, dVL, dl, dIL</i>	-1	24
<i>f</i>	-4	28	<i>dS, dSL, dIL</i>	-1	28	<i>dV, dl, dIL</i>	-1	24	<i>dS, dVL, dl, f</i>	1	24
<i>dS, dSL</i>	-3	24	<i>dS, dSL, f</i>	-3	24	<i>dV, dl, f</i>	-2	24	<i>dS, dVL, dIL, f</i>	-2	24
<i>dS, dV</i>	-1	24	<i>dS, dV, dVL</i>	-1	24	<i>dV, dIL, f</i>	-1	24	<i>dS, dl, dIL, f</i>	0	24
<i>dS, dVL</i>	0	28	<i>dS, dV, dl</i>	-1	24	<i>dVL, dl, dIL</i>	-1	28	<i>dSL, dV, dVL, dl</i>	-1	24
<i>dS, dl</i>	-1	24	<i>dS, dV, dIL</i>	-1	24	<i>dVL, dl, f</i>	-2	24	<i>dSL, dV, dVL, dIL</i>	1	28
<i>dS, dIL</i>	-2	24	<i>dS, dV, f</i>	-2	28	<i>dVL, dIL, f</i>	-1	24	<i>dSL, dV, dVL, f</i>	-1	28
<i>dS, f</i>	-1	28	<i>dS, dVL, dl</i>	-1	28	<i>dl, dIL, f</i>	-2	24	<i>dSL, dV, dl, dIL</i>	-1	28
<i>dSL, dV</i>	-2	24	<i>dS, dVL, dIL</i>	0	24	<i>dS, dSL, dV, dVL</i>	0	24	<i>dSL, dV, dl, f</i>	1	24
<i>dSL, dVL</i>	-1	24	<i>dS, dVL, f</i>	-2	28	<i>dS, dSL, dV, dl</i>	0	28	<i>dSL, dV, dIL, f</i>	0	24
<i>dSL, dl</i>	-1	24	<i>dS, dl, dIL</i>	-1	24	<i>dS, dSL, dV, dIL</i>	-1	24	<i>dSL, dVL, dl, dIL</i>	-1	28
<i>dSL, dIL</i>	-2	24	<i>dS, dl, f</i>	-2	24	<i>dS, dSL, dV, f</i>	0	24	<i>dSL, dVL, dl, f</i>	0	24
<i>dSL, f</i>	-3	24	<i>dS, dIL, f</i>	-1	24	<i>dS, dSL, dVL, dl</i>	-1	24	<i>dSL, dVL, dIL, f</i>	0	24
<i>dV, dVL</i>	-1	24	<i>dSL, dV, dVL</i>	-1	24	<i>dS, dSL, dVL, dIL</i>	0	24	<i>dSL, dl, dIL, f</i>	0	24
<i>dV, dl</i>	-1	24	<i>dSL, dV, dl</i>	0	28	<i>dS, dSL, dVL, f</i>	-2	24	<i>dV, dVL, dl, dIL</i>	0	24
<i>dV, dIL</i>	0	24	<i>dSL, dV, dIL</i>	0	24	<i>dS, dSL, dl, dIL</i>	-1	24	<i>dV, dVL, dl, f</i>	0	28
<i>dV, f</i>	-1	24	<i>dSL, dV, f</i>	-1	28	<i>dS, dSL, dl, f</i>	-2	24	<i>dV, dVL, dIL, f</i>	-1	28
<i>dVL, dl</i>	0	28	<i>dSL, dVL, dl</i>	-1	24	<i>dS, dSL, dIL, f</i>	0	24	<i>dV, dl, dIL, f</i>	-1	20
<i>dVL, dIL</i>	-1	24	<i>dSL, dVL, IL</i>	-1	24	<i>dS, dV, dVL, dl</i>	-1	24	<i>dVL, dl, dIL, f</i>	0	28
<i>dVL, f</i>	-2	28	<i>dSL, dVL, f</i>	-2	24						

Tabla 6.7. Parámetros seleccionados por el método de validación cruzada. Combinaciones de uno a cuatro descriptores. Caso del localizador SVM-3F

### 6.6.4.2 Resultados de las pruebas de precisión

En las tablas 6.8 a 6.11 se presentan los resultados de las pruebas de cada localizador, usando los datos de la última columna de la tabla 6.3.

Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba
<i>dS</i>	0,9468	<i>dl, dIL</i>	1,0000	<i>dSL, dl, dIL</i>	1,0000	<i>dS, dV, dVL, dIL</i>	1,0000
<i>dSL</i>	0,7126	<i>dl, f</i>	0,9921	<i>dSL, dl, f</i>	0,9914	<i>dS, dV, dVL, f</i>	0,9964
<i>dV</i>	0,9899	<i>dIL, f</i>	0,9856	<i>dSL, dIL, f</i>	0,9871	<i>dS, dV, dl, dIL</i>	0,9993
<i>dVL</i>	0,9698	<i>dS, dSL, dV</i>	1,0000	<i>dV, dVL, dl</i>	0,9986	<i>dS, dV, dl, f</i>	1,0000
<i>dl</i>	0,9964	<i>dS, dSL, dVL</i>	1,0000	<i>dV, dVL, dIL</i>	1,0000	<i>dS, dV, dIL, f</i>	0,9943
<i>dIL</i>	0,9382	<i>dS, dSL, dl</i>	1,0000	<i>dV, dVL, f</i>	0,9935	<i>dS, dVL, dl, dIL</i>	1,0000
<i>f</i>	0,8930	<i>dS, dSL, dIL</i>	1,0000	<i>dV, dl, dIL</i>	1,0000	<i>dS, dVL, dl, f</i>	0,9964
<i>dS, dSL</i>	1,0000	<i>dS, dSL, f</i>	0,9957	<i>dV, dl, f</i>	0,9957	<i>dS, dVL, dIL, f</i>	0,9928
<i>dS, dV</i>	1,0000	<i>dS, dV, dVL</i>	1,0000	<i>dV, dIL, f</i>	0,9899	<i>dS, dl, dIL, f</i>	0,9964
<i>dS, dVL</i>	1,0000	<i>dS, dV, dl</i>	1,0000	<i>dVL, dl, dIL</i>	1,0000	<i>dSL, dV, dVL, dl</i>	1,0000
<i>dS, dl</i>	0,9978	<i>dS, dV, dIL</i>	1,0000	<i>dVL, dl, f</i>	0,9957	<i>dSL, dV, dVL, dIL</i>	1,0000
<i>dS, dIL</i>	1,0000	<i>dS, dV, f</i>	0,9978	<i>dVL, dIL, f</i>	0,9871	<i>dSL, dV, dVL, f</i>	0,9864
<i>dS, f</i>	0,9921	<i>dS, dVL, dl</i>	1,0000	<i>dl, dIL, f</i>	0,9943	<i>dSL, dV, dl, dIL</i>	1,0000
<i>dSL, dV</i>	0,9899	<i>dS, dVL, dIL</i>	1,0000	<i>dS, dSL, dV, dVL</i>	1,0000	<i>dSL, dV, dl, f</i>	0,9907
<i>dSL, dVL</i>	0,9921	<i>dS, dVL, f</i>	0,9971	<i>dS, dSL, dV, dl</i>	1,0000	<i>dSL, dV, dIL, f</i>	0,9892
<i>dSL, dl</i>	0,9993	<i>dS, dl, dIL</i>	1,0000	<i>dS, dSL, dV, dIL</i>	1,0000	<i>dSL, dVL, dl, dIL</i>	1,0000
<i>dSL, dIL</i>	0,9727	<i>dS, dl, f</i>	1,0000	<i>dS, dSL, dV, f</i>	0,9950	<i>dSL, dVL, dl, f</i>	0,9907
<i>dSL, f</i>	0,9878	<i>dS, dIL, f</i>	0,9964	<i>dS, dSL, dVL, dl</i>	1,0000	<i>dSL, dVL, dIL, f</i>	0,9921
<i>dV, dVL</i>	1,0000	<i>dSL, dV, dVL</i>	0,9899	<i>dS, dSL, dVL, dIL</i>	1,0000	<i>dSL, dl, dIL, f</i>	0,9914
<i>dV, dl</i>	1,0000	<i>dSL, dV, dl</i>	1,0000	<i>dS, dSL, dVL, f</i>	0,9964	<i>dV, dVL, dl, dIL</i>	0,9993
<i>dV, dIL</i>	0,9986	<i>dSL, dV, dIL</i>	0,9986	<i>dS, dSL, dl, dIL</i>	1,0000	<i>dV, dVL, dl, f</i>	0,9943
<i>dV, f</i>	0,9914	<i>dSL, dV, f</i>	0,9907	<i>dS, dSL, dl, f</i>	0,9950	<i>dV, dVL, dIL, f</i>	0,9842
<i>dVL, dl</i>	1,0000	<i>dSL, dVL, dl</i>	1,0000	<i>dS, dSL, dIL, f</i>	0,9928	<i>dV, dl, dIL, f</i>	0,9935
<i>dVL, dIL</i>	0,9914	<i>dSL, dVL, IL</i>	0,9799	<i>dS, dV, dVL, dl</i>	1,0000	<i>dVL, dl, dIL, f</i>	0,9943
<i>dVL, f</i>	0,9835	<i>dSL, dVL, f</i>	0,9871				

Tabla 6.8. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FT

Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba
<i>dS</i>	0,9591	<i>dl, dIL</i>	0,9976	<i>dSL, dl, dIL</i>	0,9984	<i>dS, dV, dVL, dIL</i>	0,9976
<i>dSL</i>	0,9663	<i>dl, f</i>	0,9800	<i>dSL, dl, f</i>	0,9760	<i>dS, dV, dVL, f</i>	0,9840
<i>dV</i>	0,9832	<i>dIL, f</i>	0,9792	<i>dSL, dIL, f</i>	0,9712	<i>dS, dV, dl, dIL</i>	1,0000
<i>dVL</i>	0,9936	<i>dS, dSL, dV</i>	1,0000	<i>dV, dVL, dl</i>	0,9968	<i>dS, dV, dl, f</i>	0,9824
<i>dl</i>	0,9968	<i>dS, dSL, dVL</i>	0,9944	<i>dV, dVL, dIL</i>	1,0000	<i>dS, dV, dIL, f</i>	0,9832
<i>dIL</i>	0,9984	<i>dS, dSL, dl</i>	0,9936	<i>dV, dVL, f</i>	0,9808	<i>dS, dVL, dl, dIL</i>	1,0000
<i>f</i>	0,8750	<i>dS, dSL, dIL</i>	0,9976	<i>dV, dl, dIL</i>	0,9928	<i>dS, dVL, dl, f</i>	0,9824
<i>dS, dSL</i>	0,9936	<i>dS, dSL, f</i>	0,9808	<i>dV, dl, f</i>	0,9792	<i>dS, dVL, dIL, f</i>	0,9872
<i>dS, dV</i>	1,0000	<i>dS, dV, dVL</i>	0,9976	<i>dV, dIL, f</i>	0,9816	<i>dS, dl, dIL, f</i>	0,9848
<i>dS, dVL</i>	0,9960	<i>dS, dV, dl</i>	1,0000	<i>dVL, dl, dIL</i>	0,9992	<i>dSL, dV, dVL, dl</i>	1,0000
<i>dS, dl</i>	1,0000	<i>dS, dV, dIL</i>	0,9976	<i>dVL, dl, f</i>	0,9808	<i>dSL, dV, dVL, dIL</i>	1,0000
<i>dS, dIL</i>	0,9968	<i>dS, dV, f</i>	0,9848	<i>dVL, dIL, f</i>	0,9808	<i>dSL, dV, dVL, f</i>	0,9760
<i>dS, f</i>	0,9688	<i>dS, dVL, dl</i>	1,0000	<i>dl, dIL, f</i>	0,9752	<i>dSL, dV, dl, dIL</i>	1,0000
<i>dSL, dV</i>	0,9976	<i>dS, dVL, dIL</i>	1,0000	<i>dS, dSL, dV, dVL</i>	0,9976	<i>dSL, dV, dl, f</i>	0,9768
<i>dSL, dVL</i>	0,9976	<i>dS, dVL, f</i>	0,9864	<i>dS, dSL, dV, dl</i>	1,0000	<i>dSL, dV, dIL, f</i>	0,9824
<i>dSL, dl</i>	0,9976	<i>dS, dl, dIL</i>	0,9968	<i>dS, dSL, dV, dIL</i>	0,9984	<i>dSL, dVL, dl, dIL</i>	1,0000
<i>dSL, dIL</i>	0,9968	<i>dS, dl, f</i>	0,9872	<i>dS, dSL, dV, f</i>	0,9840	<i>dSL, dVL, dl, f</i>	0,9776
<i>dSL, f</i>	0,9583	<i>dS, dIL, f</i>	0,9864	<i>dS, dSL, dVL, dl</i>	0,9976	<i>dSL, dVL, dIL, f</i>	0,9784
<i>dV, dVL</i>	1,0000	<i>dSL, dV, dVL</i>	0,9936	<i>dS, dSL, dVL, dIL</i>	1,0000	<i>dSL, dl, dIL, f</i>	0,9808
<i>dV, dl</i>	1,0000	<i>dSL, dV, dl</i>	0,9880	<i>dS, dSL, dVL, f</i>	0,9856	<i>dV, dVL, dl, dIL</i>	1,0000
<i>dV, dIL</i>	0,9984	<i>dSL, dV, dIL</i>	1,0000	<i>dS, dSL, dl, dIL</i>	1,0000	<i>dV, dVL, dl, f</i>	0,9792
<i>dV, f</i>	0,9760	<i>dSL, dV, f</i>	0,9760	<i>dS, dSL, dl, f</i>	0,9864	<i>dV, dVL, dIL, f</i>	0,9776
<i>dVL, dl</i>	1,0000	<i>dSL, dVL, dl</i>	0,9992	<i>dS, dSL, dIL, f</i>	0,9848	<i>dV, dl, dIL, f</i>	0,9784
<i>dVL, dIL</i>	0,9968	<i>dSL, dVL, IL</i>	0,9984	<i>dS, dV, dVL, dl</i>	0,9968	<i>dVL, dl, dIL, f</i>	0,9776
<i>dVL, f</i>	0,9784	<i>dSL, dVL, f</i>	0,9792				

Tabla 6.9. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FF

Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba
<i>dS</i>	0,9311	<i>dl, dIL</i>	0,9992	<i>dSL, dl, dIL</i>	1,0000	<i>dS, dV, dVL, dIL</i>	0,9944
<i>dSL</i>	0,9495	<i>dl, f</i>	0,9832	<i>dSL, dl, f</i>	0,9952	<i>dS, dV, dVL, f</i>	0,9984
<i>dV</i>	0,9663	<i>dIL, f</i>	0,9816	<i>dSL, dIL, f</i>	0,9936	<i>dS, dV, dl, dIL</i>	1,0000
<i>dVL</i>	0,9904	<i>dS, dSL, dV</i>	1,0000	<i>dV, dVL, dl</i>	1,0000	<i>dS, dV, dl, f</i>	0,9968
<i>dl</i>	0,9752	<i>dS, dSL, dVL</i>	1,0000	<i>dV, dVL, dIL</i>	1,0000	<i>dS, dV, dIL, f</i>	0,9968
<i>dIL</i>	0,9848	<i>dS, dSL, dl</i>	1,0000	<i>dV, dVL, f</i>	0,9928	<i>dS, dVL, dl, dIL</i>	1,0000
<i>f</i>	0,8798	<i>dS, dSL, dIL</i>	1,0000	<i>dV, dl, dIL</i>	1,0000	<i>dS, dVL, dl, f</i>	0,9928
<i>dS, dSL</i>	1,0000	<i>dS, dSL, f</i>	0,9960	<i>dV, dl, f</i>	0,9960	<i>dS, dVL, dIL, f</i>	0,9968
<i>dS, dV</i>	1,0000	<i>dS, dV, dVL</i>	1,0000	<i>dV, dIL, f</i>	0,9944	<i>dS, dl, dIL, f</i>	0,9968
<i>dS, dVL</i>	1,0000	<i>dS, dV, dl</i>	1,0000	<i>dVL, dl, dIL</i>	1,0000	<i>dSL, dV, dVL, dl</i>	1,0000
<i>dS, dl</i>	1,0000	<i>dS, dV, dIL</i>	1,0000	<i>dVL, dl, f</i>	0,9960	<i>dSL, dV, dVL, dIL</i>	1,0000
<i>dS, dIL</i>	1,0000	<i>dS, dV, f</i>	0,9968	<i>dVL, dIL, f</i>	0,9968	<i>dSL, dV, dVL, f</i>	0,9952
<i>dS, f</i>	0,9639	<i>dS, dVL, dl</i>	1,0000	<i>dl, dIL, f</i>	0,9936	<i>dSL, dV, dl, dIL</i>	1,0000
<i>dSL, dV</i>	1,0000	<i>dS, dVL, dIL</i>	1,0000	<i>dS, dSL, dV, dVL</i>	1,0000	<i>dSL, dV, dl, f</i>	0,9960
<i>dSL, dVL</i>	1,0000	<i>dS, dVL, f</i>	0,9968	<i>dS, dSL, dV, dl</i>	1,0000	<i>dSL, dV, dIL, f</i>	0,9944
<i>dSL, dl</i>	1,0000	<i>dS, dl, dIL</i>	1,0000	<i>dS, dSL, dV, dIL</i>	1,0000	<i>dSL, dVL, dl, dIL</i>	1,0000
<i>dSL, dIL</i>	1,0000	<i>dS, dl, f</i>	0,9968	<i>dS, dSL, dV, f</i>	0,9960	<i>dSL, dVL, dl, f</i>	0,9960
<i>dSL, f</i>	0,9848	<i>dS, dIL, f</i>	0,9968	<i>dS, dSL, dVL, dl</i>	1,0000	<i>dSL, dVL, dIL, f</i>	0,9944
<i>dV, dVL</i>	1,0000	<i>dSL, dV, dVL</i>	1,0000	<i>dS, dSL, dVL, dIL</i>	1,0000	<i>dSL, dl, dIL, f</i>	0,9960
<i>dV, dl</i>	1,0000	<i>dSL, dV, dl</i>	1,0000	<i>dS, dSL, dVL, f</i>	0,9968	<i>dV, dVL, dl, dIL</i>	1,0000
<i>dV, dIL</i>	1,0000	<i>dSL, dV, dIL</i>	1,0000	<i>dS, dSL, dl, dIL</i>	1,0000	<i>dV, dVL, dl, f</i>	1,0000
<i>dV, f</i>	0,9776	<i>dSL, dV, f</i>	0,9912	<i>dS, dSL, dl, f</i>	0,9968	<i>dV, dVL, dIL, f</i>	0,9944
<i>dVL, dl</i>	1,0000	<i>dSL, dVL, dl</i>	1,0000	<i>dS, dSL, dIL, f</i>	0,9968	<i>dV, dl, dIL, f</i>	0,9968
<i>dVL, dIL</i>	0,9992	<i>dSL, dVL, IL</i>	1,0000	<i>dS, dV, dVL, dl</i>	1,0000	<i>dVL, dl, dIL, f</i>	0,9984
<i>dVL, f</i>	0,9880	<i>dSL, dVL, f</i>	0,9920				

