

Identificación de cargas eléctricas residenciales utilizando características basadas en la transformada fraccionaria de Fourier

Henry Mauricio Cala Castro

Trabajo de Investigación para optar al título de Magíster en Ingeniería Electrónica

Director

César Antonio Duarte Gualdrón

Ingeniero Electricista - Doctor en filosofía, Ph.D.

Codirector

Rafael Ángel Torres Amarís

Físico - Doctor en ciencias, D.Sc.

Universidad Industrial de Santander

Facultad de Ingenierías Físico-Mecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones

Magíster en ingeniería Electrónica

Bucaramanga

2018

**Tabla de Contenido**

Introducción . . . . .	17
1 Identificación no intrusiva de cargas eléctricas . . . . .	22
2 Diseño del sistema de identificación con características basadas en la transformada de Fourier fraccionaria . . . . .	32
2.1 Adquisición de señales y construcción de la base de datos . . . . .	32
2.2 Operadores de extracción de características basados en la transformada fraccionaria de Fourier . . . . .	35
2.2.1 Transformada de Fourier fraccionaria. . . . .	35
2.2.2 Extracción de características utilizando la relación entre la distribución de Wigner-Ville y la FRFT . . . . .	39
2.2.3 Extracción de características utilizando el centroide espectral fraccionario . . . . .	42
2.3 Modelos de estimación . . . . .	43
2.3.1 Máquinas de Soporte Vectorial . . . . .	44
2.3.2 Redes Neuronales Artificiales . . . . .	47
2.4 Selección del conjunto de características . . . . .	49
2.5 Sistemas propuestos para la identificación de cargas eléctricas basados en la transformada de Fourier fraccionaria . . . . .	52
2.5.1 Modelos de estimación para información de régimen estacionario . . . . .	52
2.5.2 Modelos de estimación para información de régimen transitorio . . . . .	53

SISTEMAS DE IDENTIFICACIÓN CON FRFT	5
3 Análisis del desempeño de los sistemas de identificación de cargas eléctricas diseñados	59
3.1 Resultados de los sistemas utilizando la FRFT . . . . .	59
3.2 Resultados de los sistemas utilizando la transformada Wavelet . . . . .	60
3.3 Resultados de los sistemas utilizando la transformada S . . . . .	63
3.4 Análisis de los resultados obtenidos en los sistemas de identificación . . . . .	66
3.4.1 Análisis de los resultados en identificación de los sistemas basados en eventos de régimen estacionario . . . . .	69
3.4.2 Análisis de los resultados en identificación de los sistemas basados en eventos de régimen transitorio . . . . .	74
3.5 Comparación de los sistemas de identificación . . . . .	77
4 Conclusiones . . . . .	82
Referencias Bibliográficas . . . . .	87
Apéndices . . . . .	92

### Lista de Tablas

1	Modelos de estimación, conjuntos de características propuestos y dispositivos estudiados de algunos sistemas de identificación de cargas revisados. . . . .	28
2	Etiquetas asignadas a dispositivos en uno o varios estados de operación y número total de eventos transitorios registrados por cada etiqueta. Este número incluye los transitorios registrados también cuando el dispositivo se encuentra funcionando en combinación con otros dispositivos como se muestra en la tabla 3. Un evento transitorio corresponde a la conexión o desconexión de un dispositivo a la vez, mientras los demás dispositivos de una combinación dada permanecen conectados. . . . .	36
3	Etiquetas asignadas para las combinaciones de uno, dos o tres dispositivos y número de señales registradas en estado estacionario por cada combinación. Las descripción de las etiquetas utilizadas para cada combinación se encuentra en la tabla 2. . . . .	37
4	Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen estacionario utilizando SVMs . . . . .	54
5	Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen estacionario utilizando ANNs . . . . .	55
6	Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen transitorio utilizando SVMs . . . . .	57

7	Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen transitorio utilizando ANNs . . . . .	58
8	Parámetros estadísticos de la distribución del error por clase de los sistemas de identificación con extracción de características basada en FRFT Wavelet y S y estimación con SVMs y ANNs. . . . .	69
9	Ranking de mayor a menor en base en el error de estimación para los sistemas con de extracción de características basada la FRFT y las transformadas Wavelet y S con estimación con SVMs. El ranking se realiza para cada conjunto de combinaciones de cargas y método de extracción. . . . .	71
10	Ranking de 5 elementos de mayor a menor en base en el error de estimación para los sistemas con de extracción de características basada la FRFT y las transformadas Wavelet y S con estimación con ANNs. El ranking se realiza para cada conjunto de combinaciones de cargas y método de extracción. . . . .	73
11	Error en la identificación de los dispositivos con eventos de combinaciones de una (1), dos (2) y tres (3) cargas de los sistemas con los 3 métodos de extracción de características (FRFT, Wavelet, S) y estimación con SVM. . . . .	74
12	Exactitud en identificación en validación cruzada de los sistemas de identificación con SVMs y ANNs para eventos del régimen estacionario y transitorio con las técnicas de extracción de características desarrolladas . . . . .	77

13 Error en la estimación en cada partición del proceso de validación cruzada de los sistemas de identificación de eventos de régimen estacionario con estimación con SVMs. El valor de  $z_{a_t}$  y  $z_{b_t}$  corresponden a la diferencia entre el sistema de FRFT con Wavelet y S respectivamente. . . . . 78

14 Error en la estimación en cada partición del proceso de validación cruzada de los sistemas de identificación de eventos de régimen transitorio con estimación con SVMs. El valor de  $z_{a_t}$  y  $z_{b_t}$  corresponden a la diferencia entre el sistema de FRFT con Wavelet y S respectivamente. . . . . 78

15 Valor de estadístico  $|t_{est}|$  para la prueba de hipótesis en la comparación de los sistemas de régimen estacionario y transitorio con SVMs. . . . . 79

16 Tiempo de cómputo en segundos de la estimación de un evento en los sistema de identificación de eventos transitorios con tamaño de evento  $M = 153965$ . . . . . 80

17 Tiempo de cómputo en segundos de la estimación de un evento en los sistema de identificación de eventos estacionarios con tamaño de evento  $M = 2500$ . . . . . 80

18 Error de identificación (%) en validación cruzada de los sistema A y B de eventos de régimen transitorio con estimación realizada con SVMs. El sistema A corresponde a los eventos transitorios de apagado de los dispositivos y el sistema B a los demás eventos . 81

19 Especificaciones del módulo de adquisición de datos NI9225 . . . . . 92

20 Especificaciones del módulo de adquisición de datos NI9227 . . . . . 93

21 Especificaciones del módulo de adquisición de datos NI9239 . . . . . 93

22 Especificaciones de la fuente programable CHROMA 61511 . . . . . 94

23 Especificaciones de la pinza amperimétricas Fluke i30s . . . . . 95

24 Etiquetas asignadas para las combinaciones de uno, dos o tres dispositivos y número de señales registradas en estado transitorio por cada combinación. La descripción de las etiquetas utilizadas para cada combinación se encuentra en la tabla 2. . . . . 99

### Lista de Figuras

1	Diagrama de flujo del proceso de diseño del sistema de identificación de cargas con extracción de características basada en la FRFT. . . . .	33
2	Diagramas de medición directa e indirecta de voltaje y corriente de las combinaciones de cargas con la fuente de alimentación de voltaje, $v_{in}$ , CHROMA 61511. . . . .	34
3	Algoritmo de cómputo de la FRFT basado en la definición de FRFT para señales discretas de la ecuación 2.3. . . . .	39
4	Relación entre las distribuciones de energía de $x(t)$ y $\mathcal{F}_\alpha[x(t)]$ con un orden fraccionario $\alpha$ . . . . .	40
5	a) Topología de una Red Neuronal Artificial con una capa oculta y Z neuronas. b) Topología de una Red Neuronal Artificial con dos capas ocultas con Z y J neuronas respectivamente. . . . .	48
6	algoritmos de diseño del modelo final de ANNs para los sistemas de identificación desarrollados. . . . .	50
7	Diagrama de flujo del algoritmo de eliminación de características recursivas hacia atrás	51
8	Error en validación cruzada de cada iteración en el proceso de selección de características utilizando SVMs para experimentos en régimen estacionario. Se determina en este caso seleccionar 12 características. . . . .	53

10	Error en validación cruzada de cada iteración en el proceso de selección de características utilizando SVMs para experimentos en régimen transitorio. Se determina en este caso seleccionar 17 características. . . . .	56
11	Error en validación cruzada de cada iteración en el proceso de selección de características utilizando ANNs para experimentos en régimen transitorio. Se determina en este caso seleccionar 55 características. . . . .	57
12	Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando los conjuntos finales de características de la FRFT y estimación realizada con ANNs y SVMs. . . . .	61
13	Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando los conjuntos finales de características de la FRFT y estimación realizada con ANNs y SVMs. . . . .	62
14	Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando conjuntos de características extraídos con la transformada Wavelet y estimación realizada con ANNs y SVMs. . . . .	64
15	Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando conjuntos de características extraídos con la transformada Wavelet y estimación realizada con ANNs y SVMs. . . . .	65
16	Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando conjuntos de características extraídos con la transformada S y estimación realizada con ANNs y SVMs. . . . .	67

SISTEMAS DE IDENTIFICACIÓN CON FRFT	12
17 Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando conjuntos de características extraídos con la transformada S y estimación realizada con ANNs y SVMs. . . . .	68
18 Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la FRFT y estimación con SVM . . . . .	71
19 Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la transformada Wavelet y estimación con SVM . . . . .	72
20 Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la transformada S y estimación con SVM . . . . .	72
21 Error en la identificación de dispositivos con eventos de combinaciones de cargas de los sistemas de identificación de régimen transitorio con estimación realizada con SVM para los 3 métodos de extracción de características (FRFT, Wavelet, S). . . . .	75
22 Error en la identificación de eventos transitorios por categoría de dispositivos con los 3 métodos de extracción de características (FRFT, Wavelet, S) y estimación con SVM . . . . .	76
23 Diagrama de flujo de los algoritmos genéticos para la resolución de los problemas óptimos 2.12 y 2.16 . . . . .	96
24 Adquisición de la señal de corriente del encendido y apagado de la licuadora Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz]. . . . .	97

- 25 Evento de régimen transitorio de la señal de corriente del encendido de la licuadora  
Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz]. . . . . 98
- 26 Evento de régimen estacionario de la señal de corriente del encendido de la licuadora  
Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz]. . . . . 98
- 27 Eventos de régimen transitorio y estacionario de la secuencia 1 para el computador de  
escritorio Dell con frecuencia de muestreo 50[kHz]. . . . . 100
- 28 Eventos de régimen transitorio y estacionario de la secuencia 2 para el computador de  
escritorio Dell con frecuencia de muestreo 50[kHz]. . . . . 100
- 29 Eventos de régimen transitorio y estacionario de la secuencia 3 para el computador de  
escritorio Dell con frecuencia de muestreo 50[kHz]. . . . . 101