Tabla 6.10. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-FFT

Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba
<i>dS</i>	0,8582	<i>dl, dIL</i>	0,9038	<i>dSL, dl, dIL</i>	0,9976	<i>dS, dV, dVL, dIL</i>	1,0000
<i>dSL</i>	0,8606	<i>dl, f</i>	0,8990	<i>dSL, dl, f</i>	0,9351	<i>dS, dV, dVL, f</i>	0,9399
<i>dV</i>	0,9663	<i>dIL, f</i>	0,9231	<i>dSL, dIL, f</i>	0,9471	<i>dS, dV, dl, dIL</i>	0,9880
<i>dVL</i>	0,9663	<i>dS, dSL, dV</i>	0,9832	<i>dV, dVL, dl</i>	0,9880	<i>dS, dV, dl, f</i>	0,9303
<i>dl</i>	0,8558	<i>dS, dSL, dVL</i>	0,9712	<i>dV, dVL, dIL</i>	0,9639	<i>dS, dV, dIL, f</i>	0,9519
<i>dIL</i>	0,8486	<i>dS, dSL, dl</i>	0,9976	<i>dV, dVL, f</i>	0,9303	<i>dS, dVL, dl, dIL</i>	1,0000
<i>f</i>	0,8822	<i>dS, dSL, dIL</i>	0,9663	<i>dV, dl, dIL</i>	0,9952	<i>dS, dVL, dl, f</i>	0,9567
<i>dS, dSL</i>	0,8462	<i>dS, dSL, f</i>	0,9063	<i>dV, dl, f</i>	0,9495	<i>dS, dVL, dIL, f</i>	0,9447
<i>dS, dV</i>	0,9832	<i>dS, dV, dVL</i>	0,9784	<i>dV, dIL, f</i>	0,9591	<i>dS, dl, dIL, f</i>	0,9567
<i>dS, dVL</i>	0,9591	<i>dS, dV, dl</i>	0,9976	<i>dVL, dl, dIL</i>	0,9712	<i>dSL, dV, dVL, dl</i>	0,9976
<i>dS, dl</i>	0,9856	<i>dS, dV, dIL</i>	0,9615	<i>dVL, dl, f</i>	0,9495	<i>dSL, dV, dVL, dIL</i>	0,9399
<i>dS, dIL</i>	0,9928	<i>dS, dV, f</i>	0,9399	<i>dVL, dIL, f</i>	0,9663	<i>dSL, dV, dVL, f</i>	0,9279
<i>dS, f</i>	0,9159	<i>dS, dVL, dl</i>	0,9591	<i>dl, dIL, f</i>	0,9543	<i>dSL, dV, dl, dIL</i>	0,9832
<i>dSL, dV</i>	0,9663	<i>dS, dVL, dIL</i>	0,9880	<i>dS, dSL, dV, dVL</i>	0,9904	<i>dSL, dV, dl, f</i>	0,9543
<i>dSL, dVL</i>	0,9736	<i>dS, dVL, f</i>	0,9303	<i>dS, dSL, dV, dl</i>	0,9976	<i>dSL, dV, dIL, f</i>	0,9495
<i>dSL, dl</i>	0,9663	<i>dS, dl, dIL</i>	0,9856	<i>dS, dSL, dV, dIL</i>	0,9952	<i>dSL, dVL, dl, dIL</i>	0,9952
<i>dSL, dIL</i>	0,9976	<i>dS, dl, f</i>	0,9327	<i>dS, dSL, dV, f</i>	0,9591	<i>dSL, dVL, dl, f</i>	0,9423
<i>dSL, f</i>	0,9567	<i>dS, dIL, f</i>	0,9399	<i>dS, dSL, dVL, dl</i>	0,9952	<i>dSL, dVL, dIL, f</i>	0,9519
<i>dV, dVL</i>	0,9471	<i>dSL, dV, dVL</i>	0,9784	<i>dS, dSL, dVL, dIL</i>	0,9736	<i>dSL, dl, dIL, f</i>	0,9567
<i>dV, dl</i>	0,9976	<i>dSL, dV, dl</i>	0,9688	<i>dS, dSL, dVL, f</i>	0,9495	<i>dV, dVL, dl, dIL</i>	0,9615
<i>dV, dIL</i>	0,9712	<i>dSL, dV, dIL</i>	0,9976	<i>dS, dSL, dl, dIL</i>	0,9784	<i>dV, dVL, dl, f</i>	0,9351
<i>dV, f</i>	0,9615	<i>dSL, dV, f</i>	0,9471	<i>dS, dSL, dl, f</i>	0,9495	<i>dV, dVL, dIL, f</i>	0,9327
<i>dVL, dl</i>	0,9639	<i>dSL, dVL, dl</i>	0,9904	<i>dS, dSL, dIL, f</i>	0,9663	<i>dV, dl, dIL, f</i>	0,9519
<i>dVL, dIL</i>	0,9976	<i>dSL, dVL, IL</i>	1,0000	<i>dS, dV, dVL, dl</i>	1,0000	<i>dVL, dl, dIL, f</i>	0,9615
<i>dVL, f</i>	0,9255	<i>dSL, dVL, f</i>	0,9423				

Tabla 6.11. Resultados de precisión en la prueba para todas las combinaciones de uno a cuatro descriptores. Caso del localizador SVM-3F

En la tabla 6.12, se presentan los resultados promedio de cada combinación de descriptores, en caso de todas las faltas. Para el cálculo de los valores, se utilizó un promedio aritmético de los cuatro resultados. De esta tabla se puede apreciar, cuales combinaciones resultan ser más efectivas para la localización de todos los tipos de faltas.

Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba	Conjunto de descriptores	Precisión en la prueba
<i>dS</i>	0,9238	<i>dl, dIL</i>	0,9752	<b><i>dSL, dl, dIL</i></b>	<b>0,9990</b>	<i>dS, dV, dVL, dIL</i>	0,9980
<i>dSL</i>	0,8723	<i>dl, f</i>	0,9636	<i>dSL, dl, f</i>	0,9744	<i>dS, dV, dVL, f</i>	0,9797
<i>dV</i>	0,9765	<i>dIL, f</i>	0,9674	<i>dSL, dIL, f</i>	0,9747	<i>dS, dV, dl, dIL</i>	0,9968
<i>dVL</i>	0,9800	<i>dS, dSL, dV</i>	0,9958	<i>dV, dVL, dl</i>	0,9958	<i>dS, dV, dl, f</i>	0,9774
<i>dl</i>	0,9560	<i>dS, dSL, dVL</i>	0,9914	<i>dV, dVL, dIL</i>	0,9910	<i>dS, dV, dIL, f</i>	0,9815
<i>dIL</i>	0,9425	<i>dS, dSL, dl</i>	0,9978	<i>dV, dVL, f</i>	0,9743	<b><i>dS, dVL, dl, dIL</i></b>	<b>1,0000</b>
<i>f</i>	0,8825	<i>dS, dSL, dIL</i>	0,9910	<i>dV, dl, dIL</i>	0,9970	<i>dS, dVL, dl, f</i>	0,9821
<i>dS, dSL</i>	0,9599	<i>dS, dSL, f</i>	0,9697	<i>dV, dl, f</i>	0,9801	<i>dS, dVL, dIL, f</i>	0,9804
<i>dS, dV</i>	0,9958	<i>dS, dV, dVL</i>	0,9940	<i>dV, dIL, f</i>	0,9813	<i>dS, dl, dIL, f</i>	0,9837
<i>dS, dVL</i>	0,9888	<b><i>dS, dV, dl</i></b>	<b>0,9994</b>	<i>dVL, dl, dIL</i>	0,9926	<b><i>dSL, dV, dVL, dl</i></b>	<b>0,9994</b>
<i>dS, dl</i>	0,9959	<i>dS, dV, dIL</i>	0,9898	<i>dVL, dl, f</i>	0,9805	<i>dSL, dV, dVL, dIL</i>	0,9850
<i>dS, dIL</i>	0,9974	<i>dS, dV, f</i>	0,9798	<i>dVL, dIL, f</i>	0,9827	<i>dSL, dV, dVL, f</i>	0,9713
<i>dS, f</i>	0,9602	<i>dS, dVL, dl</i>	0,9898	<i>dl, dIL, f</i>	0,9793	<i>dSL, dV, dl, dIL</i>	0,9958
<i>dSL, dV</i>	0,9885	<i>dS, dVL, dIL</i>	0,9970	<i>dS, dSL, dV, dVL</i>	0,9970	<i>dSL, dV, dl, f</i>	0,9794
<i>dSL, dVL</i>	0,9908	<i>dS, dVL, f</i>	0,9776	<b><i>dS, dSL, dV, dl</i></b>	<b>0,9994</b>	<i>dSL, dV, dIL, f</i>	0,9789
<i>dSL, dl</i>	0,9908	<i>dS, dl, dIL</i>	0,9956	<i>dS, dSL, dV, dIL</i>	0,9984	<b><i>dSL, dVL, dl, dIL</i></b>	<b>0,9988</b>
<i>dSL, dIL</i>	0,9918	<i>dS, dl, f</i>	0,9792	<i>dS, dSL, dV, f</i>	0,9835	<i>dSL, dVL, dl, f</i>	0,9766
<i>dSL, f</i>	0,9719	<i>dS, dIL, f</i>	0,9799	<i>dS, dSL, dVL, dl</i>	0,9982	<i>dSL, dVL, dIL, f</i>	0,9792
<i>dV, dVL</i>	0,9868	<i>dSL, dV, dVL</i>	0,9905	<i>dS, dSL, dVL, dIL</i>	0,9934	<i>dSL, dl, dIL, f</i>	0,9812
<b><i>dV, dl</i></b>	<b>0,9994</b>	<i>dSL, dV, dl</i>	0,9892	<i>dS, dSL, dVL, f</i>	0,9821	<i>dV, dVL, dl, dIL</i>	0,9902
<i>dV, dIL</i>	0,9920	<b><i>dSL, dV, dIL</i></b>	<b>0,9990</b>	<i>dS, dSL, dl, dIL</i>	0,9946	<i>dV, dVL, dl, f</i>	0,9771
<i>dV, f</i>	0,9766	<i>dSL, dV, f</i>	0,9762	<i>dS, dSL, dl, f</i>	0,9819	<i>dV, dVL, dIL, f</i>	0,9722
<i>dVL, dl</i>	0,9910	<i>dSL, dVL, dl</i>	0,9974	<i>dS, dSL, dIL, f</i>	0,9852	<i>dV, dl, dIL, f</i>	0,9802
<i>dVL, dIL</i>	0,9962	<i>dSL, dVL, IL</i>	0,9946	<b><i>dS, dV, dVL, dl</i></b>	<b>0,9992</b>	<i>dVL, dl, dIL, f</i>	0,9829
<i>dVL, f</i>	0,9688	<i>dSL, dVL, f</i>	0,9751				

Tabla 6.12. Resultados de precisión promedio en la prueba para todas las combinaciones de uno a cuatro descriptores

### 6.6.5 Resultados del MBM

Como resultados del MBM se presentan, los obtenidos mediante el método propuesto por R. Das en las tablas 6.13 a 6.16. En la tablas sólo se muestra la distancia a la cual ocurre la falta, pero se debe tener en cuenta que ésta puede coincidir con múltiples zonas del circuito. La distancia que se resenta es la calculada desde el nodo de medición.

Para el caso presentado, se muestran las faltas que pueden coincidir con dos o más zonas, debido a la estimación de la distancia eléctrica desde la subestación y a la presencia de ramales en el circuito. En la última columna se presentan las zonas con las que coincide esta estimación.

Tipo de falta	Datos reales de la falta			Estimación de la falta	
	Resistencia [ $\Omega$ ]	Nodo	Distancia [millas]	Distancia [millas]	Zonas
Monofásica	0,05	826	22,222	23,066	Zona 4, Zona 5
	10	826	22,222	23,511	Zona 4, Zona 5
	25	826	22,222	24,178	Zona 4, Zona 5
	0,05	820	29,157	30,265	Zona 5, Zona 4
	10	820	29,157	30,731	Zona 5, Zona 4
	25	820	29,157	31,315	Zona 5, Zona 4
	0,05	844	35,095	36,253	Zona 7, Zona 8
	10	844	35,095	37,166	Zona 7, Zona 8
	25	844	35,095	37,903	Zona 7, Zona 8
	0,05	864	33,989	35,145	Zona 6, Zona 7, Zona 8
	10	864	33,989	36,436	Zona 6, Zona 7, Zona 8
	25	864	33,989	36,742	Zona 6, Zona 7, Zona 8

Tabla 6.13. Resultados de la pruebas con el MBM en caso de faltas a tierra

Tipo de falta	Datos reales de la falta			Estimación de la falta	
	Resistencia [ $\Omega$ ]	Nodo	Distancia [millas]	Distancia [millas]	Zonas
Bifásica	0,05	836	35,676	36,960	Zona 8, Zona 7
	10	836	35,676	38,138	Zona 8, Zona 7
	25	836	35,676	38,530	Zona 8, Zona 7
	0,05	844	35,095	36,569	Zona 7, Zona 8
	10	844	35,095	37,025	Zona 7, Zona 8
	25	844	35,095	37,868	Zona 7, Zona 8
	0,05	888	32,754	33,802	Zona 9, Zona 7, Zona 8
	10	888	32,754	34,850	Zona 9, Zona 7, Zona 8
	25	888	32,754	35,243	Zona 9, Zona 7, Zona 8

Tabla 6.14. Resultados de la pruebas con el MBM en caso de faltas bifásicas

Tipo de falta	Datos reales de la falta			Estimación de la falta	
	Resistencia [ $\Omega$ ]	Nodo	Distancia [millas]	Distancia [millas]	Zonas
Bifásica a tierra	0,05	836	35,676	37,174	Zona 8, Zona 7
	10	836	35,676	37,638	Zona 8, Zona 7
	25	836	35,676	38,494	Zona 8, Zona 7
	0,05	844	35,095	36,218	Zona 7, Zona 8
	10	844	35,095	37,341	Zona 7, Zona 8
	25	844	35,095	37,762	Zona 7, Zona 8
	0,05	888	32,754	33,956	Zona 9, Zona 7, Zona 8
	10	888	32,754	34,621	Zona 9, Zona 7, Zona 8
	25	888	32,754	35,407	Zona 9, Zona 7, Zona 8

Tabla 6.15. Resultados de la pruebas con el MBM en caso de faltas bifásicas a tierra

Tipo de falta	Datos reales de la falta			Estimación de la falta	
	Resistencia [ $\Omega$ ]	Nodo	Distancia [millas]	Distancia [millas]	Zonas
Trifásica	0,05	836	35,676	37,210	Zona 8, Zona 7
	10	836	35,676	37,781	Zona 8, Zona 7
	25	836	35,676	38,780	Zona 8, Zona 7
	0,05	844	35,095	36,253	Zona 7, Zona 8
	10	844	35,095	37,060	Zona 7, Zona 8
	25	844	35,095	37,832	Zona 7, Zona 8
	0,05	888	32,754	34,031	Zona 9, Zona 7, Zona 8
	10	888	32,754	34,588	Zona 9, Zona 7, Zona 8
	25	888	32,754	35,440	Zona 9, Zona 7, Zona 8

Tabla 6.16. Resultados de la pruebas con el MBM en caso de faltas trifásicas

### 6.6.6 Resultados del híbrido

Para este caso y a partir de los resultados obtenidos con cada método, se presenta la integración en el híbrido. Los resultados de la integración de los métodos se presentan en las tablas 6.17 a 6.20, para cada tipo de falta.

Tipo de falta	Resistencia de falta [ $\Omega$ ]	Nodo de falta	Localización real de la falta		Resultado del híbrido	
			Distancia [millas]	Zona	Distancia [millas]	Zona
Monofásica	0.05	826	22,222	Zona 4	23.066	Zona 4
	10	826	22,222	Zona 4	23.511	Zona 4
	25	826	22,222	Zona 4	24.178	Zona 4
	0.05	820	29,157	Zona 5	30.265	Zona 5
	10	820	29,157	Zona 5	30.731	Zona 5
	25	820	29,157	Zona 5	31.315	Zona 5
	0.05	844	35,095	Zona 7	36.253	Zona 7
	10	844	35,095	Zona 7	37.166	Zona 7
	25	844	35,095	Zona 7	37.903	Zona 7
	0.05	864	33,989	Zona 6	35.145	Zona 6
10	864	33,989	Zona 6	36.436	Zona 6	
25	864	33,989	Zona 6	36.742	Zona 6	

Tabla 6.17. Resultados de la pruebas con el método híbrido integrado. Caso de faltas monofásicas a tierra

Tipo de falta	Resistencia de falta [ $\Omega$ ]	Nodo de falta	Localización real de la falta		Resultado del híbrido	
			Distancia [millas]	Zona	Distancia [millas]	Zona
Bifásica	0.05	836	35,676	Zona 8	36.960	Zona 8
	10	836	35,676	Zona 8	38.138	Zona 8
	25	836	35,676	Zona 8	38.530	Zona 8
	0.05	844	35,095	Zona 7	36.569	Zona 7
	10	844	35,095	Zona 7	37.025	Zona 7
	25	844	35,095	Zona 7	37.868	Zona 7
	0.05	888	32,754	Zona 9	33.802	Zona 9
	10	888	32,754	Zona 9	34.850	Zona 9
25	888	32,754	Zona 9	35.243	Zona 9	

Tabla 6.18. Resultados de la pruebas con el método híbrido integrado. Caso de faltas bifásicas

Tipo de falta	Resistencia de falta [ $\Omega$ ]	Nodo de falta	Localización real de la falta		Resultado del híbrido	
			Distancia [millas]	Zona	Distancia [millas]	Zona
Bifásica a tierra	0.05	836	35,676	Zona 8	37.174	Zona 8
	10	836	35,676	Zona 8	37.638	Zona 8
	25	836	35,676	Zona 8	38.494	Zona 8
	0.05	844	35,095	Zona 7	36.218	Zona 7
	10	844	35,095	Zona 7	37.341	Zona 7
	25	844	35,095	Zona 7	37.762	Zona 7
	0.05	888	32,754	Zona 9	33.956	Zona 9
	10	888	32,754	Zona 9	34.621	Zona 9
25	888	32,754	Zona 9	35.407	Zona 9	

Tabla 6.19. Resultados de la pruebas con el método híbrido integrado. Caso de faltas bifásicas a tierra

Tipo de falta	Resistencia de falta [ $\Omega$ ]	Nodo de falta	Localización real de la falta		Resultado del híbrido	
			Distancia [millas]	Zona	Distancia [millas]	Zona
Trifásica	0.05	836	35,676	Zona 8	37.210	Zona 8
	10	836	35,676	Zona 8	37.781	Zona 8
	25	836	35,676	Zona 8	38.780	Zona 8
	0.05	844	35,095	Zona 7	36.253	Zona 7
	10	844	35,095	Zona 7	37.060	Zona 7
	25	844	35,095	Zona 7	37.832	Zona 7
	0.05	888	32,754	Zona 9	34.031	Zona 9
	10	888	32,754	Zona 9	34.588	Zona 9
25	888	32,754	Zona 9	35.440	Zona 9	

Tabla 6.20. Resultados de la pruebas con el método híbrido integrado. Caso de faltas trifásicas

### **6.6.7 Análisis de resultados**

A partir de la aplicación de los MBM a la localización de faltas en el sistema IEEE de 34 nodos, se puede apreciar que existe un error en la localización, propio del método, pero también asociado a la estimación de la zona o sección en la cual se presenta la falta. Las tablas 6.13 a 6.16 muestran que para todos los puntos seleccionados, donde se puede ver la sección o zona real de falta y las zonas identificadas por el método.

En las tablas 6.17 a 6.20, se presentan los resultados conjuntos, en los cuales se aprecia que el único error que aún permanece es el asociado a la estimación por el MBM.

### **6.7 Conclusiones**

Las estructuras híbridas planteadas muestran una alternativa de integración de los Métodos Basados en el Modelo - MBM y los Métodos de Clasificación Basados en el Conocimiento – MCBC, para resolver el problema de estimación del sitio de falta.

La integración de técnicas permite utilizar la información del tipo teórico con la de tipo empírico, disponible para la identificación de faltas en sistemas de distribución.

El principio utilizado se fundamenta en que se pueden definir zonas lo suficientemente pequeñas, que consideren solamente un circuito radial, de tal forma que no haya posibilidad de múltiple estimación dentro de la zona. A partir de la definición de las zonas y la estimación de la distancia mediante el método basado en la impedancia, se elimina el problema de la múltiple estimación

En caso que no se puedan definir zonas pequeñas, por la ausencia de datos o por el pobre desempeño de las SVM, de todas maneras se aprecia que la región de búsqueda se reduce considerablemente. Esto también favorece la reducción en el tiempo de restauración del circuito bajo falta.

La integración de las SVM como localizador de faltas, junto con un MBM, tal como el propuesto por R. Das, muestra una alta efectividad para la reducción del problema de la múltiple estimación.

## 7 Conclusiones, recomendaciones y trabajo futuro

### 7.1. Conclusiones generales

La calidad de la energía eléctrica es un tema de gran interés para los operadores de red y para los usuarios. La importancia actual de estos aspectos está asociada a que la participación del capital privado en el sector eléctrico obliga a establecer pautas de remuneración e índices mínimos de calidad. Entre los aspectos más importantes de la calidad se tienen en cuenta, la forma de onda, la continuidad del servicio y la atención al cliente.

Los índices de continuidad del suministro están directamente asociados con la localización de faltas. A partir de esto, se puede disminuir considerablemente el tiempo de restauración del sistema, reduciendo los índices de duración y de frecuencia de las interrupciones.

El problema de la localización de faltas en sistemas eléctricos no es nuevo. La localización de faltas en líneas de transmisión a partir de las medidas tomadas en ambos extremos de una línea homogénea está relativamente resuelto. Sin embargo, los esquemas regulatorios de distribución han establecido penalizaciones directas a las empresas, que junto con la nueva estructura competitiva del mercado eléctrico, han incentivado los estudios relacionados con la localización del sitio de falta en sistemas de distribución, donde el problema es complejo y aún no resuelto completamente.

Para solucionar el problema de localización de faltas, se han planteado metodologías que lo abordan desde muchas perspectivas. Algunas utilizan la inspección visual, otras las medidas del fundamental de tensión y corriente en la subestación, otras utilizan componentes de alta frecuencia y ondas viajeras, y las últimas, elementos de sensado de corrientes y tensiones ubicados en toda la red y telecomunicados con la subestación. Cada metodología tiene ventajas y desventajas asociadas principalmente a su precisión, a su coste de instalación, operación y mantenimiento, y al aprovechamiento que hace de la información disponible.

Adicionalmente a la detección de la falta, la tarea de estimación del sitio de ocurrencia se ha abordado principalmente usando métodos que se basan en la estimación de la reactancia de falta vista desde la subestación, y obtenida a partir de las medidas del fundamental de tensión y de corriente, que en este documento se conocen como Métodos Basados en el Modelo (MBM), tal como los que se presentan en el capítulo dos y se prueban en el capítulo cinco. Estos métodos han evolucionado a partir de aquellos que han sido implementados y probados con éxito en los sistemas de transmisión, y que se han adaptado a características tales como variedad en el calibre del conductor, presencia de cargas intermedias y laterales, medidas en un solo terminal de la línea y alta ramificación del sistema, entre otros, propias del sistema de distribución. Las grandes desventajas de estos métodos son dos: a) la necesidad de conocer los parámetros del sistema para tener buenos resultados en la estimación, y b) el problema de múltiple estimación del sitio de falta, asociada a la naturaleza altamente ramificada del sistema.

La alternativa propuesta en esta tesis posibilita el uso de información extraída sólo de las medidas de tensión y corriente registradas en la subestación, mediante la utilización de los métodos de clasificación o agrupamiento que se conocen en este documento como Métodos de Clasificación Basados en el Conocimiento (MCBC), que se presentan en el capítulo tres y se aplican al problema de localización en el capítulo cinco. Esta metodología se fundamenta en el establecimiento de una relación no lineal entre la entrada (señales caracterizadas), y la salida (la zona de la falta), a partir de una etapa de entrenamiento. Con un clasificador entrenado, se le puede presentar una nueva situación para la cual debe dar respuesta. En el caso de localización de faltas en sistemas de distribución, los clasificadores están posibilitados para identificar la zona en la cual ocurrió la falta. La principal desventaja de estos métodos, es que requieren mucha más información, no siempre disponible en las empresas distribuidoras. Por ello se propone la utilización de datos provenientes de simulación, para complementar los datos en la etapa de entrenamiento, mientras para la etapa de validación se usan los datos de faltas reales en el sistema.

Para la adecuada utilización de los MCBC que usan la información de las señales de tensión y corriente medidas en la subestación, éstas tienen que ser adecuadamente caracterizadas para obtener la información relevante asociada con la localización de la falta. A partir de esta caracterización, se obtienen los descriptores, que se utilizan como entrada a los métodos de clasificación, tal como se presenta en el capítulo cuatro.

Tanto los métodos que se fundamentan en la estimación de la impedancia (MBM), que usan los parámetros del sistema, como aquellos que se fundamentan en técnicas de clasificación (MCBC), y usan como información a los descriptores, han mostrado su validez para resolver el problema de localización de faltas. Las dos metodologías propuestas tienen ventajas y desventajas, pero apartir de la teoría de integración presentada, se han determinado las opciones de complementariedad, permitiendo así la formulación de híbridos para obtener mejores resultados, como se presenta en el capítulo seis. Existen varias posibilidades de integración, entre las cuales se resaltan las siguientes: a) Localización de la región probable de falta mediante los MCBC y selección de la distancia que coincide con esa zona, de las múltiples estimadas con un MBM; b) Determinación de la región probable de falta y trabajar con un MBM simplificado para hallar la falta sólo en esa región; c) Establecimiento de estructuras combinadas donde entradas y salidas intermedias se comparten entre los métodos, como se propone en el capítulo seis de esta tesis.

## 7.2. Conclusiones asociadas a los MBM

Existen muchos Métodos Basados en el Modelo (MBM), que permiten la estimación aproximada de la distancia eléctrica, desde la subestación hasta el sitio de falta, siempre que los parámetros de la red sean perfectamente conocidos. Aquellos métodos tienen en cuenta las cargas laterales, el acoplamiento mutuo, la presencia de laterales, la resistencia de falta, la presencia de múltiples calibres de conductor y el desbalance tanto del sistema como de la carga. Esto muestra la alta dependencia de un buen modelo del sistema de distribución y del conocimiento de todos sus parámetros. Por tanto, la exactitud de los métodos está estrechamente relacionada con el conocimiento de la red, tal como se muestra como aporte en las pruebas realizadas a diez métodos que se presentan en el capítulo cinco y en [MORA05-b] [MORA06-i].

Los MBM estiman una distancia eléctrica desde la subestación hasta el sitio de falta. Como los sistemas de distribución son altamente ramificados, la distancia eléctrica estimada, puede coincidir en varios sitios en los diferentes ramales del sistema. Este problema es conocido como el de múltiple estimación del sitio de falta, tal como se muestra en el capítulo seis de esta tesis.

Las propuestas de solución para el problema de la múltiple estimación están asociadas a medidores de tensión o registradores del paso de falta, que deben estar localizados en sitios estratégicos del sistema de distribución. Esta alternativa aumenta el coste de instalación y mantenimiento de la solución propuesta por estas metodologías, tal como se describe en el capítulo uno.

## 7.3. Conclusiones asociadas a los MCBC

El problema de localización de faltas en sistemas de distribución puede ser abordado por Métodos de Clasificación Basados en el Conocimiento (MCBC), conocidos genéricamente como “clasificadores”. Estos métodos tienen la capacidad de establecer una relación, normalmente no lineal, entre un conjunto de entrada (descriptores), y un conjunto de salida (tipo y región de la falta), tal como se presenta en el capítulo tres y aplica en los capítulos cinco y seis.

Los métodos de clasificación se caracterizan porque tienen dos etapas básicas: el entrenamiento y la prueba. El entrenamiento puede ser supervisado, no supervisado o semi-supervisado, y es la etapa donde se crea una relación entre la entrada y la salida, a partir de un conjunto de datos que sirve como ejemplos. La prueba consiste en la presentación de datos de entrada para que el clasificador obtenga una salida similar a la deseada, para así evaluar la calidad de la técnica y del entrenamiento ante este problema en particular. El error en la prueba se obtiene a partir de la estimación de un grado de afinidad entre la entrada y la salida. Para la realización de esta tesis se utilizó entrenamiento supervisado, tal como se presenta como aporte en [MORA06-a] [MORA06-b] [MORA06-e] y semi-supervisado, tal como se presenta en [MORA06-f].

Existen muchas técnicas de clasificación, cada una con sus ventajas y desventajas. Para la adecuada selección de una técnica en una aplicación en particular, se deben verificar al menos las siguientes características: tipo de atributos disponibles en la base de datos de entrenamiento y validación, interpretabilidad de los resultados, velocidad de aprendizaje y/o de reconocimiento, posibilidad de realizar aprendizaje secuencial y grado de aceptación de la técnica. Para el desarrollo de la tesis se utilizaron las técnicas LAMDA y las Máquinas de Soporte Vectorial (SVM), tal como se presentan en [MORA06-a] [MORA06-f].

Considerando el problema de localización de faltas, si se tiene una base de datos compuesta por registros de corriente y tensión, plenamente relacionada con el sitio de falta, se puede entrenar un clasificador para que realice la localización de la misma. Los registros de tensión y corriente se tienen que procesar para obtener lo que en este documento se conoce como descriptores.

#### 7.4. Conclusiones asociadas a la caracterización de señales de tensión y corriente

La aplicación de los MCBC, al problema de localización de faltas en sistemas de distribución, se fundamenta en la necesidad de aprovechar el conocimiento empírico disponible en la información registrada en la subestación. Esta información se obtiene en forma de descriptores, a partir del tratamiento de las señales de tensión y de corriente medidas en la subestación, antes, durante y después de la falta, tal como se presenta como aporte en [MORA05-c] [MORA06-g].

Un descriptor es una característica que contiene información significativa sobre un evento. En este caso particular, los descriptores que se proponen como aporte científico de esta tesis se obtienen de las señales de corriente y tensión. Así definidos para el caso de estudio propuesto, los descriptores están directamente asociados con la distancia a la falta y por tanto proporcionan la información necesaria para el aprendizaje de los MCBC usados.

Los descriptores básicos, utilizados para localización de faltas mediante los MCBC, son la frecuencia del transitorio, el valor máximo de la matriz de correlación de corrientes, la variación de la potencia aparente, la reactancia vista desde la subestación, la variación, y profundidad del hueco de tensión, y la variación de la magnitud de corriente. Los cinco últimos descriptores son calculados para valores de línea y fase.

La frecuencia está asociada directamente a la distancia eléctrica desde el punto de medida, hasta al nodo en falta. Esta característica permite discriminar la localización, sin importar el valor de la resistencia de falta.

Tanto la reactancia como la variación de la misma, permiten hallar la distancia a la falta. Sin embargo este descriptor no es completamente independiente de la resistencia de falta, debido a que se ve afectada por la reactancia de la carga.

Los huecos de tensión y las variaciones del valor eficaz de la corriente son descriptores que están asociados con la localización de la falta, aunque están fuertemente influenciados por la resistencia de la misma. Sin embargo, se ha demostrado en esta tesis que la combinación de estos descriptores contiene información que permite relacionar las señales de tensión y corriente con la localización de la falta.

La potencia aparente, expresada en términos de la variación entre los estados de falta y pre-falta, es un descriptor que es más estable que los dos anteriores. Esto le permite ser un poco menos dependiente de la resistencia de falta, conservando su relación con la distancia.

Finalmente, el máximo valor propio de la matriz de correlación de las muestras de corriente, es un descriptor único que está asociado con la distancia, pero que se muestra muy influenciado por la resistencia de la falta. Este descriptor aporta información para la localización, pero por si solo no es muy bueno como entrada de un clasificador.

Los descriptores se extraen tanto de las señales de línea como de las de fase, esto se debe a su relación directa con al distancia de la falta, tal como se plantea en los algoritmos de los relés de distancia para los diferentes tipos de faltas. Así, para las faltas monofásicas, los descriptores extraídos de las señales de fase en falta, muestran un comportamiento claramente marcado por la distancia. Para el caso de faltas bifásicas y bifásicas a tierra, aquellos descriptores obtenidos de las

señales de línea, representan mejor la relación del mismo con la distancia a la falta. Para el caso de faltas trifásicas, tanto los descriptores de fase como los de línea son adecuados.

Los descriptores usan información de cada fase, incluso aquellas que no están en falta, debido que esta información permite resolver el problema de la múltiple estimación, porque el comportamiento de las fases que no están en falta es diferente en cada circuito y ofrecen información adicional. Así mismo, las fases que no están en falta ayudan a diferenciar entre falta de baja impedancia en la parte final de un alimentador, y faltas de alta impedancia en la región cercana al punto de medida.

Los descriptores aquí propuestos no están influenciados por el ángulo de inserción de la falta, debido que la mayoría se calculan con estimadores de estado estable. Para el caso de análisis del transitorio, la frecuencia no se afecta por esta causa, aunque si la magnitud del mismo. Esta es la razón por la que en esta tesis se toma la frecuencia de las tres fases.

## 7.5. Conclusiones asociadas a la aplicación de los MBM y los MCBC

Los MBM localizan la distancia eléctrica desde el punto de medida hasta el sitio de falta tal como se presenta como aporte en [MORA06-i]. Adicionalmente los MCBC localizan una zona dentro del sistema de potencia en la cual se encuentra la falta como se aporta en [MORA06-b] [MORA06-e] [MORA06-f]. Estas características básicas muestran una alta complementariedad de las dos metodologías para resolver el problema de la múltiple estimación del sitio de falta.

La estimación de la distancia a la falta, obtenida a partir de los métodos basados en la impedancia (MBM), permite obtener una buena aproximación de la distancia entre el punto de medida (generalmente la subestación), y el nodo bajo falta. El principal problema de estos métodos está asociado a la disponibilidad de un buen modelo del sistema y a la múltiple estimación de la distancia, debido a la presencia de ramales que cumplen con el valor de impedancia o reactancia estimado, tal como se aporta en [MORA06-i].

Los MCBC usados son las técnicas LAMDA y SVM. Éstas son una buena opción de solución del problema de localización de faltas en los sistemas de distribución. La metodología se desarrolla con base en la información monitorizada en la subestación de cabecera de los circuitos o en simulaciones, tal como se presenta en [MORA06-c] [MORA06-d].

La técnica basada en SVM tiene mejor desempeño para resolver el problema de localización de faltas que la técnica LAMDA. Esto se debe principalmente a la presencia de los “kernels”, los cuales permiten transformar la dimensión del espacio de clasificación en uno en el cual las clases son separables linealmente. Los resultados presentados en [MORA06-e] y en [MORA06-f] validan esta afirmación y confirman el aporte propuesto.

Según las pruebas realizadas, se puede apreciar que el localizador de faltas basado SVM es una excelente alternativa. Los resultados obtenidos en esta tesis muestran una alta efectividad ante las diferentes circunstancias probadas con 63 combinaciones diferentes de descriptores para el sistema prototipo, como se presenta en el capítulo cinco.

Las SVM se probaron con dos estructuras diferentes, la dependiente del tipo de falta (DTF) y la no dependiente del tipo de falta (NDTF). En ambos casos se utilizó aprendizaje supervisado y con un “kernel” de función de base radial (RBF). Para la selección de la constante de penalización  $C$  y el parámetro  $\sigma$  se utiliza la búsqueda en malla y la validación cruzada. Los resultados obtenidos entrenando con aproximadamente el 20% del total de los datos y probando con el restante 80% muestran la potencialidad de las SVM y la adecuada selección de descriptores, en la aplicación como localizador de faltas.

Las diferentes combinaciones de descriptores probadas en el sistema prototipo muestran un excelente comportamiento para la localización de la zona en falta, reduciendo el problema de la múltiple estimación. Sin embargo, descriptores como el valor propio no presentan tan buenos resultados debido su alta variabilidad con la resistencia de falta, tal como se presenta en el aporte del capítulo cuatro y se corrobora en [MORA06-j]. Asimismo, el descriptor asociado a la reactancia de falta varía considerablemente con la carga del sistema, lo cual lo hace poco robusto para este tipo de aplicación, según se muestra en la parte final del capítulo cinco.

## 7.6. Conclusiones asociadas a los métodos híbridos

La mayoría de las actividades se fundamentan en la explotación de gran cantidad de información, hechos y experiencias, que corresponden a grandes grupos de conocimiento: el empírico y el teórico. Los MBM permiten la representación del conocimiento teórico y los MCBC permiten la representación y la adquisición del conocimiento empírico. A partir de esta concepción, se formulan los sistemas híbridos, con los cuales se busca la complementación de los dos grupos de conocimiento, de los que normalmente se dispone cuando se aborda un problema, tal como se propone como aporte en este documento y en [MORA05-d] [MORA05-e] [MORA06-a].

La estructura híbrida propuesta y validada en esta tesis muestra una alternativa de integración de los MBM con los MCBC, para resolver el problema de estimación del sitio de falta. El principio utilizado se fundamenta en que se pueden definir zonas lo suficientemente pequeñas para el MCBC, que consideren solamente un circuito radial, de tal forma que no haya posibilidad de múltiple estimación dentro de la zona. Con suficientes datos de entrada a la MCBC, a partir de la definición de las zonas y la estimación de la distancia con el MBM, se elimina el problema de la múltiple estimación.

La implementación de los MBM muestra que su precisión de la distancia a la falta, es muy buena para los sistemas presentados, pero la múltiple estimación se presenta cada vez que hay una derivación con una reactancia igual a la estimada para el ramal en falta. Los MCBC presentan una precisión perfecta para condiciones nominales, y por lo tanto un alto desempeño en la localización de la zona en falta. La complementación de estas estrategias permite la reducción del problema de múltiple estimación tal como se presenta para el sistema prototipo definido para esta tesis y el sistema de 34 nodos de la IEEE propuesto por el “Distribution System Analysis Subcommittee”.

El desempeño del método híbrido formulado en esta tesis, muestra cómo se pueden integrar dos técnicas que provienen de aproximaciones diferentes, para conformar un localizador de faltas. Éste brinda la información adecuada para la rápida restauración del servicio a partir de su aplicación a los sistemas de distribución, tal como se presenta en el capítulo seis de este documento. Los sistemas usados para la prueba son reales y reconocidos como prototipos de prueba en el campo de la investigación de problemas de la distribución de la energía eléctrica.

Las conclusiones presentadas en este capítulo justifican los aportes de la tesis, resumidos en el numeral 1.6.

## 7.7. Recomendaciones de aplicación

Algunas recomendaciones importantes que se deben seguir para el desarrollo de aplicaciones similares y que surgen a partir de la experiencia del autor, se presentan en esta sección.

El problema de la localización de faltas en distribución no es sencillo de abordar, principalmente por las características del sistema y por los escasos instrumentos de medición disponibles. El desempeño de los MBM y los MCBC aquí presentados está sujeto a la buena calidad del modelo en el primer caso, y a la disponibilidad de datos de faltas en el segundo. Estos datos deben ser pre-tratados de igual manera para las etapas de entrenamiento y la prueba, ya que son las entradas del MCBC y su desempeño está directamente asociado a la calidad de los descriptores.

Los MBM aquí implementados y probados, presentan un buen desempeño en el sistema prototipo. Para su implementación se requiere del conocimiento de los parámetros (resistencia, impedancia, capacitancia y longitud), de cada sección de línea, la localización de las cargas y la ubicación de los ramales. Se recomienda que en la codificación de los mismos, se debe considerar que la metodología para localizar la falta, debe repetirse para cada derivación del sistema, pues éstos hacen una múltiple estimación.

Los descriptores aquí presentados como uno de los principales aportes de esta tesis, están propuestos por su estrecha relación con la localización de la falta. Sin embargo, como ya se demostró, no todas las combinaciones son adecuadas para un sistema en particular. Por tanto, para la aplicación de los MCBC a un nuevo sistema, se debe plantear una estrategia que combine diferentes descriptores, con el fin de encontrar las mejores alternativas de precisión del método.

Los resultados obtenidos de clasificación para un mismo problema, con un conjunto determinado de descriptores y una pareja  $C$  y  $\sigma$ , pueden variar significativamente dependiendo de la normalización de los datos de entrada. Los descriptores o datos de entrada, para el caso de las SVM, se deben normalizar a un intervalo comprendido entre uno y menos uno. Adicionalmente, una recomendación fundamental es normalizar con los mismos valores a cada una de las ternas de descriptores de tensión, de corriente, de potencia y de reactancia, ya que solo así se mantiene la relación entre fases impuesta por el sistema de potencia.

Para la aplicación de los MCBC se requiere zonificar el sistema de distribución. En esta tesis se plantean varias alternativas para realizar esta zonificación, sin embargo siempre hay un compromiso entre el tamaño de las zonas, la disponibilidad de datos de falta y la precisión del localizador. Para un nuevo sistema, siempre se recomienda abordar el problema a partir de una zonificación grande y con muchos nodos, e ir reduciendo paulatinamente su tamaño hasta que los valores de precisión sigan siendo lo suficientemente altos como para ser aceptados como buenos.

Para la definición de la estructura híbrida aquí presentada, se propone una alternativa validada en otras aplicaciones. Para el caso particular de la localización de faltas, la estructura propuesta permite encontrar la zona y la distancia a partir de las SVM y el método propuesto por R. Das, respectivamente. Sin embargo, para la estimación de la distancia, se pueden usar métodos simplificados como el de la reactancia, ya que según las pruebas, han mostrado ser efectivos en sistemas con características similares a las que se muestran en las gráficas del comportamiento de los descriptores, presentadas en la parte final del capítulo cuatro.

## 7.8. Trabajo futuro

A partir de la investigación y las propuestas presentadas en este documento, se abren nuevos temas de investigación que sirven para complementar el aporte de esta tesis. Algunos de los más importantes se presentan en esta sección.

La metodología aquí presentada ha sido validada para dos sistemas de prueba. Sin embargo, como trabajo futuro se deben establecer convenios con entidades que suministren datos reales con información del tipo y localización de la falta, buenos modelos del sistema, para realizar las pruebas y los ajustes de la metodología propuesta. Este tipo de trabajos deben responder a la eterna pregunta sobre la diferencia entre las aproximaciones y limitaciones del modelado y su valor en el sistema real.

La aplicación de la estrategia de localización de faltas ayuda a determinar el sitio en la cual ésta ocurrió. Es claro que este tipo de propuestas contribuyen a disminuir los índices de duración y frecuencia equivalentes de las interrupciones. Sin embargo no hay cuantificaciones reales. Un estudio que se debe realizar es el de cuantificar económicamente y en términos reales al impacto de estas estrategias, para asociarlas a las características básicas de los circuitos como niveles de automatización, tamaño y tipos de cargas, entre otras.

A partir de la localización, se puede estructurar una estrategia operativa bajo condiciones de falta, que permita un restablecimiento del sistema, aislando rápidamente la zona bajo falta. Una tarea de alto grado de aplicabilidad está asociada al desarrollo de alternativas de configuración adecuada de elementos de protección y ubicación de mecanismos seccionalizadores en el circuito de distribución, que permitan aislar la zona falta en un tiempo corto. Esto permite el mantenimiento de la continuidad del suministro a los usuarios cercanos al nodo en falta.