**Lista de apéndices**

Apéndice A. Especificaciones de los instrumentos utilizados en la adquisición de la base de datos. . . . . 92

Apéndice B. Algoritmos genéticos usados para resolver el problema de órdenes óptimos . 95

Apéndice C. Extracción de eventos en las señales sin procesar y descripción de los eventos de régimen transitorio y secuencias de eventos para el computador de escritorio. . . . . 96

## Resumen

**TÍTULO:** IDENTIFICACIÓN DE CARGAS ELÉCTRICAS RESIDENCIALES UTILIZANDO CARACTERÍSTICAS BASADAS EN LA TRANSFORMADA FRACCIONARIA DE FOURIER<sup>1</sup>

**AUTOR:** HENRY MAURICIO CALA CASTRO<sup>2</sup>

**PALABRAS CLAVE:** MONITORIZACIÓN NO INTRUSIVA DE CARGAS, IDENTIFICACIÓN DE CARGAS, TRANSFORMADA DE FOURIER FRACCIONARIA, EXTRACCIÓN DE CARACTERÍSTICAS, MÁQUINAS DE SOPORTE VECTORIAL, REDES NEURONALES ARTIFICIALES.

### DESCRIPCIÓN:

Con la información del consumo de energía individual de los dispositivos se pueden generar estrategias de gestión para optimizar la eficiencia energética en una casa. En esta área, se presenta la monitorización no intrusiva de cargas (NILM) como un método de bajo costo, el cual consiste en realizar una identificación de los dispositivos activos para posteriormente definir los consumos individuales. En el presente trabajo de investigación, se describe el diseño de sistemas de identificación de cargas basados en eventos transitorios y estacionarios de los dispositivos con características extraídas con la transformada fraccionaria de Fourier (FRFT). En el diseño propuesto, se definen operadores de extracción de características con las propiedades de las señales en los dominios fraccionarios y se seleccionan conjuntos subóptimos de características con un proceso de eliminación recursivo. Luego, se implementan los sistemas de identificación con estimación realizada con Máquinas de Soporte Vectorial (SVMs) y Redes Neuronales Artificiales (ANNs) y los resultados obtenidos en validación cruzada se comparan con los de sistemas con extracción de características basada en las transformadas Wavelet y S. Como resultados, se tiene que los algoritmos con SVMs logran exactitudes en identificación más elevadas en comparación de los algoritmos con ANNs. Además, en los sistemas con SVMs, la exactitud en identificación con los enfoques basados en Wavelet y FRFT son muy similares logrando una ventaja sobre el enfoque basado en la transformada S. En procesamiento, el sistema de FRFT obtiene una mayor velocidad de cómputo seguido del sistema con transformada S. Considerando lo anterior, los sistemas con extracción de características basada en la FRFT muestran el mejor balance entre exactitud y velocidad entre las alternativas comparadas.

---

<sup>1</sup>Trabajo de investigación

<sup>2</sup>Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Maestría en Ingeniería Electrónica, Grupo GISEL. Director Ph.D. César Antonio Duarte, Ingeniero Electricista. Codirector D.Sc. Rafael Ángel Torres, Físico.

**Abstract**

**TITLE:** RESIDENTIAL LOADS IDENTIFICATION WITH FEATURE EXTRACTION BASED ON THE FRACTIONAL FOURIER TRANSFORM <sup>3</sup>

**AUTHOR:** HENRY MAURICIO CALA <sup>4</sup>

**KEYWORDS:** NONINTRUSIVE LOAD MONITORING, LOAD IDENTIFICATION, FRACTIONAL FOURIER TRANSFORM, FEATURE EXTRACTION, SUPPORT VECTOR MACHINES, ARTIFICIAL NEURAL NETWORKS

**DESCRIPTION:**

With the individual energy consumptions of the devices, management plans can be generated in order to optimize energy efficiency in a home. In this area, Nonintrusive Load Monitoring (NILM) is presented as a low cost method, which consists in realizing an identification of the active loads, and then defining the individual power consumptions using prior information. In this research work, the design of load identification systems is presented based on stationary and transient events with features that are extracted using the Fractional Fourier Transform (FRFT). Feature extraction operators are proposed using some properties and patterns in the fractional Fourier domains. Then, a recursive feature elimination method is implemented to identify suboptimum feature groups from the complete feature set. With the final features groups, load identification systems are designed and implemented, where the estimation is solved with Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs). Finally, the cross-validation scores of the systems is compared with approaches using feature extraction based on S and Wavelet transforms. As results, the algorithms with SVMs has a higher accuracy than the ANNs approaches. Also, in SVMs systems, accuracies are similar for the proposals based on Wavelet and FRFT, and they both exhibit better accuracy than the system based on S transform. Regarding computation time, the FRFT system is the fastest algorithm followed by the S transform one. As final conclusion, the systems with feature extraction based on the FRFT have the best average performance in speed and accuracy among the systems tested.

---

<sup>3</sup>Research work.

<sup>4</sup>Faculty of physical-Mechanical Engineering, School of Electrical, Electronic and Telecommunications Engineering. Master in Electronic Engineering, Group GISEL. Director Ph.D. César Duarte, Electrical engineer. Codirector D.Sc. Rafael Ángel Torres, physicist.