Muchos de los MBM están fundamentados en la estimación de la reactancia hasta el sitio de falta. Sin embargo, el efecto de las compensaciones capacitivas no ha sido analizado dentro del desempeño de este tipo de aproximaciones. Si se considera que la reactancia capacitiva va a alterar la estimación basada en reactancia inductiva, reduciendo la distancia probable a la falta, el resultado no sería tan bueno como en casos de sistemas con sólo el componente inductivo.

La aplicación de los MCBC a un circuito en particular, está sujeta a una etapa de entrenamiento con datos del mismo circuito. Un análisis importante es la caracterización de los diferentes circuitos y la aplicación de un entrenamiento básico de un circuito con las mismas características y determinar así las mejores estrategias para lograr un buen desempeño.

Las SVM es una técnica que permite ser usada para clasificación, tal como se presenta en esta tesis. También pueden ser usadas como técnica de regresión, a partir de lo cual se estima que se puede encontrar la distancia de falta. De comprobarse la efectividad de esta propuesta, se utilizarán las SVM como estimadores de la distancia y de la zona.

Como trabajo posterior, se deba analizar la mejor alternativa de implementación física del localizador, en términos de requerimientos de “hardware” y “software”, satisfaciendo requisitos de alta confiabilidad técnica y también de bajo coste de implementación. Estos dos últimos aspectos favorecen el desarrollo de prototipos con mercado viable.

Como se presentó en el capítulo uno, el problema de localización de faltas en sistemas de transmisión no está completamente resuelto. Aunque existen medidas en ambos terminales de la línea, bajo circunstancias de faltas evolutivas y de alta impedancia aun se presentan errores en la estimación de la distancia. Una alternativa de aplicación viable es usar los MCBC, y utilizar los descriptores de ambos terminales para alimentar un localizador basado en SVM.

## Bibliografía

- [AGGA97] R. K. Aggarwal, Y. Aslan, and A.T. Johns. "An interactive approach to fault location on overhead distribution lines with load taps". IEE Developments in Power System Protection. Conference Publication No. 434.1997, pp 184-187.
- [AGUA98] J. C. Aguado. "A Mixed Qualitative-Quantitative Self-Learning Classification Technique Applied to Situation Assessment in Industrial Process Control". Ph. D. Thesis Universitat Politècnica de Catalunya, 1998.
- [AGUI80] J. Aguilar, M. Balssa, R. López. "Estimation recursive d'une partition. Exemples d'apprentissage et auto-apprentissage dans  $R^n$  et  $I^n$ ". Rapport technique 880I39, LAAS/CNRS, 1980.
- [AGUI82] J. Aguilar and R. López de Mántaras: "The process of classification and learning the meaning of linguistic descriptors of concepts". Approximate Reasoning in Decision Analysis. p. 165-175. North Holland, 1982.
- [AGRA98] R. Agrawal, J. Gehrke, D. Gunopulos, P. Raghavan. "Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications". In Proc. of the 1998 ACM-SIGMOD Conf. On the Management of Data, 94-105. 1998
- [ANCE94] G. Ancell, N. Pahalawaththa. "Maximum likelihood estimation of fault location on transmission lines using travelling waves" IEEE Transactions on Power Delivery, vol. 9, 1994. Pages. 680-689.
- [ANDE95] P. M. Anderson. "Analysis of Faulted Power Systems". Wiley-IEEE Press. 1995
- [ANGU01] C. Angulo. "Aprendizaje con máquinas de núcleo en entornos de multclasificación". Tesis Doctoral. Universidad Politècnica de Catalunya. 2001.
- [BALL65] G. Ball, D. Hall. "ISODATA, a novel method of data analysis and classification". Tech. Rep. Stanford University, Stanford, CA. 1965
- [BARR06] V. Barrera "Sistemas de distribución: Metodología para localización de fallas mediante la aplicación de inteligencia artificial". Disertación de Maestría en Ingeniería Eléctrica. Tutor. G. Carrillo. Universidad Industrial de Santander, Colombia. 2006
- [BENN92] K. Bennett, O. Mangasarian. "Robust linear programming discrimination of two linearly separable set" Optimizations methods and Software. 1992. Pags:23-24
- [BEZD81] J. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms". Plenum Press, New York, NY. 1981.
- [BIAN95] M. Bianchini, P. Frasconi, M. Gori. "Learning without local minima in radial basis function networks", IEEE Trans. Neural Networks, vol. 6, no 3, 749-756.1995
- [BILL84] R. Billinton, R. Allan. "Reliability Evaluation of Power Systems". New York, Plenum Press, 1984.
- [BLAC98] L. Blackburn. "Protective Relaying Principles and Applications", Marcel Dekker Inc. 1998.
- [BLUM98] A. Blum, T. Mitchell. "Combining labeled and unlabeled data with co-training. COLT": Proceedings of the Workshop on Computational Learning Theory, Morgan Kaufmann, 1998, p. 92-100.
- [BOLL93] M. Bollen. "Literature Search for reliability Data Of Components in Electric Distribution Networks". The Eindhoven University of Technology Netherlands . August 1993.
- [BOLL00] M. Bollen. "Understanding Power Quality Problems: Voltage Sags and Interruptions". IEEE Press. 2000.
- [BOSE92] B. Boser, I. Guyon, V. Vapnik. "A training algorithm for optimal margin classifier" Proc V ACM Workshop on Computational Learning Theory.1992. Pags: 144-152
- [BURG98] C. Burges. "A tutorial on support vector machines for pattern recognition". Data Mining and Knowledge Discovery, 1998. Pags: 121-167.
- [CHIE02] C. Chien; S. Chen; Y. Lin; "Using Bayesian network for fault location on distribution

- feeder” IEEE Transactions on Power Delivery Vol. 17, Issue 3, 2002 Page(s):785 – 793.
- [CHEN95] C. Cheng, “A branch-and-bound clustering algorithm”. IEEE Trans. Syst. Man Cybern. 25, 895–898. 1995.
- [CHOI04] M. Son Choi, S. Lee, D. Lee and B. Jin “A new fault location algorithm using direct circuit analysis for distribution systems”. IEEE Transactions on Power Systems. 2004, pages 35-41.
- [CORT95] C. Cortes, V. Vapnik. “Support vector networks”. Machine learning. 1995. Pags: 273-297
- [CORM06] J. Cormane. “Modelo estadístico para la localización de fallas en sistemas de distribución de energía eléctrica”. Disertación de Maestría en Ingeniería Eléctrica. Tutor. H. R. Vargas. Universidad Industrial de Santander, Colombia. 2006.
- [CREG02] Comisión de Regulación de Energía y Gas CREG, Resoluciones CREG 070 de 1998, CREG 096 de 2000, CREG 084 de 2002.
- [CROZ99] C. Crozier, W. Wisdom, “A power quality and reliability index based on customer interruption costs” Power Engineering Review, IEEE. Vol 19, Issue 4, April 1999. pages:59 – 61.
- [DAGE00] J. B Dagenhart. “The 40- Ground-Fault Phenomenon” IEEE Transactions on Industry Applications, vol. 36, no. 1, pages 30-32, 2000
- [DAS95] R. Das, M. Shadev, T Shidu. “A technique for estimating locations on shunt faults on distribution lines” Conferecnce proceedings. IEEE WESCANEX, 1. 1995. pp 6-11.
- [DAS98] R. Das. “Determining the locations of faults in distribution systems”. Canada, 1998, 206 p. Doctoral disertation. University of Saskatchewan Saskatoon.
- [DEMP77] A. Dempster, N. Laird, D. Rubin. “Maximum likelihood from incomplete data via the EM algorithm”. J. Royal Stat. Soc. B. 39, 1, 1–38. 1977.
- [DESR97] P. Desroches “Syclare: Systeme de classification avec apprentissage et reconnaissance des formes. Manuel d’utilization” Centre d’Etudis Avanzats de Blanes, 1997.
- [DIET95] T. Dietterich, G. Bakiri. “Solving multiclass learning problems via error-correcting output codes”. Journal of Artificial Intelligence Research. 1995. Pags:263-286.
- [DONG04] M. Dong, D. Xu, M. Li, Z. Yan. “Fault diagnosis model for power transformer based on statistical learning theory and dissolved gas analysis” IEEE International Symposium on Electrical Insulation, 2004. Pag: 85- 88
- [DUBO94] D. Dubois, H. Prade. “Fuzzy sets - a convenient fiction for modeling vagueness and possibility”. IEEE Transactions on Power Systems, 1994. Tome 2, Pags 16-21.
- [EMPT87] Leuven EMTP Center (LEC), Alternative Transient Program, ATP, Rule Book, 1987
- [ESTE96] M. Ester, H. Kriegel, J. Sander, X. Xu. “A density-based algorithm for discovering clusters in large spatial data sets with noise”. Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining. Portland, OR, pp. 226–231. 1996
- [EVER93] B. Everitt. “Cluster analysis”. Halsted Press, New York, 1974, 1993.
- [FRIE97] J.H. Friedman, “On bias, variance, 0/1 loss, and the curse-of-dimensionality”. Data Mining and Knowledge Discovery Journal, 1997. Tome 1: 55-77
- [FOGE93] D. Fogel, P. Simpson. “Evolving fuzzy clusters”. In Proceedings of the International Conference on Neural Networks (San Francisco, CA), pag: 1829–1834. 1993.
- [GAIN04] Z. Gaing. “Wavelet-Based Neural Network for Power Disturbance Recognition and Classification” IEEE Transactions on Power Delivery, Vol. 19. 2004. pp 1560-1568.
- [GANY05] V. Ganyun, C. Haozhong, Z. Haibao and D. Lixin. “Fault diagnosis of power transformer based on multi-layer SVM classifier” Electric Power Systems Research. Elsevier B.V. Vol. 74, Issue 1, April 2005, Pag: 1-7.
- [GARC04] V. García “Aplicación de la DWT en la detección e identificación de eventos de calidad de energía eléctrica”. Thesis. Universidad Industrial de Santander. 2004.
- [GE04] M. Ge, R. Du, G. Zhang, Y. Xu “Fault diagnosis using support vector machine with an application in sheet metal stamping operations”. Mechanical Systems and Signal Processing. Elsevier Ltd. Volume 18, Issue 1, January 2004, Pag: 143-159.

- [GELL02] C. Gellings, "Power delivery system of the future" Power Engineering Review, IEEE Volume 22, Issue 12, Dec. 2002 Page(s):7 – 12
- [GIAC92] P. Giacometti. "Modèles Hybrides de l'Expertise". Tesis doctoral. Lab. LIFIA - IMAG, Grenoble / ENST Paris - Francia, Noviembre. 1992
- [GIR93] A. Girgis, C. Fallon, D. Lubkerman "A fault location technique for rural distribution feeders" IEEE Transactions on Industry and Applications. Vol 26, 1993. Pages 1170-1175.
- [GOLD89] D. Goldberg. "Genetic Algorithms in Search, Optimization and Machine Learning". Addison-Wesley Publishing Co., Inc., Red-wood City, CA. 1989.
- [GOME05] Gómez A., Morales G. "Estudio e implementación de una herramienta basada en máquinas de soporte vectorial aplicada a la localización de fallas en sistemas de distribución". Trabajo para optar por el título de Ingeniero Electricista, dirigido por PhD. Herman Vargas. Universidad Industrial de Santander. 2005.
- [GREE91] A. Greewood. "Electrical Transients in Power Systems". Renselarch Polytechnic Institute, Second Edition. 1991
- [GUHA98] S. Guha, R. Rastogi, K Shim. "CURE: An Efficient Clustering Algorithm for Large Data sets", Published in the Proceedings of the ACM SIGMOD Conference. 1998
- [GUYO98] I. Guyon, J. Makhoul, R. Schwartz, V. Vapnik "What Size Test Set Gives Good Error Rate Estimates?". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, 1998. Pags:52-64
- [HAN01] J. Han, M. Kamber, A. Tung. "Spatial clustering methods in data mining: A survey." In Miller, H. and Han, J. (Eds.) Geographic Data Mining and Knowledge Discovery, Taylor and Francis. 2001.
- [HART75] J. Hartigan. "Clustering Algorithms". John Wiley & Sons, New York, NY. 1975.
- [HART79] J. Hartigan, and M. Wong,. "Algorithm AS136: A k-means clustering algorithm". Applied Statistics, 28, 100-108. 1979
- [HEYD01] G. T. Heydt "Power Quality Engineering". Power Engineering Review, IEEE , Volume: 21 Issue: 9 , Sept.2001 Page(s): 5 –7
- [HOLL75] J. Holland. "Adaption in Natural and Artificial Systems". University of Michigan Press, Ann Arbor, MI. 1975
- [HSU90] Y. Hsu, F. Liu, Y. Chien, J. Liu, J.. Lin, H. Yu, and R. Kuo, "An expert system for locating distribution system faults" in Proc. IEEE/PES Summer Meeting, July 1990, pp. 366–372.
- [HUAN02] S. Huang; "Application of immune-based optimization method for fault-section estimation in a distribution system". IEEE Transactions on Power Delivery, Vol. 17, Issue 3, July 2002 Page(s):779 – 784.
- [IEEE93] IEEE Task Force on Load Representation for Dynamic Performance "Load representation for dynamic performance analysis of power systems" IEEE Transactions on Power Systems Vol: 8, May 1993.
- [IEEE00] IEEE Distribution System Analysis Subcommittee. "Radial Test Feeders". 2000 <http://www.ewh.ieee.org/soc/pes/dsacom/testfeeders.html>.
- [IEEE04] IEEE Std C37.114 "IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines" Power System Relaying Committee 2004.
- [FU82], K. Fu "Introduction in Applications of Pattern Recognition", K.S. press. 1982
- [INVA03] F. Invancic "Modeling Analysis of Hybrid Systems". Thesis doctoral. University of Pennsylvania. 2003.
- [JAIN88] A. Jain, R. Dubes. "Algorithms for Clustering data". Prentice-Hall, Inc., 1988.
- [JAIN99] A. Jain, M. Murty, P. Flynn, "Data Clustering: A Review", ACM Computing Surveys, Vol.31, No3. 1999
- [JARV94] P. Jarventausta, P. Verho, and J. Partanen, "Using fuzzy sets to model uncertainty in fault location process of distribution networks," IEEE Trans. Power Delivery, vol. 9, pp. 954–960, Apr. 1994.

- [KAUF90] L. Kaufman, P. Rousseeuw. "Finding groups in data: an introduction to cluster analysis". John Wiley & Sons, 1990.
- [KARA99] G. Karypis, E. Han, V Kumar "CHAMELEON: A hierarchical clustering algorithm using dynamic modeling" *Computer* 32(8): 68-75, 1999.
- [KIRK83] S. Kirkpatrick, C. Gelat, M. Vecchi. "Optimization by simulated annealing". *Science* 220, 4598 (May), 671–680. 1983
- [KOH01] T. Kohonen. "Self-Organizing Maps". Springer Series in Information Sciences, 30, Springer. 2001
- [KOON75] W. Koontz, K Fukunaga, P Narendra. "A branch and bound clustering algorithm". *IEEE Trans. Comput.* 23, 908–914. 1975.
- [KEMP06] T. Kempowsky, A. Subias, J. Aguilar "Process situation assessment: from a fuzzy partition to a finite state machine" *Engineering Applications of Artificial Intelligence*. 2006. Vol.19, N°5, pp.461-477.
- [KUHN51] H. Kuhn, A Tucker. "Nonlinear programming". *Proceedings II Berkeley Symposium on Mathematical Statistics and probabilistics*,. Pags: 481-492. University of California Press. 1951
- [KRIS89] S. Krishnaswamy and A. St-Jacques. "A new fault location algorithm for radial transmission lines with loads". *IEEE Transactions on Power Delivery*.1989, pp 1676-1682.
- [LAKE89] E. Lakervi & E. J. Holmes, "Electricity Distribution network Design", Peter Peregrinus Ltd., 1989.
- [LEE04] S. Lee, M. Choi, S. Kang, B. Jin, D. Lee, B. Ahn "An intelligent and Efficient Fault Location and Diagnosis Scheme for Radial Distribution Systems *IEEE Transactions on Power Delivery*, Vol. 19, No. 2, April 2004. Pages 524-531.
- [LEUV87] Leuven EMTP Center (LEC). "Alternative Transient Program ATP", Rule Book, 1987
- [LEWI87] J. Lewis "Protective relaying-Principles and applications", Ed Marcel Dekker, Inc. New York 1987, pp 390-409.
- [LE5497] Ley 54 de 1997 "Ley del sector eléctrico". Comisión Nacional de Energía. Tercera edición 2001
- [LIM00] T. S. Lim, W. Y. Loh "A comparison of prediction accuracy, complexity and training time of thirty-three old and new classification algorithms". *Machine learning*, 2000, Tome 27(1)
- [LIN04] C. Lin, C. Chang, and C. Hsu, "A Practical Guide to Support Vector Classification". National Taiwan University, 2004.
- [LIU02] Y. Liu; N. Schulz, "Knowledge-based system for distribution system outage locating using comprehensive information". *IEEE Transactions on Power Systems*, Volume 17, Issue 2, May 2002 Page(s):451 – 456
- [LOON97] C. Looney "Pattern Recognition using Neural Networks: Theory and algorithms for engineers and scientist". Oxford University Press, Inc., New York, 1997.
- [LU78] S. Lu, K. Fu "A sentence-to-sentence clustering procedure for pattern analysis". *IEEE Transactions on Systems Man and Cybernetics*. 8, 381–389. . 1978
- [MACQ67] J. MacQueen. "Some methods for classification and analysis of multivariate data", *Proc. 5<sup>th</sup> Berkeley Symposium on Probability and Statistic*, University of California Press, Berkeley. 1967
- [MAHA04] R. Mahanty; P. Gupta; "Application of RBF neural network to fault classification and location in transmission lines" *IEE Proceedings Generation, Transmission and Distribution*, - Volume 151, Issue 2, 2 March 2004 Page(s):201 - 212
- [MART02] M. Martínez. "Análisis y medida de procesos no estacionarios en el dominio tiempo frecuencia" Tesis Doctoral. Universidad Politécnica de Madrid.2002
- [MART03] W. Martínez, A. Martínez. "Computacional Statistics Handbook with MatLab", Chapman & Hall/CRC, Boca Ratón, Florida, 2003.
- [MCLA88] G. McLachlan, T. Krishnan "The EM algorithm and extensions", JohnWiley & Sons. 1997

- [MINS88] M. Minsky, S. Papert, "Perceptrons" Expanded Edition, MIT Press, Cambridge, MA. 1988
- [MORA96] J. Mora, J. Manrique. "Estimación de magnitudes eléctricas utilizando la Transformada Discreta de Fourier". Proyecto de grado para obtener el título de Ingeniero electricista. Dirigido por G. Ordóñez. Universidad Industrial de Santander UIS. 1996
- [MORA03] J. Mora. "Voltage Sag Characterization and Classification for Diagnosis in Electric Power Quality Domain", Master Thesis. Directed by J. Colomer. University of Girona, Spain. 2003.
- [MORA04-a] J. Mora, P. Warren, J. Colomer "Hybrid Approach Based On Temporal Representation And Classification Techniques Used To Determine Unstable Conditions In A Blast Furnace", Current Topics in Artificial Intelligence (ISBN: 3-540-22218-9), May 2004. pag 125-136.
- [MORA04-b] J. Mora, J. Meléndez. "Caracterización de huecos de tensión para localización de fallas en sistemas de distribución - un primer paso hacia un método híbrido". Revista Scientia et Technica No 25. (ISSN 0122-1701). Noviembre de 2004. Pags 45-52
- [MORA05-a] J. Mora, C. Cortes, J. Gutiérrez "Técnica de localización de faltas para un sistema de potencia radial, con cargas laterales desequilibradas y circuitos no homogéneos" Revista Scientia et Técnica. N.28. (ISSN 0122-1701). Septiembre 2005, pages 56-62.
- [MORA05-b] J. Mora, S. Muñoz, G. Carrillo "Técnicas algorítmicas de localización de faltas como alternativa para reducir el efecto de las salidas en sistemas de potencia – Una revisión". III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov 2005.
- [MORA05-c] J. Mora, H. Serrano, G. Ordóñez "Mejoramiento de los Índices de Calidad del Servicio mediante la Caracterización de Señales para Localización de Fallas en Distribución". III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov 2005.
- [MORA05-d] J. Mora, J. Meléndez, G. Carrillo "Arquitectura Genérica para el Desarrollo de Sistemas Híbridos. Propuesta de localizador de fallas para mejorar los índices de calidad". III International Symposium of Power Quality. Colombia, Bogotá, CD proceedings. Nov 2005.
- [MORA05-e] J. Mora, M. Granada, L. Marín. "Métodos de representación del conocimiento en inteligencia artificial y su integración en sistemas híbridos de localización de fallas". Revista Ternura (ISSN 012-3921X), año 9, No. 17, II Semestre de 2005, pp 98 - 109
- [MORA05-f] J. Mora, J. Manrique, G. Ordóñez "Efecto del deslizamiento de frecuencia en la estimación de magnitudes eléctricas mediante la DFT". Revista Scientia et Technica N.27. (ISSN 0122-1701). Abril 2005. Pag 27-34.
- [MORA06-a] J. Mora, J. Rodríguez, S. Pérez. "Método híbrido basado en la estructura de agentes para localización de fallas en sistemas de distribución de energía eléctrica". Revista Scientia et Technica. No 29 de 2006. (ISSN 0122-1701). Mayo de 2006. Pags 59-65
- [MORA06-b] J. Mora, G. Carrillo, L. Pérez. "Fault location in power distribution systems using ANFIS nets and current patterns". In proceedings of 2006 IEEE PES Transmission and Distribution Conference and Exposition, Caracas, August, 2006
- [MORA06-c] J. Mora, J Melendez, J. Bedoya. "Extensive Events Database Development using ATP and Matlab to Fault Location in Power Distribution Systems". In proceedings of 2006 IEEE PES Transmission and Distribution Conference and Exposition, August, Caracas, 2006.
- [MORA06-d] J. Mora, J Meléndez, J. Bedoya. "Implementación de Protecciones y Simulación Automática de Eventos para Localización de Fallas en Sistemas de Distribución de Energía". Revista Ingeniería y Competitividad. Número 1, Agosto de 2006 (ISSN 0123-3033). Pag 43-51
- [MORA06-e] J. Mora, J Meléndez, G. Carrillo. "Fault Location in Power Distribution Systems Based on Signal Descriptors" In proceedings of 2006 IEEE 12th International Conference on Harmonics and Quality of Power, Porto, 2006.
- [MORA06-f] J. Mora, V. Barrera, G. Carrillo. "Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis". IEEE Transaction on Power Delivery. 2006 (Accepted paper)
- [MORA06-g] J. Mora, G. Olguin, H. Serrano. "Fault Location in Power Distribution Systems Using Characterization of Voltages and Currents". IEEE Transaction on Power Delivery. 2006 (RI. paper)

- [MORA06-h] J. Mora, J. Cormane, G. Ordóñez. "Improving of Power Continuity Indexes in Distribution Systems by using a Statistical Model for Fault Location". ETEP European Transactions on Power Systems. 2006 (R1. paper)
- [MORA06-i] J. Mora, J. Meléndez, G. Carrillo. "Comparison between power distribution system fault location methods". Electric Power System Research. 2006 (Submitted paper)
- [MORA06-j] J. Mora, J. Meléndez, G. Carrillo. "SVM based Fault Locator applied in Power Distribution Systems". IEEE Transaction on Power Delivery. 2006 (Submitted paper)
- [MURT83] F. Murtagh. 1983. A survey of recent advances in hierarchical clustering algorithms. Computer Journal, 26, 4, 354-359.
- [NOVO98] D. Novosel, D. Hart, J. Myllymaki "System for locating faults and estimating fault resistance in distribution networks with tapped loads". 1998. US Patent number 5,839,093.
- [OLGU05] G. Olguin "Stochastic Assessment of Voltage Dips Caused by faults in Large transmission Systems". Doctoral Thesis. Chalmers University of Technology. 2005.
- [OPPE97] A. Oppenheim, S. Willsky. S. Hamid Nawab "Signals and Systems", International Edition. Prentice Hall. 1997
- [PENG04] J. Peng, C.Chien; T. Tseng,"Rough set theory for data mining for fault diagnosis on distribution feeder" IEE Proceedings Generation, Transmission and Distribution, - Volume 151, Issue 6, 12 Nov. 2004 Page(s):689 – 697
- [PHAD95] A. Phadke, S. Horowitz "Power System Relaying". Research Studies Press. Baldock, England 1995.
- [PIER87] N. Piera. "Connectius de lògiques no estandard com a operadors d'agregació en classificació multivariable i reconeixement de formes". Doctorate dissertation at the Universitat Politècnica de Catalunya, 1987.
- [PIER89] N. Piera, P. Desroches, J. Aguilar. "An incremental conceptual clustering system. Report technique 89420, LAAS/CNRS, 1989.
- [RAVI77] B. Ravindranath and M. Chander. "Power System Protection and Switchgear". Wiley Eastern Ltd., New Delhi, India, 1977.
- [RD1955] Real Decreto 1955/2000. "Por el que se regulan las actividades de transporte, distribución, comercialización, suministro y procedimientos de autorización de instalaciones de energía eléctrica. Boletín Oficial del Estado de 27 diciembre 2000
- [RIVI99] J. Rivier "Calidad del servicio. Regulación y optimización de inversiones" Tesis doctoral. Instituto de Investigación Tecnológica. Universidad Pontificia Comillas. 1999.
- [RENC95] A. Rencher "Methods of Multivariable Analysis". Jhons Wiley and Sons, INC. Brigham Young University. Utah. 1995. Chapter 12, Pages 415-443
- [SAFA91] S. Safavian, D. Landgrebe. "A Survey of Decision Tree Classifier Methodology". IEEE Transactions on Systems, Man, and Cybernetics, Vol. 21, No. 3, pp 660-674, May 1991
- [SAHA02] M. Saha, E. Rosolowski. "Method and device of fault location for distribution networks. 2002". US Patent number 6,483,435 B2.
- [SAHD98] M. Shadev, R Agarwal. "A technique for estimating transmission line fault locations from digital impedance relay measurements" IEEE Transactions on Power Delivery, 3, N. 1, 1998, pp 121-129
- [SANT98] F. Santos "Un Système Hybride Neuro-Symbolique pour l'Apprentissage Automatique Constructif". Tesis Doctoral. Laboratoire LEIBNIZ / INPG, Grenoble – França. 1998.
- [SCHO02] B. Scholkopf, A. Smola. "Learning with kernels support vector machines, Regularization, Optimization and Beyond". The MIT Press, Cambridge. 2002
- [SCHW81] H. Schwefel. "Numerical Optimization of Computer Models". John Wiley and Sons, Inc., New York, NY. 1981
- [SEPA96] A. Sepälä "Load research and estimation in electricity distribution". Doctoral thesis. Technical research centre of Finland. 1996
- [SETH91] I. Sethi, A. Jain. "Artificial Neural Networks and Pattern Recognition: Old and New Connections". Elsevier Science Inc., New York, NY. 1991.

- [SHUI04] Y. Shui, G. Ming, Q. Hai, J. Lee, X. Yangsheng. "Intelligent diagnosis in electromechanical operation systems". Proceedings of IEEE International Conference on Robotics and Automation, 2004. ICRA '04. 2004. Milwaukee, WI, USA .Vol.3. Pag: 2267-2272.
- [SILV79] W. Silvert. "Symmetric summation: a class of operation in fuzzy sets. IEEE transactions in Systems, Man and Cybernetics, 1979. Issue 9. Pags 657-659
- [SHEI98] C. Sheikholeslami, S. Chatterjee, A.Zhang, "WaveCluster: A-MultiResolution Clustering Approach for Very Large Spatial Data set". Proc. of 24th VLDB Conference. 1998
- [SOUS02] L. Sousa, V. Fernao, V. Alegria. "A new accurate fault location method using  $\alpha\beta$  space vector algorithm". XIV Power System Control Conference. 2002
- [SOUS03] L. Sousa, J. Martins, V. Fernao, V. Alegria. "A network Distribution Power System Fault location based on Neural Eigenvalue Algorithm", IEEE Bologna Power Tech. 2003.
- [SRIN89] K. Srinivasan and A. St-Jacques. "A new fault location algorithm for radial transmission lines with loads". 1989, pp 1676-1682. IEEE Transactions on Power Delivery.
- [STEV75] A. Stevenson, D. William, "Elements of Power System Analysis". McGraw-Hill, Inc. 1975.
- [TAKA81] T. Takagi, Y. Yamakoshi, J. Baba, K. Uemura and T. Sakaguchi "A new algorithm of an accurate fault location for EHV/UHV transmission lines: Part—Fourier transform method," IEEE Transactions Power App. Syst., vol. PAS-100, Mar 1981. Pages. 1316–1323.
- [VAPN00] V. Vapnik "The nature of Statistical Learning Theory" Second Edition, Springer Verlag, 2000.
- [VIAW06] F. Viawan. "Steady State Operation and Control of Power Distribution Systems in the Presence of Distributed Generation". Thesis for the Degree of Licentiate Engineering. Chalmers University Of Technology. Göteborg, Sweden 2006.
- [WAIS00] J. Waissman. "Construction d'un modele comportemental pour la supervision de procedes: Aplication a une station de traitement des eaux" Doctoral thesis - Institut National Polytechnique de Toulouse, Pgs 17-28. 2000
- [WANG97] W. Wang, J. Yang and R. Muntz. " STING: A Statistical Information Grid Approach to Spatial Data Mining", Proceedings of the 23rd VLDB Conference Athens, Greece, 1997.
- [WARR68] A. R. Warrington, C. Van "Protective relays. Their theory and practice, Vol I, Chapman and Hall Ltd. London, 1968
- [WEIW02] Y. Weiwu, S. Huihe "Application of support vector machine nonlinear classifier to fault diagnoses" Proceedings of the 4th World Congress on Intelligent Control and Automation, 2002. vol.4, Pag: 2697- 2700.
- [WEST99] J. Weston, C. Watkins. "Multi-Class Support Vector Machines". Proceedings on VII European Symposium on Artificial Neural Networks, 1999. Pags: 219-224.
- [WEST64] Westinghouse Electric Corporation "Electrical Transmission and Distribution Reference Book". Pittsburgh Pennsylvania. USA 1964.
- [XU05] Y. Xu, L. Wang. "Fault Diagnosis System Based on Rough Set Theory and Support Vector Machine" Lecture Notes in Computer Science. Springer Berlin / Heidelberg. 2005. Pag: 980-988.
- [YANG98] L. Yang. "One terminal fault location system that corrects for fault resistance effects". 1998. US Patent number 5,773,980.
- [ZAHN71] C. Zahn "Graph-theoretical methods for detecting and describing gestalt clusters". IEEE Transactions on Computing. C-20 (Apr 1971..), 68–86.
- [ZHAN97] T. Zhang, R. Ramakrishnan and M. Linvy. (1997), BIRCH: An efficient data clustering method for very large data sets. Data Mining and Knowledge Discovery, 1(2): 141–182.
- [ZIMM80] H. Zimmermann, Zynso, P. "Latent connectives in human decision making" Fuzzy sets and systems, 1980. Tome 4. Pags: 37-51
- [ZHU97] J. Zhu, D. Lubkeman and A. Girgis. "Automated fault location and diagnosis on electric power distribution feeders". IEEE Transactions on Power Delivery 1997, pp. 801–809. .

## Anexo A. Sistema de pruebas prototipo

### A.1 Introducción

El sistema de distribución seleccionado para la realización de las pruebas de los diferentes métodos es un circuito de 25kV de Saskatoon Power and Light de la ciudad de Saskatoon, Canadá. Este circuito es presentado en [DAS98] y frecuentemente citado y utilizado por otros investigadores del problema de la localización de faltas.

### A.2 Diagrama Unifilar

En la figura A1 se presenta el diagrama unifilar del sistema prototipo.

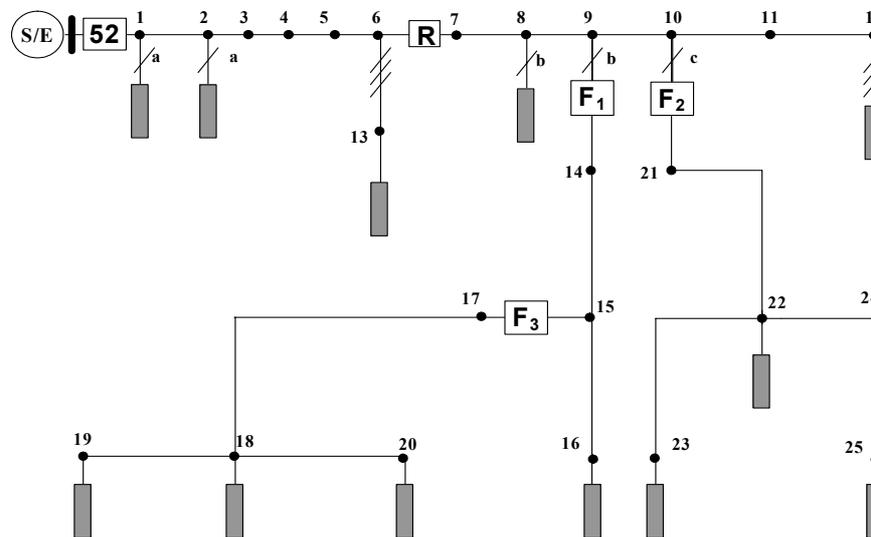


Figura A.1. Diagrama unifilar del sistema de 25 kV de Saskatoon Power and Light seleccionado para pruebas.

### A.3 Parámetros de las cargas

Los parámetros de las cargas se presentan en la tabla A.1.

Numero de Nodo	Fase	Carga conectada (KVA)	Composición [%]		
			Calentamiento	Iluminación	Motor
1	A	15.0	99.8	0.1	0.1
2	A	15.0	99.8	0.1	0.1
8	B	15.0	99.8	0.1	0.1
12	A,B,C	1000.0	0.1	0.1	99.8
13	A,B,C	67.5	99.8	0.1	0.1
16	B	15.0	99.8	0.1	0.1
18	B	15.0	99.8	0.1	0.1
19	B	7.5	99.8	0.1	0.1
20	B	15.0	99.8	0.1	0.1
22	C	25.0	99.8	0.1	0.1
23	C	15.0	99.8	0.1	0.1
25	C	15.0	99.8	0.1	0.1

Tabla A.1. Datos de las cargas del sistema de prueba.

### A.4 Parámetros de las líneas

Los parámetros de las líneas se presentan en la tabla A.2.

Sección	Longitud (km)	Impedancia serie (Ohms/km)		Admitancia paralelo (Mhos/km)	
		Secuencia + / -	Secuencia cero	Secuencia + / -	Secuencia cero
1-2	2.414	0.3480 + j0.5166	0.5254 + j1.704	j3.74E-6	j2.49E-6
2-6	16.092	0.3480 + j0.5166	0.5254 + j1.704	j3.74E-6	j2.49E-6
6-7	Reconector	-	-	-	-
7-8	4.023	0.3480 + j0.5166	0.5254 + j1.704	j3.74E-6	j2.49E-6
8-9	5.150	0.5519 + j0.5390	0.7290 + j1.727	j3.59E-6	j2.39E-6
9-10	2.414	0.5519 + j0.5390	0.7290 + j1.727	j3.59E-6	j2.39E-6
10-11	4.506	0.5519 + j0.5390	0.7290 + j1.727	j3.59E-6	j2.39E-6
11-12	2.414	0.3480 + j0.5166	0.7290 + j1.727	j3.74E-6	j2.49E-6
6-13	2.414	0.3480 + j0.5166	0.7290 + j1.727	j3.74E-6	j2.49E-6
9-14	Fusible				
14-15	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
15-16	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
15-17	Fusible				
13-15	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
15-16	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
17-18	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
18-19	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
18-20	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
10-21	Fusible				
21-22	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
22-23	2.414	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
22-24	3.219	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6
24-25	3.219	7.3977 + j0.8998	7.3977 + j0.8998	j2.51E-6	j2.51E-6

Tabla A.2. Datos de las líneas del sistema de prueba.

## **Anexo B. Resumen de algunos de los algoritmos de aprendizaje**

### **B.1 Introducción**

Las técnicas de clasificación usualmente se asocian a los métodos de aprendizaje utilizados, tres de los cuales se usan para el ajuste de los parámetros de un clasificador. Estos métodos son el aprendizaje supervisado, el no supervisado o agrupamiento (también conocido como aprendizaje sin instructor) y el agrupamiento semi-supervisado. En este anexo se presentan algunas de las técnicas más comúnmente usadas y referenciadas en diversos campos de aplicación para cada método de aprendizaje, según lo presentado en la figura 3.1.

### **B.2 Algoritmos de clasificación con aprendizaje supervisado**

Entre los algoritmos que utilizan aprendizaje supervisado se encuentran: Redes neuronales, árboles de decisión y modelos mixtos o mezclas finitas.

#### **B.2.1 Redes neuronales**

Las redes neuronales artificiales han sido utilizadas extensamente en aplicaciones de clasificación y agrupamiento [SETH91]. Una de sus principales ventajas radica en la capacidad de trabajar adecuadamente en problemas que presentan poca información previa y datos con ruido.

Las redes neuronales consisten de un conjunto de unidades de procesamiento enlazadas a través de pesos, los cuales son modificados durante un proceso de entrenamiento. Los datos para el entrenamiento están constituidos por varios pares de patrones de entrada y salida. El algoritmo compara la salida de la red y el valor esperado y ajusta los pesos y ganancias con el fin de minimizar el error que se presenta. Este proceso iterativo es el que entrena la red [LOON97].

Entre algunos de los tipos de redes que utilizan aprendizaje supervisado se encuentran: Redes perceptron multicapa y las redes de base radial.

##### **a. Redes perceptron multicapa**

Las redes perceptron de múltiples capas tienen una capa de entrada que contiene los nodos, varias capas ocultas y una capa de salida. La selección del número de neuronas en las capas ocultas depende de la aplicación en particular y de la complejidad computacional que se desee en el entrenamiento [MINS88].

Uno de los algoritmos utilizados para entrenar estas redes usa la retropropagación del error (Backpropagation), para ajustar los parámetros de la red. Los valores de las salidas se comparan con los correspondientes valores esperados, para calcular el valor de alguna función de error predefinida y ajustar algunos parámetros de la red. Esta red puede suministrar representaciones compactas de un grupo complejo de datos, sin embargo su entrenamiento es muy lento (consumo de tiempo), necesita una cantidad apreciable de datos de entrada/salida y en algunos casos puede caer en un mínimo local de la función error.

##### **b. Redes de base radial**

La arquitectura de estas redes se caracteriza por tener tres capas: una de entrada que contiene los nodos, una sola capa oculta con cada neurona con un tipo especial de función de activación, y una capa de salida. A diferencia de las redes perceptrón de múltiples capas, donde en las capas ocultas se calcula la suma ponderada de las entradas y se le aplica una función sigmoideal, las neuronas de las redes de base radial calculan la distancia euclídea entre el vector de pesos y la entrada, y sobre esta distancia se aplica una función de tipo radial con forma gaussiana [BIAN95].

Este tipo de redes tienen la ventaja de no caer en un mínimo local de la función de error, debido a que los únicos parámetros que son ajustados durante el proceso de entrenamiento son los mapeos lineales desde la capa oculta a la capa de salida.

### **B.2.2 Árboles de decisión**

Este tipo de aprendizaje se conoce como inductivo [SAFA91]. Las reglas se inducen a partir de los datos históricos disponibles, para lo cual procede a clasificar en la clase correspondiente diferentes objetos, basándose en el valor de las características o atributos que los definen. El resultado final es una estructura en forma de árbol que consiste de una gran cantidad de nodos, que representan las diferentes clases.

### **B.2.3 Modelos mixtos o mezclas finitas**

Un modelo mixto es la representación de una función objetivo como una combinación lineal de funciones de densidad de probabilidad. En algunos casos, la función objetivo es desconocida y representa las características de los datos. Los parámetros para cada componente de probabilidad de densidad (media, desviación estándar) son fijadas utilizando algoritmos iterativos tales como “Expectation Maximization - EM”, que permite determinar la función de membresía de los datos con respecto al modelo escogido.

### **B.2.4 Máquinas de soporte vectorial**

Los clasificadores de soporte vectorial están basados en hiperplanos que separan los datos de entrenamiento en dos subgrupos que tienen su propia etiqueta. Se fundamentan en el principio de que en medio de todos los posibles planos de separación entre las dos clases, etiquetadas como  $\{-1, +1\}$ , existe un único hiperplano óptimo de separación (OSH), de forma que la distancia entre el hiperplano óptimo y el patrón de entrenamiento más cercano sea máxima, con la intención de forzar la generalización de la máquina de aprendizaje [BURG98].

## **B.3 Algoritmos de clasificación con aprendizaje semi-supervisado**

El aprendizaje semi-supervisado utiliza para el entrenamiento, tanto datos de los cuales se les conoce la clase a la que pertenecen, como datos sin etiqueta o sin ningún tipo de información sobre su clasificación [BASU2004]. Entre algunas de estas metodologías, se encuentran aquellas que se basan en la adaptabilidad de la medición de la similitud buscando satisfacer algunas restricciones.