## Introducción

La demanda global de energía eléctrica se incrementa progresivamente por factores como el crecimiento poblacional y la masificación de dispositivos electrónicos, entre otros. Con esto, surgen interrogantes sobre la capacidad de los sistemas eléctricos para poder suplir la demanda energética en los próximos años, que a su vez, han motivado el desarrollo de investigaciones en nuevas fuentes de generación y tecnologías para el uso racional de la energía eléctrica (Ehrhardt-Martinez, Donnelly, Laitner, y cols., 2010). Con planes de gestión energética más elaborados, se puede incentivar al consumidor a dar un uso más consciente a la energía eléctrica. Para la elaboración de estos planes, es necesario el desarrollo de sistemas que entreguen información detallada y automática del consumo de los usuarios. Con esta información se pueden identificar los hábitos de consumo de los clientes y realizar campañas que motiven los hábitos positivos y reduzcan la frecuencia de los negativos. Estudios muestran que una reducción de alrededor del 10 al 15 % en el consumo de energía residencial se puede obtener aplicando este tipo de estrategia de gestión (Darby y cols., 2006), (Fischer, 2008), (Ehrhardt-Martinez y cols., 2010).

Para obtener información de los hábitos de consumo de los usuarios se han propuesto sistemas de procesamiento de datos que pueden diferir en el nivel de detalle de la información que entregan. En este aspecto, es de gran interés la implementación de sistemas que identifiquen la actividad y el consumo de cada dispositivo en una casa o edificio. Para realizar la estimación del consumo individual de las cargas se han desarrollado dos alternativas principales: la monitorización intrusi-

va de cargas (*ILM: Intrusive Load Monitoring*) y la monitorización no intrusiva de cargas (*NILM: Non-Intrusive Load Monitoring*). Con los métodos ILM se obtiene el consumo individual de cada dispositivo mediante múltiples sensores acoplados en los puntos de conexión de las cargas, es decir, se realiza una medición directa en cada dispositivo. Por otra parte, en los métodos NILM solo se requiere un sensor conectado al panel principal, y con información de las señales agregadas y algoritmos de procesamiento de señales se busca identificar los consumos individuales de los dispositivos (Abubakar, Khalid, Mustafa, Shareef, y Mustapha, 2015). Al comparar ambas alternativas, se tiene que los métodos ILM tienen una mayor exactitud en la estimación del consumo individual debido a que miden directamente en el punto de conexión del dispositivo. Por otra parte, la implementación de un sistema ILM es más costosa porque requiere un sensor por cada dispositivo conectado a la red (Zoha, Gluhak, Imran, y Rajasegarar, 2012). Computacionalmente, el diseño de un sistema NILM es más complejo en el sentido que requiere modelos que describan la actividad y/o el consumo de cada dispositivo utilizando información agregada de los dispositivos. Considerando las ventajas y desventajas de cada alternativa, las investigaciones se han centrado en mejorar el rendimiento de los sistemas NILM y reducir su complejidad computacional con el propósito de obtener un sistema eficiente y de bajo costo monetario.

Una propuesta para la implementación de un sistema NILM consiste en, inicialmente, identificar los dispositivos activos (identificación de cargas) y luego describir el consumo promedio de los electrodomésticos utilizando información conocida de las cargas (Abubakar y cols., 2015). Con este enfoque, el rendimiento del sistema se determina por la exactitud en la identificación de cargas, la cual consiste en extraer características en las señales agregadas y con algoritmos de reco-

nocimiento de patrones realizar la estimación los dispositivos que generaron tales características. En la literatura estudiada en el capítulo 1, la extracción de características generalmente se realiza calculando parámetros de eventos estacionario y/o transitorio en el dominio del tiempo y/o dominios transformados. Para eventos en el régimen estacionario, se puede destacar como conjuntos de características la representación en armónicos de Fourier, los valores de potencia activa y reactiva, así como los cambios que se presenten en estos valores. En eventos transitorios, se acostumbra a realizar transformaciones tiempo-frecuencia (S, Wavelet y STFT) a los datos y extraer parámetros de estas representaciones. En cuanto a la estimación, existe una tendencia a utilizar algoritmos de aprendizaje supervisado entre los cuales se destacan las Máquinas de Soporte Vectorial (SVMs), Redes Neuronales Artificiales (ANNs) y vecinos cercanos.

La transformada fraccionaria de Fourier (FRFT) es una herramienta para analizar procesos no estacionarios en un sentido amplio utilizando la noción clásica de armónico de la transformada de Fourier. Con esta transformación, se puede representar las señales en dominios intermedios entre el tiempo y la frecuencia, donde dichas representaciones están asociadas a proyecciones de la distribución de energía de las señales (Almeida, 1994), (R. T. Amarís, 2008). considerando esto, surgió la pregunta de investigación ¿Existen características en la representación de las señales en dominios fraccionarios que mejoren el rendimiento de los sistemas de identificación? la cual motivó el planteamiento de la presente investigación. En esta, se busca proponer y explorar diferentes conjuntos de características extraídas de la representación de las señales en dominios fraccionarios y evaluar su desempeño en la identificación de cargas eléctricas residenciales. El desarrollo de la propuesta de investigación aborda el cumplimiento del objetivo general de diseñar y evaluar

sistemas de identificación de cargas eléctricas con características extraídas a partir de la FRFT y estimación realizada con SVMs y ANNs. El cual se puede dividir en la realización de los siguientes objetivos secundarios:

- Proponer y seleccionar características de la representación de las señales con la FRFT para el régimen transitorio y estacionario de señales de un grupo de cargas residenciales.
- Diseñar sistemas de identificación de cargas con ANNs y SVMs que utilicen los conjuntos de características propuestos.
- Contrastar el desempeño del sistema de identificación de cargas mediante la exactitud en identificación y el tiempo de cómputo con sistemas con diferente propuesta de extracción de características.