Entre algunas de las técnicas que utilizan este tipo de aprendizaje están las máquinas de soporte vectorial, las mezclas finitas ya explicadas y los métodos basados en grafos.

Los métodos basados en grafos definen un árbol donde los datos etiquetados y no etiquetados corresponden a los nodos y las ramas reflejan la similitud entre los datos.

## **B.4 Algoritmos de clasificación con aprendizaje no supervisado**

En el aprendizaje no supervisado, el proceso de agrupar datos con base en características similares se conoce como Agrupamiento. De acuerdo con [FU82], el agrupamiento se realiza en los siguientes pasos:

- a) Medición de la similitud: Inicialmente se divide el conjunto de datos en  $N$  subgrupos, de acuerdo con un criterio (por ejemplo, la distancia de cada elemento al centro del grupo).
- b) Prueba de la partición: Se realiza para determinar si los subgrupos son diferentes entre sí, o si se requieren nuevas particiones.
- c) Terminación: Normalmente se realiza cuando se alcanza el criterio de parada, por ejemplo: no se presentan cambios en los subgrupos.

Las diferentes técnicas de agrupamiento se definen de acuerdo con la forma como se realice cada uno de los pasos anteriores.

Existe mucha información sobre la utilización de técnicas de agrupamiento o “clustering” en áreas tales como la minería de datos, el análisis de patrones y las máquinas de aprendizaje, entre otras. A pesar que se encuentran varias técnicas de clasificación, es equivocado considerar que el

desempeño de un método es superior a otro bajo todas las circunstancias. Un método puede ser superior a otro, sólo en algún problema específico.

Dentro de las diversas clasificaciones de los métodos de agrupamiento encontradas, se destacan las realizadas por [EVER93][KAUF90][JAIN99][JAIN88] en las que se realizan análisis de estas metodologías de acuerdo con la estructura algorítmica utilizada y a su operación. Debido a la cantidad y diversidad de metodologías existentes, cualquier intento de categorizarlas sería incompleta.

Con base en las revisiones realizadas, la mayoría de las técnicas de agrupamiento se clasifican en dos grandes grupos: jerárquicas y de partición. Debido al auge de nuevas herramientas computacionales como redes neuronales y técnicas de optimización, el panorama se amplía a los algoritmos utilizados por las máquinas de aprendizaje.

Las metodologías jerárquicas agrupan datos en subgrupos en una forma estructurada, utilizando algún criterio de similitud; mientras que las de partición obtienen solo una división del conjunto de datos. Lo anterior hace que las metodologías jerárquicas sean mucho más versátiles ya que los datos pueden concentrarse en subgrupos no uniformes, bien distanciados unos de otros y la clasificación se realiza adecuadamente. De otra parte, los algoritmos de partición entregan buenos resultados en la medida en que los subgrupos sean de características similares, pero tienen la ventaja de que requieren un menor tiempo para entregar una respuesta. Durante el proceso, esta técnica realiza un ajuste continuo de las agrupaciones, lo que constituye una ventaja con respecto a las metodologías jerárquicas.

A continuación se presenta un resumen de algunos de los algoritmos utilizados como técnicas de agrupamiento.

#### **B.4.1 Agrupamiento jerárquico**

Se incluyen aquellos algoritmos que construyen gradualmente, una jerarquía de grupos de datos con características similares (agrupaciones), en forma de estructuras de árboles conocidos como dendogramas, para representar la interrelación entre las agrupaciones. Esta jerarquización se puede realizar iterativamente, en forma aglomerativa (cada patrón del conjunto de datos es definido inicialmente como una agrupación, y posteriormente se combinan sucesivamente), o en forma divisiva (se parte de una sola agrupación inicial, que contiene todos los datos y posteriormente se van desagregando en otras agrupaciones), hasta que un determinado criterio se cumpla. Generalmente el criterio de parada utilizado, es cuando se alcanza un número requerido de agrupaciones.

A diferencia de los métodos divisivos, la mayoría de los métodos jerárquicos pertenecen a la categoría de las aglomerativas. Entre ellas difieren sólo en el criterio que utilizan para definir la similitud entre agrupaciones.

Entre los principales criterios utilizados por estos algoritmos para determinar si las agrupaciones deben combinarse o dividirse, se encuentra el de la medición de la similaridad. Este concepto es fundamental para determinar cuan próximos o distantes se encuentran dos patrones,  $x_i$  y  $x_j$ . Entre los más populares está la distancia Euclidiana, utilizada para determinar la proximidad entre objetos en un espacio bidimensional o tridimensional. Algunos autores justifican su uso, siempre y cuando las agrupaciones de datos sean compactas y aisladas unas de otras.

Para la definición de la proximidad entre agrupaciones, en los algoritmos jerárquicos aglomerativos, se utilizan los conceptos de enlace simple, enlace completo y enlace promedio [MURT83].

La proximidad por enlace simple entre dos agrupaciones, está definida como la distancia más corta existente entre dos miembros cualesquiera que pertenecen a agrupaciones diferentes. Si la proximidad es por enlace completo, se toma como la más grande distancia entre dos miembros cualesquiera que pertenecen a agrupaciones diferentes. En la proximidad por enlace promedio, se toma como la distancia promedio entre dos miembros cualesquiera que pertenecen a agrupaciones diferentes.

La desventaja que presenta el enlace simple es la sensibilidad al ruido, y la del enlace completo es que puede tener problemas con las formas convexas.

Entre algunos de los algoritmos desarrollados que utilizan el agrupamiento jerárquico se encuentran:

a. **BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies)**

Este algoritmo propuesto en [ZHAN97], utiliza una estructura de datos jerárquica conocida como un árbol de pesos balanceados, en el cual se almacenan las características de las agrupaciones. Esta metodología es razonablemente rápida, ya que puede arrojar buenos resultados (buenas agrupaciones), con unos pocos barridos de los datos, lo cual lo hace una herramienta adecuada para ser utilizada con otros algoritmos. Sin embargo no arroja tan buenos resultados, cuando los “clusters” no presentan formas esféricas, debido a que utiliza el concepto de radio o diámetro para definir los límites de las agrupaciones.

b. **CURE (Clustering Using Representatives)**

Existen métodos basados en centroides/medoides (los medoides son objetos representativos de un grupo de datos), que utilizan un centro para representar una agrupación. El algoritmo propuesto por [GUHA98] utiliza un número constante de puntos representativos para definir una agrupación. Por tal razón esta metodología puede encontrar agrupaciones de tamaños y formas arbitrarias, tales como elipsoidales, espirales y cilíndricas. Su principal desventaja radica en que no puede ser aplicado directamente a grandes cantidades de datos.

c. **CHAMELEON**

Este algoritmo propuesto en [KARA99], determina las agrupaciones en el conjunto de datos, en dos etapas: Inicialmente genera un grafo que contiene enlaces entre un punto y sus vecinos más cercanos, posteriormente agrupa los datos en subgrupos y repetidamente utiliza un algoritmo aglomerativo para combinarlos.

## B.4.2 Agrupamiento por partición

En esta categoría se incluyen aquellos algoritmos en los cuales se parte de una agrupación inicial, que contiene todos los patrones e iterativamente se realiza y particiones en otras agrupaciones, hasta que se consigue optimizar una función. Esta metodología de agrupamiento, es recomendada cuando se cuenta con una gran cantidad de datos, ya que realizarlo jerárquicamente requería de mayor tiempo computacional.

Según la función objetivo que se utiliza en esta metodología, se definen los siguientes algoritmos:

a. **Algoritmos de error cuadrático**

Este algoritmo se encarga de determinar la diferencia que existe con respecto al centro de una agrupación, para cada uno de los patrones que pertenecen a la misma. Por esta razón, este algoritmo funciona bien para agrupaciones compactas pero separadas de otras agrupaciones.

Uno de los métodos que utilizan este algoritmo, es el *k-means* [MACQ67] [HART75] [HART79] [JAIN99], que inicia con una partición inicial del grupo de patrones en *k* agrupaciones definidas previamente, e iterativamente reasigna cada patrón al agrupamiento cuyo centro sea el más próximo a él, hasta que un criterio de convergencia es alcanzado (por ejemplo, es minimizada una función objetivo, para este caso la función error cuadrático). Entre sus principales desventajas se encuentra: la alta dependencia que tiene de la selección de la partición inicial ya que ésta puede llevar al algoritmo a que su función objetivo alcance un mínimo local.

Otro de los métodos es el *k-medoids* [KAUF90], en el cual cada grupo es representado por los objetos más cercanos al centro de gravedad. Una de las variantes de este algoritmo, se encuentra ISODATA [BALL65], que permite realizar divisiones y combinaciones de las agrupaciones resultantes.

**b. Algoritmos con agrupamiento difuso**

También conocido como agrupamiento no exclusivo. Tradicionalmente, las técnicas de agrupamiento generan particiones y en cada una de ellas, un patrón pertenece a solo un cluster. A esta técnica, se le conoce como agrupamiento *Hard*.

Cuando se desea definir el grado de pertenencia de cada uno de los patrones con respecto a cada una de las agrupaciones del sistema, se utiliza el agrupamiento difuso. Uno de los más utilizados ha sido el *Fuzzy C-Means* (FCM) [BEZD81] [BEZD84], el cual está basado en *k-means*. En este tipo de agrupamiento, cada agrupación es un grupo difuso de todos los patrones. Así, cada agrupación se representa por el patrón y el valor de membresía del mismo con respecto a esa agrupación.

La definición de las funciones de membresía para este tipo de aplicaciones se puede realizar con base en los centroides de cada agrupación.

**c. Agrupamiento por teoría de grafos**

Un grafo está constituido por nodos que representan los patrones a ser clasificados y por ramas que representan la relación entre los patrones. Un árbol de expansión de un grafo es un subgrafo, que conecta todos los vértices conjuntamente. Un simple grafo puede tener diferentes árboles de expansión. Para determinar cual es el mínimo árbol de expansión de un grafo, se le asigna a cada rama un peso que indica que tan favorable es la distancia entre un punto y otro, y se calcula la mínima suma de los pesos de las ramas de cada árbol para cada árbol de expansión.

Este tipo de agrupamiento está basado en la construcción del mínimo árbol de expansión de los datos [ZAHN71], en el cual se borran las ramas con los más altos pesos para generar las agrupaciones.

**d. Agrupamiento probabilístico**

Este tipo de algoritmos utiliza una de varias distribuciones paramétricas (generalmente Gausiana si es continua, o Poisson si es discreta) para representar al conjunto completo de datos como una mezcla de esas distribuciones [MCLA88]. El objetivo de esta metodología, consiste en determinar los parámetros de cada distribución y en lo posible su número. Entre ellos se encuentra el algoritmo Expectation Maximization [DEMP77], en el cual se asume un modelo de probabilidad con parámetros que describen la probabilidad de que un patrón pertenezca a determinada agrupación. El algoritmo LAMDA también es del tipo probabilística [WAIS00].

Como desventaja de este tipo de algoritmos está el de la alta complejidad computacional.

**B.4.3 Métodos basados en la densidad**

Estos algoritmos presentados en [HAN2001], consideran que las agrupaciones son regiones de datos altamente concentrados, que están separados por regiones menos densas. Así, las agrupaciones pueden ser identificadas al buscar regiones de alta densidad, llamadas modos. Estos modos, a su vez son asociados con el centro del agrupamiento y cada patrón es asignado al agrupamiento que presente el centro más cercano. Esto permite utilizar agrupaciones de forma arbitraria y con datos arbitrariamente distribuidos.

Uno de los métodos que utiliza este algoritmo es el DBSCAN (*Density Based Spatial Clustering of Application with Noise*), que parte de la idea de que en el vecindario de cada patrón de una agrupación, se encuentra un mínimo número de patrones. Como otros de las variantes de DBSCAN, se encuentran GDBSCAN, PDSCAN [ESTE96].

**B.4.4 Agrupamiento por malla**

Estos algoritmos dividen el espacio de agrupamiento en un número finito de celdas (hiperrectángulos), sobre los cuales desarrolla las operaciones requeridas. Las celdas que contienen cierto número de puntos son tratadas como densas y la conexión entre ellas forma a su vez una agrupación. Algunas implementaciones de esta categoría son: STING (*Statistical Information Grid approach*) [WANG97], WaveCluster (utiliza la transformación wavelet para transformar el espacio de características original) [SHEI98], CLIQUE (*Clustering In QUEst*) [AGRA98].

### **B.4.5 Agrupamiento vecino más cercano**

En esta técnica se utiliza el concepto de determinar la distancia de cada patrón con respecto a su vecino más cercano [LU78]. Es un procedimiento iterativo, en el cual a cada patrón no clasificado, se le asigna al agrupamiento más cercano a él, siempre y cuando la distancia entre ellos esté por debajo de un umbral preestablecido. El proceso continúa hasta que todos los patrones están clasificados o no ocurran nuevas clasificaciones.

### **B.4.6 Técnicas de agrupamiento utilizadas en máquinas de aprendizaje**

#### **a. Redes neuronales artificiales con aprendizaje no supervisado**

Las redes neuronales con aprendizaje no supervisado, son aquellas redes que cuentan con un conjunto de datos de entrenamiento que consiste sólo en patrones de entradas. La red aprende a adaptarse basada en las experiencias recogidas de los patrones de entrenamiento anteriores. Generalmente la utilización de estas redes se realiza para: Determinar estructura de datos, y/o comprimir, codificar y/o transformar datos.

En el caso de aprendizaje no supervisado, están las redes neuronales competitivas [JAIN96] que son utilizadas para agrupar datos de entrada. En este tipo de aprendizaje, las neuronas de salida compiten para tener el derecho a responder. La arquitectura de estas redes hace que el sistema se modifique para conseguir que las unidades más cercanas respondan similarmente. Como característica principal está la simplicidad en su estructura, debido a que poseen sólo una capa de neuronas. El número de neuronas está definido como el número de clases. Las redes competitivas son bastante eficientes para resolver problemas de clasificación, sin embargo presentan problemas, tales como la adecuada elección de una tasa de aprendizaje que permita hallar un punto de equilibrio entre velocidad de convergencia y estabilidad final de los vectores de peso.

Un ejemplo muy conocido es el de redes de Kohonen. Específicamente, para el diagnóstico de fallas, estas redes han sido utilizadas con los algoritmos de mapas auto\_organizativos SOM (Self-Organized Maps) [KOH90], que utilizan un aprendizaje no supervisado y el algoritmo de aprendizaje por vector de cuantización, que utiliza una mezcla de aprendizaje supervisado y no supervisado.

#### **▪ Mapas auto organizativos o redes de Coñeen**

El aprendizaje para estas redes es de tipo *off-line*, por lo que se distinguen dos etapas: una etapa de aprendizaje realizada con aprendizaje no supervisado y otra de funcionamiento con aprendizaje supervisado.

En la etapa de aprendizaje se fijan los valores de las conexiones entre la capa de entrada y la de salida. Esta red utiliza un aprendizaje no supervisado de tipo competitivo, las neuronas de la capa de salida compiten por activarse y sólo una de ellas permanece activa ante una determinada información de entrada a la red, los pesos de las conexiones se ajustan en función de la neurona que haya resultado vencedora.

En este modelo, el aprendizaje no concluye después de presentarle una vez todos los patrones de entrada, sino que este proceso es repetitivo buscando refinar el mapa topológico de salida, para que la red pueda realizar una clasificación más selectiva.

Como una desventaja que presentan estos mapas, se encuentra la alta dependencia de la selección inicial de los pesos, ya que pueden presentarse particiones de los datos en una forma no óptima. La convergencia de este algoritmo es controlado por parámetros tales como la tasa de aprendizaje y el vecindario del nodo ganador, en donde el aprendizaje se realiza.

#### **▪ Algoritmo de aprendizaje por vector de cuantización**

Este algoritmo es una versión supervisada del vector de cuantización. Estas redes son similares a SOM, excepto que la única capa de neuronas utiliza vectores de patrones, que se corresponden con la entrada. Las clases son predefinidas y se tiene un conjunto de datos etiquetados. El propósito es determinar un conjunto de prototipos que mejor representen a cada clase [KOH86].

Esta red es un híbrido que emplea tanto aprendizaje no supervisado, como aprendizaje supervisado para clasificación de patrones. En esta red, cada neurona de la primera capa es asignada a una clase, después cada clase es asignada a una neurona en la segunda capa. El número de neuronas en la primera capa, debe ser mayor o al menos igual que el número de neuronas en la segunda capa.

Al igual que con redes competitivas, cada neurona en la primera capa de esta red aprende un vector prototipo, el cual permite a la neurona clasificar una región del espacio de entrada. Sin embargo, en lugar de calcular la distancia entre la entrada y el vector de pesos por medio del producto punto, la red calcula la distancia directamente.

#### **b. Métodos basados en técnicas evolutivas**

Las técnicas evolutivas también han tenido su participación en el agrupamiento de datos, si se considera éste como un problema de optimización, que localiza los centroides óptimos de las agrupaciones más que el determinar una partición óptima.

Dentro de estas técnicas se encuentran las estrategias de evolución [SCHW81], la programación evolutiva [FOGE65] y los algoritmos genéticos [GOLD89][HOLL75]. Siendo estos últimos, los más frecuentemente utilizados en agrupamiento, debido a que ejecutan una búsqueda globalizada en comparación con otros procedimientos de agrupamiento que ejecutan una búsqueda localizada, tales como los algoritmos: *k-means*, agrupamiento difuso y redes neuronales, entre otros.

#### **c. Métodos basados en la búsqueda**

La utilización de los métodos basados en la búsqueda, se ha tomado como un problema de optimización en problemas de clasificación. Entre ellos se encuentra la técnica de búsqueda *Branch and Bound* [KOON75][CHEN95], que aplicado al agrupamiento, permite obtener particiones óptimas de los datos pero a un excesivo costo computacional.

El enfriamiento simulado (*simulated annealing*), es una técnica basada en un tratamiento térmico controlado, de calentamiento y enfriamiento de un material para incrementar el tamaño de sus cristales y reducir sus defectos [KIRK83]. Este algoritmo puede tardar un poco en alcanzar la solución óptima, debido a que aprovecha la primera ley de la termodinámica, en cuanto a que la variación lenta de la temperatura permite evitar pseudos-óptimos, esto hace que la obtención de la solución sea lenta.

## Anexo C. Tablas con resultados de las pruebas de los MBM para localización de faltas

### C.1 Introducción

En este anexo se presentan las tablas que complementan las presentadas en el capítulo 5, para las pruebas de los Métodos de localización de faltas Basados en el Modelo (MBM).

Las pruebas que se presentan pretenden mostrar el comportamiento de los MBM ante la variación de la distancia y el valor de la resistencia de falta. De los MBM analizados, hay que recordar que no todos están definidos para los cuatro tipos de faltas.

### C.2 Tablas de resultados

Los resultados de las pruebas de localización ante faltas bifásicas, se presentan en las tablas C1, C2 y C3.

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	<b>2,414</b>	<b>6,437</b>	<b>10,460</b>	<b>14,483</b>	<b>18,506</b>	<b>22,529</b>	<b>27,679</b>	<b>30,093</b>	<b>34,599</b>	<b>37,013</b>
A. Warrington [WARR68].	2.413	6.435	10.453	14.470	18.486	22.498	27.626	30.024	34.491	36.930
K. Srinivasan et al. [SRIN89]	2.414	6.438	10.460	14.482	18.503	22.521	27.669	30.080	34.585	37.039
A. Girgins et al [GIR93]	2.414	6.438	10.460	14.482	18.503	22.522	27.669	30.080	34.585	37.039
J. Zhu et al [ZHU97]	2.489	6.511	10.532	14.547	18.561	22.544	27.656	30.040	34.463	36.538
R. Das [DAS98]	2.408	6.436	10.457	14.483	18.508	22.517	27.677	30.113	34.598	37.032
D. Novosel [NOVO98]	2.380	6.345	10.306	14.267	18.226	22.182	27.459	29.925	34.521	36.931
L. Yang [YANG98]	2.377	6.336	10.290	14.241	18.189	22.131	27.429	29.909	34.533	36.934
Saha et al [SAHA02]	2.413	6.437	10.462	14.491	18.523	22.558	27.757	30.209	34.860	37.013
M. Choi et al [CHOI04]	2.414	6.437	10.460	14.481	18.502	22.521	27.669	30.080	34.585	37.041

Tabla C.1. Resultados de las pruebas –Falta bifásica a través de una resistencia de falta  $R_f=0,05\Omega$

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	<b>2,414</b>	<b>6,437</b>	<b>10,460</b>	<b>14,483</b>	<b>18,506</b>	<b>22,529</b>	<b>27,679</b>	<b>30,093</b>	<b>34,599</b>	<b>37,013</b>
A. Warrington [WARR68].	2.300	6.465	9.451	14.494	18.502	22.507	27.614	30.004	34.450	36.827
K. Srinivasan et al. [SRIN89]	2.315	6.488	9.472	14.537	18.550	22.620	27.950	30.194	34.722	37.140
A. Girgins et al [GIR93]	2.365	6.548	9.565	14.627	18.645	22.667	27.798	30.216	34.736	37.154
J. Zhu et al [ZHU97]	2.368	6.555	9.574	14.637	18.674	22.709	27.885	30.323	34.885	37.318
R. Das [DAS98]	2.343	6.449	10.465	14.482	18.533	22.576	27.829	30.127	34.610	37.055
D. Novosel [NOVO98]	2.231	6.346	9.288	14.268	18.223	22.175	27.421	29.875	34.441	36.793
L. Yang [YANG98]	2.332	6.451	9.418	14.394	18.356	22.314	27.608	30.087	34.705	37.062
Saha et al [SAHA02]	2.300	6.478	9.477	14.546	18.581	22.631	27.855	30.341	35.096	37.013
M. Choi et al [CHOI04]	2.203	6.384	9.367	14.439	18.465	22.494	27.654	30.082	34.609	37.023

Tabla C.2. Resultados de las pruebas –Falta bifásica a través de una resistencia de falta  $R_f=5\Omega$

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	3.236	7.240	11.229	15.242	19.233	23.143	28.210	30.556	34.926	37.013
K. Srinivasan et al. [SRIN89]	2.612	6.551	10.517	14.533	18.531	22.490	29.215	30.498	35.073	37.530
A. Girgins et al [GIR93]	2.868	6.978	11.074	15.196	19.218	23.219	28.324	30.760	35.323	37.788
J. Zhu et al [ZHU97]	2.891	7.003	11.101	15.225	19.330	23.365	28.613	31.069	35.683	38.162
R. Das [DAS98]	2.377	6.361	10.368	14.363	18.366	22.546	28.386	30.129	34.740	37.153
D. Novosel [NOVO98]	2.384	6.348	10.298	14.272	18.223	22.122	27.288	29.680	34.118	36.442
L. Yang [YANG98]	2.851	6.901	10.936	14.993	19.025	23.002	28.317	30.781	35.358	37.732
Saha et al [SAHA02]	3.259	7.325	11.385	15.474	19.565	23.624	29.048	31.730	36.968	37.013
M. Choi et al [CHOI04]	1.085	5.111	9.187	13.318	17.434	21.525	26.838	29.400	34.017	36.532

Tabla C.3. Resultados de las pruebas –Falta bifásica a través de una resistencia de falta  $R_f=25\Omega$ 

Los resultados de las pruebas de localización ante faltas bifásicas a tierra, se presentan en las tablas C4, C5 y C6.

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	2.414	6.434	10.454	14.471	18.487	22.499	27.625	30.025	34.489	36.878
A. Girgins et al [GIR93]	2.415	6.437	10.461	14.483	18.504	22.521	27.667	30.082	34.585	36.994
J. Zhu et al [ZHU97]	2.463	6.489	10.511	14.532	18.550	22.562	27.712	30.125	34.643	37.076
R. Aggarwal et al [AGGA97]	2.445	6.447	10.485	14.264	19.118	23.082	28.343	30.085	37.013	37.013
R. Das [DAS98]	2.411	6.432	10.456	14.479	18.504	22.529	27.688	30.096	34.560	37.004
D. Novosel [NOVO98]	2.380	6.344	10.308	14.269	18.228	22.183	27.457	29.927	34.519	36.880
L. Yang [YANG98]	2.378	6.337	10.293	14.243	18.191	22.133	27.429	29.912	34.532	36.887
Saha et al [SAHA02]	2.414	6.437	10.464	14.493	18.525	22.559	27.756	30.211	34.858	37.013

Tabla C.4. Resultados de las pruebas –Falta bifásica a tierra a través de una resistencia de falta  $R_f=0,05\Omega$ 

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	2.540	6.468	10.545	14.497	18.512	22.519	27.630	30.014	34.453	36.831
A. Girgins et al [GIR93]	2.600	6.550	10.652	14.629	18.654	22.675	27.801	30.215	34.729	37.150
J. Zhu et al [ZHU97]	7.174	11.497	15.892	20.034	23.854	27.233	32.351	34.693	40.457	42.306
R. Aggarwal et al [AGGA97]	2.445	6.447	10.485	14.264	19.118	23.082	28.343	30.085	37.013	37.013
R. Das [DAS98]	2.397	6.464	10.504	14.539	18.552	22.580	27.806	30.230	34.801	37.190
D. Novosel [NOVO98]	2.473	6.348	10.373	14.271	18.233	22.187	27.437	29.886	34.445	36.797
L. Yang [YANG98]	2.565	6.454	10.488	14.397	18.366	22.325	27.623	30.097	34.707	37.065
Saha et al [SAHA02]	2.541	6.481	10.575	14.549	18.591	22.642	27.870	30.351	35.098	37.013

Tabla C.5. Resultados de las pruebas –Falta bifásica a tierra a través de una resistencia de falta  $R_f=5\Omega$

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	3.241	7.259	11.270	15.271	19.279	23.223	28.279	30.615	34.992	37.013
A. Girgins et al [GIR93]	2.876	7.003	11.120	15.230	19.270	23.291	28.325	30.752	35.322	37.777
J. Zhu et al [ZHU97]	24.960	28.216	31.943	39.297	43.501	47.511	53.375	56.065	61.035	63.177
R. Aggarwal et al [AGGA97]	2.445	6.447	10.485	14.301	19.118	23.082	28.306	30.085	37.013	37.013
R. Das [DAS98]	2.171	6.462	10.444	14.489	18.645	22.629	27.768	30.132	34.911	37.348
D. Novosel [NOVO98]	2.389	6.369	10.341	14.303	18.272	22.207	27.365	29.745	34.189	36.500
L. Yang [YANG98]	2.856	6.922	10.978	15.023	19.073	23.087	28.391	30.844	35.425	37.787
Saha et al [SAHA02]	3.263	7.345	11.426	15.503	19.611	23.707	29.121	31.794	37.013	37.013

Tabla C.6. Resultados de las pruebas –Falta bifásica a tierra a través de una resistencia de falta  $R_f=25\Omega$ 

Finalmente, los resultados de las pruebas de localización ante faltas trifásicas, se presentan en las tablas C7, C8 y C9.

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	2.127	6.435	10.457	14.478	18.499	22.517	27.651	30.053	34.524	36.918
K. Srinivasan et al. [SRIN89]	2.128	6.436	10.459	14.481	18.503	22.522	27.669	30.083	34.585	36.993
A. Girgins et al [GIR93]	0.861	7.878	21.386	14.653	18.653	22.668	27.755	30.074	34.585	36.904
R. Das [DAS98]	2.305	6.193	10.205	14.217	18.368	22.395	27.457	30.014	34.422	36.900
D. Novosel [NOVO98]	2.374	6.346	10.314	14.281	18.248	22.215	27.509	29.986	34.593	36.960
L. Yang [YANG98]	2.543	6.338	10.297	14.255	18.210	22.163	27.479	29.970	34.605	36.966
Saha et al [SAHA02]	2.435	6.439	10.470	14.505	18.545	22.590	27.807	30.269	34.935	37.013

Tabla C.7. Resultados de las pruebas –Falta trifásica a tierra a través de una resistencia de falta  $R_f=0,05\Omega$ 

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
	2,414	6,437	10,460	14,483	18,506	22,529	27,679	30,093	34,599	37,013
A. Warrington [WARR68].	2.451	6.469	10.488	14.507	18.522	22.532	27.506	30.044	34.683	36.875
K. Srinivasan et al. [SRIN89]	2.444	6.472	10.501	14.530	18.558	22.584	27.598	30.168	34.888	37.105
A. Girgins et al [GIR93]	5.349	28.711	11.128	13.385	1.946	20.345	33.894	30.703	42.065	43.603
R. Das [DAS98]	2.273	6.253	10.257	14.264	18.273	22.290	27.371	29.877	34.325	36.797
D. Novosel [NOVO98]	2.383	6.351	10.322	14.293	18.260	22.224	27.130	29.960	34.642	36.892
L. Yang [YANG98]	2.479	6.460	10.442	14.422	18.396	22.366	27.322	30.175	34.922	37.165
Saha et al [SAHA02]	2.453	6.486	10.527	14.573	18.620	22.680	27.567	30.427	35.329	37.013

Tabla C.8. Resultados de las pruebas –Falta trifásica a tierra a través de una resistencia de falta  $R_f=5\Omega$

Método	Nodo bajo falta y distancia real desde la subestación de distribución ( Nodo 1) [km]									
	Nodo 2	Nodo 3	Nodo 4	Nodo 5	Nodo 6	Nodo 8	Nodo 9	Nodo 10	Nodo 11	Nodo 12
<b>A. Warrington</b> [WARR68].	2,451	6.469	10.488	14.507	18.522	22.532	27.652	30.045	34.495	36.875
<b>K. Srinivasan et al.</b> [SRIN89]	2.444	6.472	10.501	14.530	18.558	22.584	27.747	30.169	34.691	37.105
<b>A. Girgins et al</b> [GIR93]	5.349	28.711	11.128	13.385	1.946	20.345	1.675	30.715	42.414	43.603
<b>R. Das</b> [DAS98]	2.275	6.289	10.291	14.283	18.274	22.262	27.384	29.824	34.275	36.716
<b>D. Novosel</b> [NOVO98]	2.383	6.351	10.322	14.293	18.260	22.224	27.498	29.962	34.537	36.892
<b>L. Yang</b> [YANG98]	2.479	6.460	10.442	14.422	18.396	22.366	27.688	30.177	34.804	37.165
<b>Saha et al</b> [SAHA02]	2.453	6.486	10.527	14.573	18.620	22.680	27.933	30.430	35.199	37.013

Tabla C.9. Resultados de las pruebas –Falta trifásica a tierra a través de una resistencia de falta  $R_f = 25\Omega$

## Anexo D. Perfeccionamiento de las redes LAMDA para localización de faltas

### D.1 Introducción

A partir de los resultados obtenidos en las pruebas preliminares que se presentan en el capítulo 5 para el método LAMDA, en este anexo se muestra el resumen de un estudio posterior en el cual, mediante la definición de nuevos descriptores y una nueva metodología, se obtienen mejores resultados, aunque inferiores a los obtenidos por las SVM.

### D.2 Descriptores utilizados

Un descriptor es una característica de los huecos de tensión que permite su clasificación con respecto a otros huecos de tensión. En la figura C.1 se hace una representación gráfica de algunos de los descriptores utilizados durante el proceso de selección del grupo de descriptores para la clasificación de huecos de tensión en un circuito de distribución. La definición de cada uno de los descriptores se presenta en la tabla C1.[MORA03].

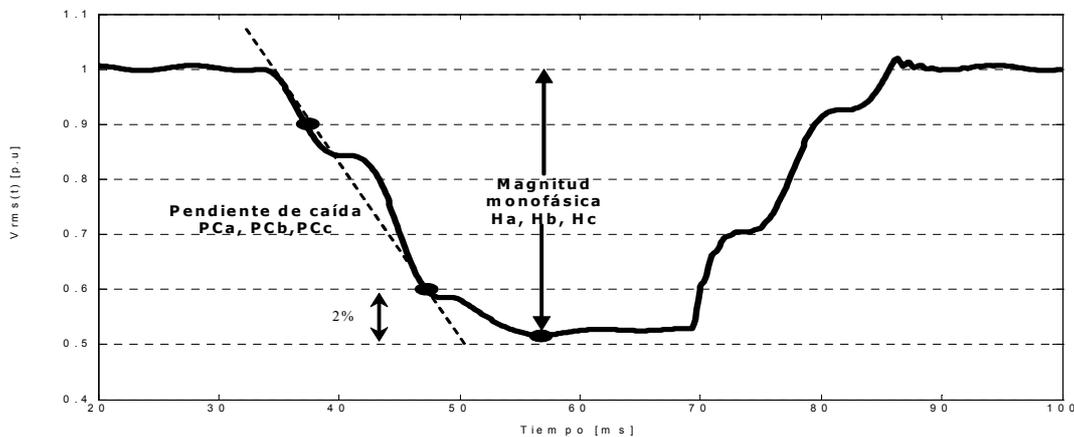


Figura D.1. Descriptores utilizados para el proceso de clasificación.

Descriptor	Descripción
Ha, Hb, Hc	Profundidad del hueco de tensión por fase [p.u]
Ia, Ib, Ic	Elevación de corriente por fase [p.u].
PCa,PCb,PCc	Pendiente de caída por fase.
maxH	Máxima profundidad de tensión monofásica [p.u].
maxI	Máxima corriente monofásica [p.u].
minPC	Mínima pendiente de caída monofásica.
$\bar{H}$	Media de las tres profundidades monofásicas [p.u].
$\bar{I}$	Media de las tres corrientes monofásicas [p.u].
$\bar{PC}$	Media de las tres pendientes de caída monofásica.
$\sigma H$	Desviación estándar de las tres profundidades monofásicas.
$\sigma I$	Desviación estándar de las tres corrientes monofásicas.
$\sigma PC$	Desviación estándar de las tres pendientes de caída monofásica

Tabla D.1. Descriptores utilizados para la localización de fallas

A partir de la tabla D.1, se define la tabla D.2, en la cual se presentan las posibles combinaciones de descriptores analizadas.

### D.3 Procesamiento de los descriptores

Como estrategia de análisis, se utilizaron para el procesamiento de los descriptores definidos en la tabla D.1 se usan dos herramientas ampliamente conocidas: la búsqueda de la proyección (PPEDA) y el análisis de componentes principales (PCA).

El PPEDA consiste en obtener diferentes proyecciones de los datos de manera aleatoria con el fin de encontrar aquella en la cual las clases de datos se encuentren claramente separadas. El grado de adecuación de la proyección es valorado a través de la maximización del índice de búsqueda de la proyección – *ppi* [MART03]. Con el objetivo de encontrar patrones que permitan clasificar las faltas con respecto a su localización dentro del sistema de distribución, se generaron observaciones para 36 distintos grupos de descriptores, y se obtuvieron las proyecciones para cada uno de ellos por medio de PPEDA.

El PCA permite analizar la representación de un conjunto de  $n$  observaciones de  $p$  variables, con un número menor de variables que son combinaciones lineales de las originales [RENC95]. A partir del uso de esta técnica se realiza un análisis para los grupos de descriptores con mayor índice *ppi*. Estos datos son los utilizados para el entrenamiento del LAMDA y se presentan en la tabla D.2.

Grupo	Descriptor						ppi
	1	2	3	4	5	6	
1	Ha	Hb	Hc	Ia	Ib	Ic	3,27
2	Ha	Hb	Hc	PCa	PCb	PCc	1,54
3	Ia	Ib	Ic	PCa	PCb	PCc	0,8
5	Ha	Hb	Hc	maxH			4,92
6	Ia	Ib	Ic	maxI			1,11
7	PCa	PCb	PCc	minPC			0,63
8	maxH	Ia	Ib	Ic			1,74
9	maxH	PCa	PCb	PCc			0,49
10	Ha	Hb	Hc	maxI			4,42
11	PCa	PCb	PCc	maxI			0,92
12	Ha	Hb	Hc	minPC			3,88
13	Ia	Ib	Ic	minPC			0,97
14	maxH	maxI	minPC				3,68
15	Ha	Hb	Hc	maxH	maxI	minPC	4,00
19	maxH	$\sigma H$	$\bar{H}$				4,32
20	maxI	$\sigma I$	$\bar{I}$				2,41
21	minPC	$\sigma PC$	$\bar{PC}$				1,70
21	maxH	$\bar{I}$	$\bar{PC}$				3,67
23	maxH	$\sigma I$	$\sigma PC$	minPC			1,43
24	maxI	$\sigma I$	maxH	$\bar{H}$			3,51
26	Ha	Hb	Hc	maxH	$\sigma H$	$\bar{H}$	9,13
27	Ia	Ib	Ic	maxI	$\sigma I$	$\bar{I}$	2,35
28	PCa	PCb	PCc	minPC	$\sigma PC$	$\bar{PC}$	1,28
29	Ha	Hb	Hc	maxI	$\sigma I$		2,13
30	Ha	Ia	Hb	Ib	Hc	Ic	1,65
35	minPC	maxI					2,54
36	minPC	maxH					3,84

Tabla D.2. Grupos de descriptores utilizados a partir de la definición de la tabla D.1.

### D.4 Entrenamiento de LAMDA

Las faltas utilizadas para el entrenamiento fueron simuladas al 5%, 50% y 95% de cada sección, mientras que las faltas de validación fueron simuladas al 27,5% y 72,5%. Las faltas se dividieron de esta manera porque permiten una mejor valoración de la capacidad de generalización de la red LAMDA, debido a que las faltas utilizadas en el proceso de validación tienen ubicaciones distintas a las utilizadas durante el proceso de entrenamiento. En la tabla D.3 se presenta el número de datos utilizados tanto para el entrenamiento como para la validación del sistema inteligente de cada uno de los cuatro tipos de falta.

Tipo de falta	Entrenamiento	Validación
Monofásica	297	198
Bifásica	216	144
Bifásica a tierra	216	144
Trifásica	72	48

Tabla D.3. Número de datos para entrenamiento y validación

El proceso de entrenamiento supervisado consiste en presentar los datos de faltas a cada una de las cinco redes basadas en la técnica LAMDA encargada de un tipo específico de falta. De esta manera, se acondicionaron los parámetros de cada una de las clases.

El número de clases de cada una de las cinco redes se determina partiendo del supuesto que cada sección del sistema es una clase, esto es, se consideró que cada una de las faltas de cada sección tiene un comportamiento diferente con respecto a las faltas de las secciones vecinas. Posteriormente, se agrupa en una misma clase cada una de las secciones que tienen un comportamiento similar, conformando finalmente lo que se denominara “zona en falta”.

Las pruebas se realizaron para los grupos de descriptores 5, 10, 12, 19 y 26. Los mejores resultados se obtienen para el grupo 26. En las figuras D.2 y D.3 se presenta en el eje vertical cada una de las clases seleccionadas por la red monofásica y bifásica. En el eje horizontal se presentan los datos de entrenamiento (297 monofásicos y 216 bifásicos) y tomando como una clase a cada una de las secciones. Considerando lo anterior, en la red monofásica (Figura D.2) hay presentes 17 posibles zonas de falta, mientras que en la red bifásica solo hay 8 (Figura D.3). La información de estas figuras se utiliza para decidir las secciones que se unen para conformar una nueva zona en falta.

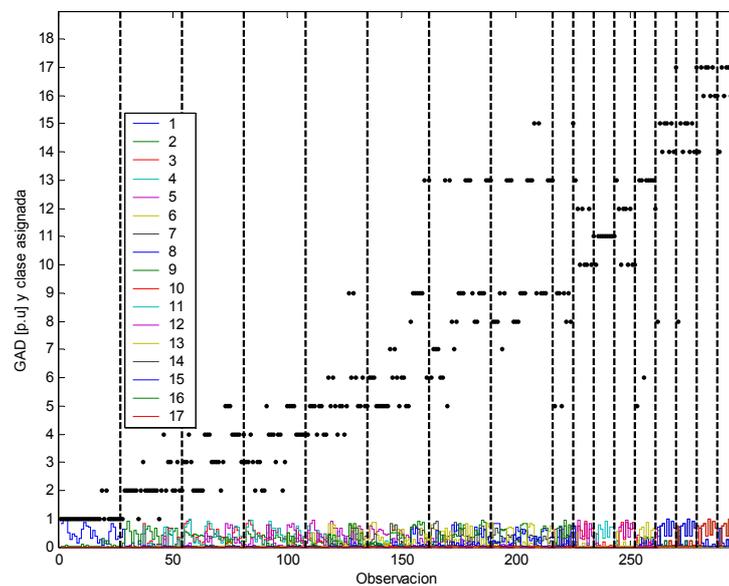


Figura D.2. GAD y sección asignada durante el proceso de entrenamiento de la red MONOFÁSICA basada en la técnica LAMDA. Grupo de descriptores No. 26 e impedancias de falta  $0,05\Omega$ ,  $5\Omega$  y  $10\Omega$ .

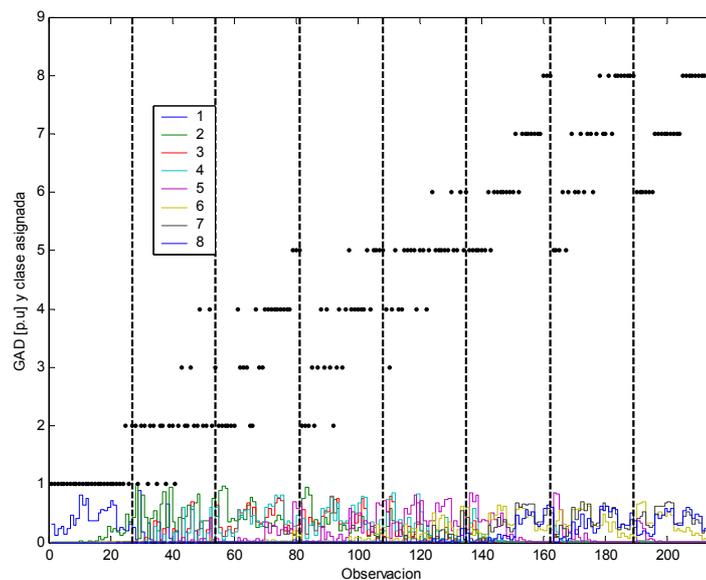


Figura D.3. GAD y sección asignada durante el proceso de entrenamiento de la red BIFÁSICA basada en la técnica LAMDA. Grupo de descriptores No. 26 e impedancias de falta  $0,05\Omega$ ,  $5\Omega$  y  $10\Omega$ .