En el actual documento, se describe el proceso de diseño y evaluación de los sistemas de identificación de cargas propuestos basados en eventos del régimen estacionario y transitorio de los dispositivos. La extracción de características se realiza calculando centroides y parámetros de las representaciones de las señales en diferentes dominios fraccionarios. Además, para el conjunto de características propuesto se implementa un proceso de selección de características recursivo en búsqueda de subconjuntos óptimos de menor dimensión. La estimación de los dispositivos se efectúa con modelos de SVMs y ANNs, y los resultados en exactitud y velocidad de cómputo obtenidos son comparados con los de sistemas de identificación con extracción de características basados en la transformada S y Wavelet. En las comparaciones realizadas se tiene que la estimación con SVMs logra mejores valores de exactitud en identificación en relación con los de ANNs. Cuando la estimación se realiza con SVMs, los sistemas basados en Wavelet y FRFT logran exactitudes similares logrando una ventaja sobre el sistema basado en la transformada S. Al

comparar la velocidad de cómputo, el sistema con FRFT y estimación con SVMs logra la mayor velocidad, seguido por el sistema con transformada S. En general, los sistemas de identificación con extracción de características con la FRFT y estimación con SVMs logran el mejor balance entre exactitud y velocidad de procesamiento.

### **Organización del documento**

El contenido de la tesis es desarrollado de la siguiente manera, en el capítulo 2 se describe el problema en la monitorización no intrusiva de cargas y una propuesta de solución con el diseño de sistemas de identificación de cargas con algoritmos de reconocimiento de patrones. La solución planteada se fundamenta en el diseño de sistemas de identificación de cargas con extracción de características utilizando las propiedades de la FRFT y con estimación realizada con SVMs y ANNs. Luego, se presenta la literatura estudiada, realizando énfasis en las diferentes propuestas de extracción de características y algoritmos de estimación en los sistemas revisados. En el capítulo 3, se desarrolla el diseño del sistema de identificación planteado, lo cual incluye la propuesta de operadores de extracción de características, el proceso de selección de características y el diseño de los algoritmos de estimación. Al final de este capítulo se enuncian los resultados obtenidos en cada etapa de diseño y los sistemas finales de identificación con extracción de características basada en la FRFT. En el capítulo 4, se muestran los principales resultados de exactitud por dispositivo y globales en la identificación de eventos transitorios y estacionarios con cada sistema desarrollado. Los resultados de los sistemas son comparados en su exactitud en la identificación y en su velocidad de procesamiento. Finalmente, en el capítulo 5 se enuncian las conclusiones obtenidas en la investigación.

## 1. Identificación no intrusiva de cargas eléctricas

El consumo total de potencia  $P_T(t)$  en un tiempo  $t$  para  $D$  dispositivos (consumiendo o no energía) en una casa se puede representar como

$$P_T(t) = P_1(t) + P_2(t) + \dots P_D(t), \quad (1.1)$$

donde  $P_j(t)$  es la potencia asociada a cada dispositivo para  $j = 1, 2, \dots, D$ . En un sistema NILM se busca encontrar el conjunto  $\{P_j(t) | j = 1, 2, \dots, D\}$  a través de observaciones o mediciones de la potencia total agregada  $P_T(t)$ . Asimismo, las potencias en la ecuación (1.1) cambian en el tiempo debido a operaciones de conexión, desconexión y cambio de estado de los dispositivos. Además, condiciones ruidosas, la presencia de distorsiones o cambios en las condiciones de operación producen variaciones en los consumos de potencia de los dispositivos. El problema de desagregación se puede resolver a partir de las potencias nominales (o estimadas) de los dispositivos, minimizando el error entre la potencia total medida y la suma de potencias nominales (o estimadas). Esta solución puede tener una incertidumbre elevada y un espacio de búsqueda considerable dado el gran número de posibles combinaciones de dispositivos. En consecuencia, en la literatura se han propuesto métodos de solución que incluyan información adicional como los algoritmos de reconocimiento de patrones (Zoha y cols., 2012), los cuales serán utilizados en este trabajo de investigación.

Un planteamiento de sistema NILM consiste en realizar un proceso de identificación de las cargas activas (que consumen energía) en la casa y posteriormente de un proceso de estimación de los consumos de potencia de cargas activas con información conocida a priori sobre estas. En el proceso de identificación de cargas o grupos de cargas, se extraen características de señales agregadas (e. g. tensión, corriente, potencia, ...), las cuales son relacionadas con las cargas o grupos de cargas mediante modelos de estimación. Es decir, los modelos de estimación predicen el conjunto de cargas activas que generan las características observadas. La identificación de cargas es altamente dependiente del diseño y selección de los modelos de estimación, así como del conjunto de características extraído para la representación de las cargas. Como modelo de estimación generalmente se utilizan modelos de aprendizaje de máquina y sus parámetros se ajustan con procesos de entrenamiento utilizando bases de datos con información del comportamiento de los dispositivos. Se espera que en el conjunto de datos de entrenamiento se incluya la mayor cantidad de eventos de los dispositivos para que el modelo de estimación tenga la mayor capacidad posible de generalización del problema.

Considerando un operador de extracción de características  $V(\cdot)$  y una señal  $x_k(t)$  medida durante un  $k$ -ésimo experimento para un dispositivo o grupo de dispositivos, se tiene que el vector  $v_k$  de características extraídas está dado por

$$v_k = V(x_k(t)). \quad (1.2)$$

Ahora, se desea contar con un modelo de estimación que permita asociar un vector de características dado,  $v_k$ , con una etiqueta  $l_k$  que permite conocer información del experimento corres-

pondiente. Por ejemplo, una etiqueta puede brindar información sobre el cambio de estado en la operación de un dispositivo dado o, por ejemplo, sobre el conjunto de dispositivos que se encuentran conectados durante una operación en estado estacionario. Para un modelo de estimación  $\Psi_{\Theta}(v_k)$  con parámetros  $\Theta$ , el error  $e$  en la estimación se puede calcular como:

$$e = \frac{1}{N} \sum_{k=1}^N b_k, \quad (1.3)$$

con

$$b_k = \begin{cases} 1 & l_k \neq \Psi_{\Theta}(v_k), \\ 0 & l_k = \Psi_{\Theta}(v_k). \end{cases}$$

Donde  $N$  es el número de experimentos considerados en la evaluación del modelo de estimación. En el diseño de un sistema de identificación se busca encontrar el operador  $V(\cdot)$  y los parámetros  $\Theta$  que minimicen el error  $e$  para un conjunto de datos de validación. Considerando lo anterior, en el diseño de sistemas de identificación de cargas se pueden destacar dos etapas principales: La extracción de características y el diseño del algoritmo de estimación. A continuación se presenta un análisis de la literatura existente sobre sistemas NILM. En este análisis se han considerado la complejidad y rendimiento de los diferentes operadores de extracción de características, los modelos utilizados para la estimación de etiquetas de información, y los dispositivos considerados en los experimentos para construir las bases de datos para entrenamiento y validación. La tabla 1 sintetiza los resultados de este análisis de literatura.

Los algoritmos de identificación generalmente son seleccionados del conjunto de algoritmos de aprendizaje de máquina, los cuales pueden ser supervisados o no supervisados (Zoha y cols.,

2012). Los algoritmos de aprendizaje no supervisado buscan tendencias y relaciones entre un conjunto de datos para agruparlos. Los datos nuevos que ingresen al algoritmo serán agrupados en los grupos definidos en el entrenamiento. Por otra parte, los algoritmos de aprendizaje supervisado requieren un conjunto de datos de entrada con sus respectivas salidas (etiquetas) para estimar los parámetros de los modelos. Una vez obtenido el modelo, éste se puede utilizar para estimar las salidas de nuevos datos (Christopher, 2016). En los sistemas de identificación de cargas es de mayor interés estimar las cargas o grupos de cargas activas a estimar posibles agrupaciones de los datos de entrenamiento, por lo que existe una mayor tendencia a utilizar algoritmos supervisados. Por ejemplo, en la literatura contemplada se utilizan diferentes algoritmos supervisados como: vecinos cercanos en [(Figueiredo, De Almeida, y Ribeiro, 2012), (Tsai y Lin, 2011), (Hamid, Barbarosou, Papageorgas, Prekas, y Salame, 2017), (Tsai y Lin, 2011), (Berges, Goldman, Matthews, y Soibelman, 2010), (Chahine y cols., 2011), (Gillis y Morsi, 2017)], redes neuronales [(Hamid y cols., 2017), (Srinivasan, Ng, y Liew, 2006), (Chang, Chien, Lin, y Chen, 2011), (Tsai y Lin, 2011), (Bianson y Plangklang, 2016), (Chang, Kuo-Lung, Su, y Lee, 2014), (Su, Lian, y Chang, 2011), (Wang y Yin, 2017), (Chang, 2012), (Du y cols., 2012), (Chang, Lian, Su, y Lee, 2013), (Yang, Chang, y Lin, 2007)], máquinas de soporte vectorial [(Srinivasan y cols., 2006), (Chahine y cols., 2011), (Lin y Tsai, 2011), (Duarte, Delmar, Goossen, Barner, y Gomez-Luna, 2012), (Figueiredo y cols., 2012), (Jimenez, Duarte, Petit, y Carrillo, 2014), (Jiang, Luo, y Li, 2012), (Du y cols., 2012)] y modelos de árboles de decisión son utilizados en los trabajos [(Bhattacharjee, Kumar, y RoyChowdhury, 2014), (Hamid y cols., 2017), (Gillis y Morsi, 2017)].

El proceso de extracción de características se fundamenta en el cálculo de parámetros en los

regímenes transitorio y estacionario de las señales eléctricas. En los trabajos [(Yang y cols., 2007), (Berges y cols., 2010), (Chang y cols., 2011), (Su y cols., 2011), (Lin y Tsai, 2011), (Du y cols., 2012), (Figueiredo y cols., 2012), (Bianson y Planklang, 2016), (Hamid y cols., 2017)] se utiliza la potencia activa y reactiva de los dispositivos como características y [(Yang y cols., 2007), (Su y cols., 2011), (Du y cols., 2012)] incluyen además información de eventos transitorios. [(Tsai y Lin, 2011), (Lin y Tsai, 2011)] utilizan únicamente parámetros de señales transitorias en el dominio del tiempo como el factor de cresta, la energía y la duración del evento. Otros conjuntos de propuestas de extracción de características consisten en transformar las señales eléctricas a otros dominios y extraer parámetros de éstas representaciones. Con este enfoque se espera que la transformación realizada a las señales permite la extracción de patrones con mayor poder de discriminación. En (Srinivasan y cols., 2006) y (Jiang y cols., 2012) se utiliza como características las componentes armónicas extraídas con la transformada de Fourier y la STFT. En (Chahine y cols., 2011) se presenta un conjunto de características basadas en los polos de las señales eléctricas calculados con la transformada  $z$ . En [(Duarte y cols., 2012), (Chang y cols., 2014), (Gillis y Morsi, 2017)] se define como vector de características la contribución de energía de las escalas en la transformada Wavelet y en (Jimenez y cols., 2014) se propone una extracción de características utilizando la transformada  $S$ . Una revisión más amplia del estado del arte de los sistemas de identificación se encuentra en [(Zoha y cols., 2012), (Tabatabaei, Dick, y Xu, 2017), (Sadeghianpourhamami, Ruysinck, Deschrijver, Dhaene, y Develder, 2017)].