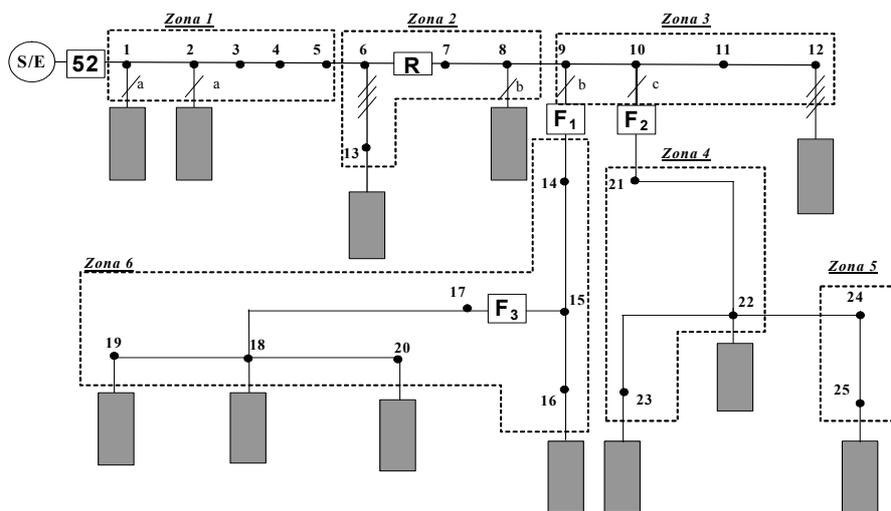
La clase 11 en los datos monofásicos de la figura D.2 tiene 100% de acierto en el entrenamiento, por tanto esta sección podría considerarse como una sola zona de falta. Las secciones 3, 4 y 5 de la misma gráfica por el contrario tienen un porcentaje de aciertos relativamente bajo, debido a que la red se confunde con las faltas de estas tres secciones, por lo cual es conveniente fusionarlas y conformar una sola.

Realizando el mismo análisis con los datos de faltas bifásicas de la figura D.3 y con los demás tipos de falta, es posible seleccionar cada una de las secciones que conformarán cada zona de falta, como se muestra en la tabla D.4.

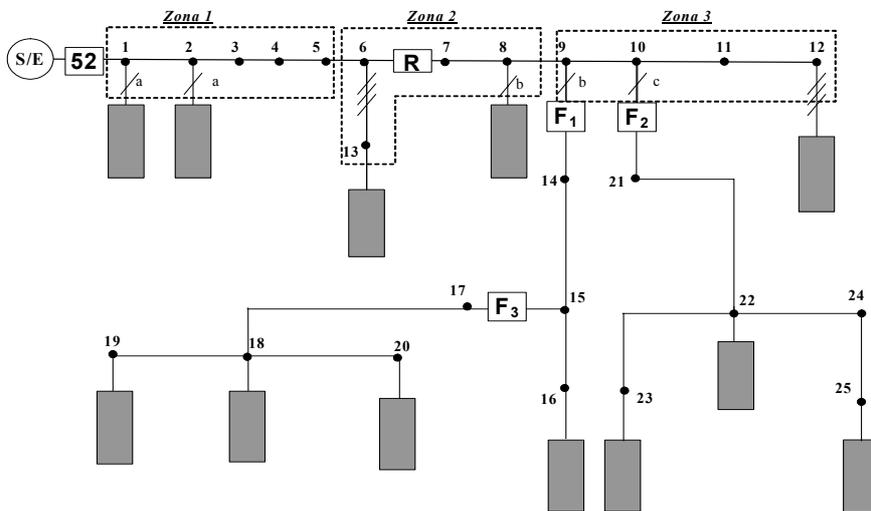
Zona en falta	Secciones por tipo de falta			
	Monofásica	Bifásica	Bifásica a tierra	Trifásica
1	1, 2		1, 2	
2	3, 4, 5		3, 4, 5	
3	6, 7, 8		6, 7, 8	
4	9, 10, 12			
5	11		No presenta esta zona	
6	13,14,15,16,17			

Tabla D.4. Secciones que conforman cada una de las zonas de falta

En la tabla 5.5 se nota que los tipos de falta bifásica, bifásica a tierra, trifásica y trifásica a tierra comparten las mismas zonas en falta, por consiguiente el sistema de distribución en estudio tendrá únicamente dos esquemas de posibles zonas en falta, uno para zonas monofásicas y otro para las demás. En la figura D.4 a) y b), se presentan los dos esquemas de zonas finalmente obtenidos.



(a)



(b)

Figura D.4. Zonas de falta. (a) Monofásicas (b) Bifásica, bifásica a tierra, trifásica, trifásica a tierra.

## D.5 Pruebas de precisión de LAMDA

En las figuras D.5, D.6 y D.7 se presentan los resultados de clasificación con los datos de validación de la red monofásica, bifásica y trifásica respectivamente. El número de zonas en falta y cada una de las secciones que la conforman están de acuerdo con lo consignado en la tabla D.4.

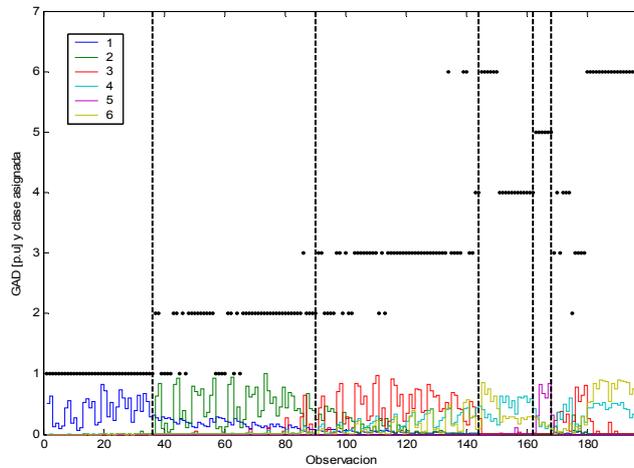


Figura D.5. GAD y zona de falta asignada durante el proceso de validación de la red MONOFÁSICA basada en la técnica LAMDA. Grupo de descriptores No. 26 e impedancias de falta  $0.05\Omega$ ,  $5\Omega$  y  $10\Omega$ .

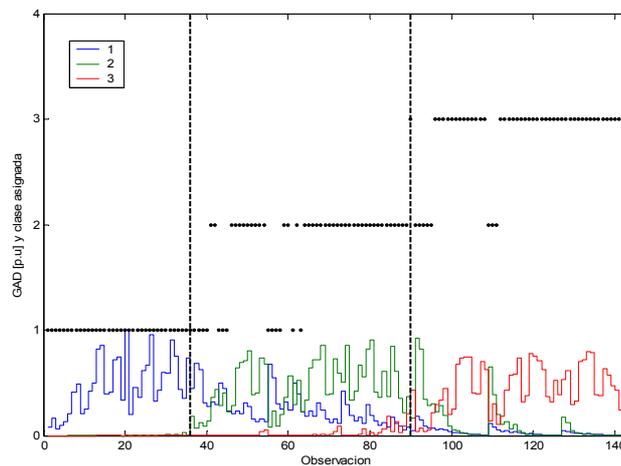


Figura D.6. GAD y zona de falta asignada durante el proceso de validación de la red BIFÁSICA basada en la técnica LAMDA. Grupo de descriptores No. 26 e impedancias de falta  $0.05\Omega$ ,  $5\Omega$  y  $10\Omega$ .

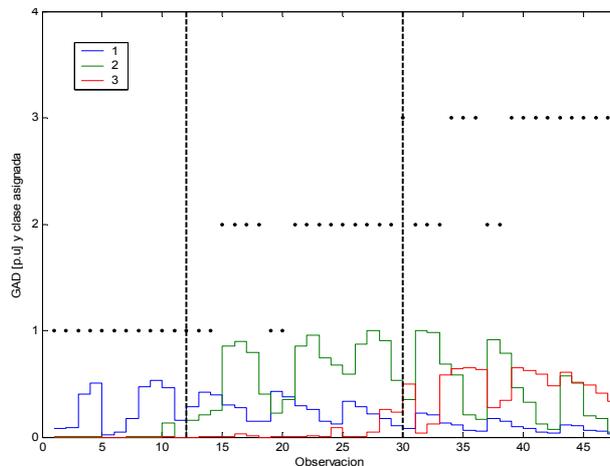


Figura D.7. GAD y zona de falta asignada durante el proceso de validación de la red TRIFÁSICA basada en la técnica LAMDA. Grupo de descriptores No. 26 e impedancias de falta  $0.05\Omega$ ,  $5\Omega$  y  $10\Omega$ .

En la parte inferior de cada una de las figuras anteriores se grafica el GAD de cada una de las zonas de falta, manifestándose un comportamiento “Gaussiano” con medias y desviaciones distintas con

respecto a las otras zonas de falta. Este comportamiento del GAD es muy importante porque permite confirmar que el clasificador es capaz de discriminar entre cada uno de los grupos de clasificación, que en este caso corresponde a las posibles zonas de falta.

En la tabla D.5 se presenta el número de aciertos de cada una de las redes durante el proceso de validación.

Zona de falta	Precisión por tipo de falta			
	Monofásica	Bifásica	Bifásica a tierra	Trifásica
1	36/36:100	36/36:100	36/36:100	12/12:100
2	41/54:76	40/54:74	39/54:72	13/18:72
3	40/54:74	46/54:85	43/54:80	13/18:72
4	12/18:67	No presenta esta zona de falta		
5	6/6:100			
6	19/30:63			
Totales	<b>154/198:78</b>	<b>122/144:85</b>	<b>118/144:82</b>	<b>38/48:79</b>

Tabla D.5. Resultados de validación de cada una de las cinco redes basadas en la técnica LAMDA

En la tabla D.5 se aprecia que utilizando los descriptores del grupo No. 26, los resultados obtenidos son buenos y que con la implementación de esta metodología es posible mejorar los índices de calidad referentes a la continuidad del suministro de energía eléctrica, como consecuencia de la reducción del tiempo de localización de la falta.

Finalmente, se concluye que la búsqueda de la proyección y el análisis de componentes principales son herramientas útiles para determinar el comportamiento de los diferentes tipos de falta en los sistemas de distribución. A través de las componentes principales de los datos de falta del grupo No. 26 (Magnitudes de los huecos de tensión por fase, magnitud máxima, desviación y promedio de las profundidades), es posible discriminar la ubicación de las faltas dentro de un sistema de distribución desbalanceado, y además, la fase o fases que intervienen en la misma.

## D.6 Pruebas adicionales.

En [BARR06] se presentan nuevas pruebas, a partir de el uso de los descriptores definidos en esta tesis y usando el pre-tratamiento con el PPEDA y PCA. Los descriptores utilizados en esta nueva prueba se presentan en la tabla D.6, mientras que los resultados se presentan en la tabla D.7, para la zonificación del sistema que se presenta en la tabla D.8.

Descriptor	Descripción
Ha, Hb, Hc	Profundidad del hueco de tensión por fase [p.u]
Ia, Ib, Ic	Elevación de corriente por fase [p.u].
Fa, Fb, Fc	Frecuencia del transitorio [Hz]
Sa, Sb, Sc	Variación de la potencia reactiva [VAR]
Vp	Valor propio máximo del la matriz de correlación de corrientes

Tabla D.6. Descriptores utilizados para la localización de fallas en las prueba adicional

Zona de falla	Monofásica	Bifásica	Bifásica a tierra	Trifásica
	[Aciertos]	[Aciertos]	[Aciertos]	[Aciertos]
1	352/360:98%	233/240:97%	221/240:92%	79/80:99%
2	25/50:50%	92/120:77%	97/120:80%	22/40:55%
3	20/40:50%			
	397/450:88%	325/360:90%	318/360:88%	101/120:84%
TOTAL	<b>1141/1290:88%</b>			

Tabla D.7. Resultados de validación de cada una de las 4 redes basadas en la técnica LAMDA para las observaciones del nuevo grupo de descriptores.

Zona de falla	Monofásica	Bifásica	Bifásica a tierra	Trifásica
	[barras]		[barras]	
1	1, 2,3,4,5,6,7,8,9,10,11,12		1,2,3,4,5,6,7,12	
2	13,14,15,16,17		8,9,10,11,12	
3	18,19,20,21			

Tabla D.8. Barras que conforman cada una de las zonas de falla

Según los resultados, se aprecia que son inferiores a los obtenidos con las SVM por dos razones principales. La precisión es inferior y la definición de un menor número de zonas.

## Anexo E. Sistema de pruebas de 34 nodos de la IEEE

### E.1 Introducción

El sistema de distribución seleccionado para la realización de las pruebas adicionales es el tomado de los “test feeders” del “Distribution System Analysis Subcommittee” del “Institute of Electrical and Electronics Engineers” [IEEE00].

Este alimentador está localizado en Arizona, y opera a una tensión nominal de 24.9 kV. Entre sus principales características están la presencia de ramales monofásicos y trifásicos, múltiples calibres de conductor, cargas desbalanceadas de naturaleza concentrada y distribuida, así como la presencia de bancos de capacitores.

En este anexo se presentan los datos utilizados para el modelado y simulación de este sistema.

### E.2 Diagrama Unifilar

En la figura E.1 se presenta el diagrama unifilar del sistema de 34 nodos de la IEEE.

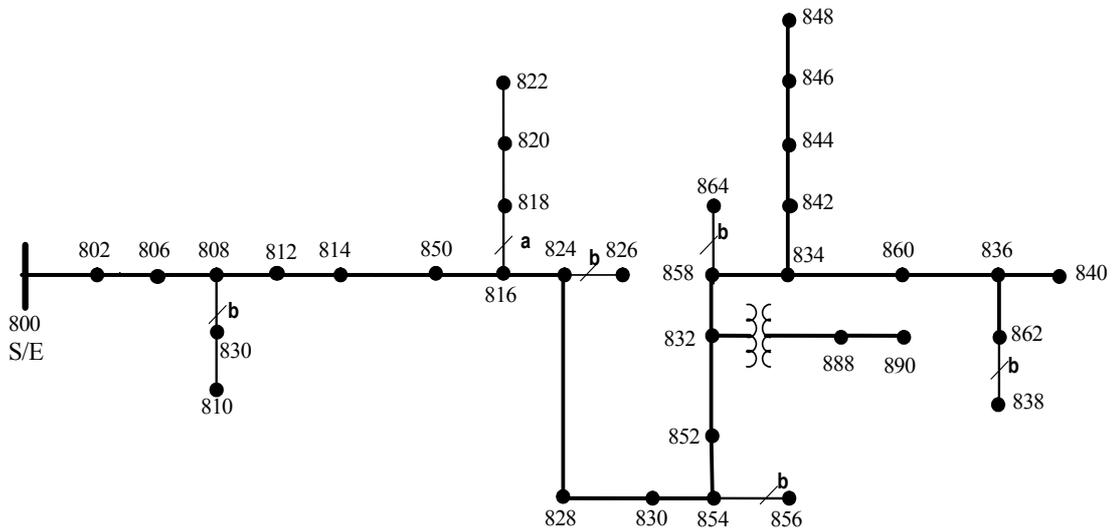


Figura E.1. Diagrama unifilar del sistema de 34 nodos y 24.9 kV de la IEEE

### E.3 Parámetros de las cargas

Los parámetros de las cargas concentradas se presentan en la tabla D.1.

Nodo	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-4
	[kW]	[kVAr]	[kW]	[kVAr]	[kW]	[kVAr]
860	20	16	20	16	20	16
840	9	7	9	7	9	7
844	135	105	135	105	135	105
848	20	16	20	16	20	16
890	150	75	150	75	150	75
830	10	5	10	5	25	10
Total	344	224	344	224	359	229

Tabla E.1. Datos de las cargas concentradas del sistema de 34 nodos de la IEEE.

Los parámetros de las cargas distribuidas se presentan en la tabla D.2.

Nodo inicial	Nodo final	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
		[kW]	[kVAr]	[kW]	[kVAr]	[kW]	[kVAr]
802	806	0	0	15	7.5	12.5	7
808	810	0	0	8	4	0	0
818	820	17	8.5	0	0	0	0
820	822	68	35	0	0	0	0
816	824	0	0	2.5	1	0	0
824	826	0	0	20	10	0	0
824	828	0	0	0	0	2	1
828	830	3.5	1.5	0	0	0	0
854	856	0	0	2	1	0	0
832	858	3.5	1.5	1	0.5	3	1.5
858	864	1	0.5	0	0	0	0
858	834	2	1	7.5	4	6.5	3.5
834	860	8	4	10	5	55	27.5
860	836	15	7.5	5	3	21	11
836	840	9	4.5	11	5.5	0	0
862	838	0	0	14	7	0	0
842	844	4.5	2.5	0	0	0	0
844	846	0	0	12.5	6	10	5.5
846	848	0	0	11.5	5.5	0	0
Total		131	66.5	120	60	110	57

Tabla E.2. Datos de las cargas distribuidas del sistema de 34 nodos de la IEEE.

#### E.4 Parámetros de las líneas

Los parámetros de las líneas se presentan en la tabla E.3.

Nodo inicial	Nodo final	Longitud del tramo [Millas]	Configuración.
800	802	0.489	300
802	806	0.328	300
806	808	6.104	300
808	810	1.099	303
808	812	7.102	300
812	814	5.631	300
814	850	0.002	301
816	818	0.324	302
816	824	1.934	301
818	820	9.119	302
820	822	2.602	302
824	826	0.574	303
824	828	0.159	301
828	830	3.871	301
830	854	0.098	301
832	858	0.928	301
832	888	0.000	XFM-1
834	860	0.383	301
834	842	0.053	301
836	840	0.163	301
836	862	0.053	301
842	844	0.256	301
844	846	0.689	301
846	848	0.100	301
850	816	0.059	301
852	832	0.002	301
854	856	4.419	303
854	852	6.975	301
858	864	0.307	303
858	834	1.104	301
860	836	0.508	301
862	838	0.920	304
888	890	2.000	300

Tabla E.3. Datos de las líneas del sistema de 34 nodos de la IEEE.

Las matrices de los parámetros  $Z$  y  $B$  se presentan a continuación

Configuración 300:

$Z$  (R +jX) en ohms por milla  
 1.3368 1.3343 0.2101 0.5779 0.2130 0.5015  
                   1.3238 1.3569 0.2066 0.4591  
                                   1.3294 1.3471  
 B en micro Siemens por milla  
 5.3350 -1.5313 -0.9943  
           5.0979 -0.6212  
                   4.8880

Configuración 301:

$Z$  (R +jX) en ohms por milla  
 1.9300 1.4115 0.2327 0.6442 0.2359 0.5691  
                   1.9157 1.4281 0.2288 0.5238  
                                   1.9219 1.4209  
 B en micro Siemens por milla  
 5.1207 -1.4364 -0.9402  
           4.9055 -0.5951  
                   4.7154

Configuración 302:

$Z$  (R +jX) en ohms por milla  
 2.7995 1.4855 0.0000 0.0000 0.0000 0.0000  
                   0.0000 0.0000 0.0000 0.0000  
                                   0.0000 0.0000  
 B en micro Siemens por milla  
 4.2251 0.0000 0.0000  
           0.0000 0.0000  
                   0.0000

Configuración 303:

$Z$  (R +jX) en ohms por milla  
 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
                   2.7995 1.4855 0.0000 0.0000  
                                   0.0000 0.0000  
 B en micro Siemens por milla  
 0.0000 0.0000 0.0000  
           4.2251 0.0000  
                   0.0000

Configuración 304:

$Z$  (R +jX) en ohms por milla  
 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000  
                   1.9217 1.4212 0.0000 0.0000  
                                   0.0000 0.0000  
 B en micro Siemens por milla  
 0.0000 0.0000 0.0000  
           4.3637 0.0000  
                   0.0000

## E.5 Parámetros de los transformadores

Los parámetros de los transformadores se presentan en la tabla E.4.

Nodo	kVA	kV-alta	kV-baja	R - %	X - %
Subestación:	2500	69 - D	24.9 - Gr. W	1	8
XFM -1	500	24.9 - Gr.W	4.16 - Gr. W	1.9	4.08

Tabla E.4. Datos de los transformadores del sistema de 34 nodos de la IEEE.