En la mayoría de documentos revisados se presenta un número relativamente bajo de dispositivos y etiquetas a identificar. Debido a esto, los modelos de estimación tienen un número reducido

de posibles salidas, con lo que se puede explicar los resultados de exactitud superiores a 80 % en las investigaciones. Un conjunto más amplio de dispositivos y combinaciones de dispositivos deben ser considerados para obtener una estimación más cercana a un escenario real del rendimiento de los sistemas. Además, se encuentra la tendencia de utilizar SVMs, ANNs y vecinos cercanos para la estimación, y en los escenarios donde son comparados, las diferencias no son muy elevadas entre estos 3 enfoques. En el régimen estacionario, existe una alta tendencia en utilizar los valores de potencia activa y reactiva, lo que indica que este conjunto de características parece ser suficiente para identificar los eventos, aunque hace falta una exploración más amplia con un mayor número de cargas y combinaciones. Para eventos de régimen transitorio, no se presentan conjuntos de características que sean frecuentemente utilizados, con lo que existe una gran variedad de propuestas con transformaciones tiempo-frecuencia. En este escenario, se busca encontrar una estrategia de extracción de características que presente la información más discriminante de las distribuciones tiempo-frecuencia de las señales, con lo que una alternativa se puede encontrar en la FRFT por sus propiedades relacionadas con las distribuciones de energía.

En la presente investigación se plantea comparar a partir de una misma base de datos la extracción de características utilizando la FRFT con las correspondientes alternativas de extracción de características utilizando representaciones tiempo-frecuencia calculadas con la transformada S y la transformada Wavelet, utilizando como modelos de estimación ANNs y SVMs.

Tabla 1.

*Modelos de estimación, conjuntos de características propuestos y dispositivos estudiados de algunos sistemas de identificación de cargas revisados.*

<b>Doc.</b>	<b>Dispositivos</b>	<b>características</b>	<b>Modelo</b>	<b>Contribución</b>	<b>Año</b>
(Srinivasan y cols., 2006)	8 cargas y combinaciones de éstas para un total de 256 etiquetas (clases) y 18 muestras por etiqueta	Componente fundamental y partes real e imaginaria de los 15 primeros armónicos de corriente en régimen estacionario	ANNs y SVMs	Comparan el desempeño de los modelos de identificación con SVMs y ANNs. con exactitudes superiores al 95 % con ambos métodos	2006
(Yang y cols., 2007)	Cargas industriales: motor de inducción, motor síncrono, banco de cargas	Energía de eventos transitorios y potencia activa y reactiva en régimen estacionario	ANNs con propagación hacia atrás y Cuantización del vector de aprendizaje	comparan el rendimiento de los sistemas con redes neuronales entrenadas con cada método. Se obtiene que con el método de propagación hacia atrás la exactitudes es 95.6 % y con Cuantización del vector de aprendizaje es 42.94 %	2007
(Berges y cols., 2010)	25 estados de operación de un grupo de 10 cargas	Envolvente de la potencia instantánea, potencia activa y reactiva	1-vecino cercano	Evalúan el rendimiento del sistema en un escenario real. La diferencia entre la energía real consumida y la estimada fue de 14.8 % en un tiempo de 5.5 días	2010
(Tsai y Lin, 2011)	3 cargas y combinaciones de éstas: ventilador, luces fluorescentes y radio	Valor pico a pico, RMS, máximo, promedio, factor de forma y cresta, energía y potencia de eventos transitorio	Vecinos cercanos, ANNs con propagación hacia atrás y Cuantización del vector de aprendizaje	Comparan los modelos propuesto donde vecinos cercanos logra exactitud de 97.5 %, ANN con propagación hacia atrás 95 % y ANN con cuantización del vector de aprendizaje 85.63 %	2011

<b>Doc.</b>	<b>Dispositivos</b>	<b>características</b>	<b>Modelo</b>	<b>Contribución</b>	<b>Año</b>
(Chahine y cols., 2011)	9 etiquetas con 100 experimentos por etiqueta	Amplitud y locación de los polos complejos de señales discretas de corriente	Vecinos cercanos, SVMs, Naive Bayes	Compara la exactitud los sistemas y se obtienen exactitudes de identificación en vecinos cercanos 99.22 %, en SVMs 96.11 % y en Naive Bayes 94 %	2011
(Chang y cols., 2011)	Motores de inducción y bancos de cargas industriales	Potencia activa y reactiva	ANNs y algoritmos genéticos	Con algoritmos genéticos se encuentran subconjuntos óptimos de características y obtienen resultados más elevados en identificación cuando se utiliza este pre-procesamiento	2011
(Su y cols., 2011)	Dos motores de inducción y un deshumidificador	Potencia activa, reactiva y energía del transitorio de encendido	ANNs	Comparan la exactitud en identificación con detección de eventos transitorios con las transformadas Wavelet y STFT, obteniendo resultados mas altos con la transformada Wavelet.	2011
(Lin y Tsai, 2011)	Radio, ventilador, luces fluorescentes	Factor de cresta, duración del transitorio y valores máximos	SVMs	Diseñan sistemas independientes para conexiones y desconexiones. La detección de eventos transitorios es realizada con información de la corriente y su derivada	2011
(Duarte y cols., 2012)	Eventos de licuadora, ventilador y lámparas de escritorio	Contribución de energía de las escalas en la transformada Wavelet	SVMs	el sistema con transformada Wavelet se compara un sistema basado en la STFT y se obtiene 80 % y 71.43 %, respectivamente.	2012

<b>Doc.</b>	<b>Dispositivos</b>	<b>características</b>	<b>Modelo</b>	<b>Contribución</b>	<b>Año</b>
(Jiang y cols., 2012)	Ventilador, pistola de calor, cargador de PC, Monitor LCD, tetera, Motor de alta velocidad, Bombillas, PC	Análisis de armónicos con la STFT en señales de corriente	SVMs	Exactitud en identificación superior al 90 % para todas las clases	2012
(Du y cols., 2012)	Bombillas, ventilador, impresora, cargadores, DVD, calentador, Microondas, televisores LCD, LED y plasma	Valor RMS, distorsión armónica, factor de cresta, tercer y quinto armónico de la señal de corriente y factor de potencia promedio	Sistema híbrido con SVMs y ANNs no supervisadas	El sistema determina la categoría del vector de entrada utilizando ANNs no supervisadas y posteriormente se identifica el dispositivo utilizando la SVMs de la respectiva categoría	2012
(Figueiredo y cols., 2012)	6 cargas individuales y 4 combinaciones de dos cargas: cafetera, tostadora, microondas, lámpara y pantallas LCD	Corriente promedio, potencia activa y reactiva	SVMs y vecinos cercanos	Comparan el rendimiento en multiclase con las estrategias “uno contra todos” y “uno contra uno” obteniendo resultados de al rededor del 90 % con ambas estrategias	2012
(Jimenez y cols., 2014)	Lámparas de escritorio, monitor, tv, luces de oficina	Parámetros estadísticos de la representación con la transformada S	SVMs	El sistema con transformada S muestra exactitudes de 86 % y es comparado con dos propuesta de extracción de características con Wavelet con exactitudes de 82 % y 75 %	2014

<b>Doc.</b>	<b>Dispositivos</b>	<b>características</b>	<b>Modelo</b>	<b>Contribución</b>	<b>Año</b>
(Chang y cols., 2014)	Motores de inducción y cargas R-L. La base de datos consiste en medidas de laboratorio y datos sintéticos	Contribución de energía de 8 escalas en la transformada Wavelet discreta	ANNs	Prueban la exactitud del sistema para 6 casos de estudio. Se contemplan escenarios de combinaciones de cargas con potencia activa y reactiva similares, así como encendidos y apagados simultáneos	2014
(Bianson y Planklang, 2016)	Aire acondicionado, televisión, refrigerador y olla arrocera	Potencia activa y reactiva	ANNs	Detectan la conexión o desconexión de las cargas observando las variaciones en la potencia y desagregan con ANNs. Obtienen un 98 % de exactitud promedio	2016
(Hamid y cols., 2017)	8 cargas con 318 eventos por etiqueta o clase	Potencia activa y reactiva	Árboles de decisión, k-vecinos cercanos, análisis de discriminante y ANNs	obtienen exactitudes con árboles de decisión de 99 % con vecinos cercanos de 98.98 % con análisis de discriminante de 74.4 % y con redes neuronales de 83.8 %	2017
(Gillis y Morsi, 2017)	4 cargas: baterías recargables, CFL, bombilla incandescente y PC	Contribución de energía de las escalas de la transformada Wavelet	Árboles de decisión y vecinos cercanos	Se realiza entrenamiento semi-supervisado en el sistema: Se entrenan ambos algoritmos con una base de datos con etiquetas y se realiza un juicio de expertos para definir las etiquetas de un segundo conjunto de entrenamiento	2017

## **2. Diseño del sistema de identificación con características basadas en la transformada de Fourier fraccionaria**

En el presente capítulo se describe el diseño de los sistemas de identificación con características basadas en la transformada fraccionaria de Fourier. Inicialmente, se especifica la base de datos utilizada en el diseño de los algoritmos de identificación, lo cual incluye la descripción del grupo de dispositivos, combinaciones realizadas, número de etiquetas (clases) y experimentos o eventos en cada etiqueta. Además, se ilustran los diagramas de conexiones realizadas para las mediciones de corriente y voltaje de los dispositivos. Luego, se presentan los operadores de extracción de características propuestos con la FRFT y el proceso de selección de parámetros en la extracción. Finalmente, se describe el diseño de los algoritmos de identificación y la metodología utilizada en la etapa de selección del conjunto final de características. La figura 1 muestra el proceso de diseño del sistema de identificación.

### **2.1 Adquisición de señales y construcción de la base de datos**

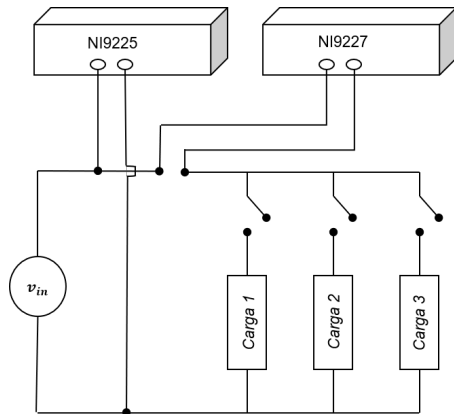
La base de datos para el diseño de los algoritmos consiste en mediciones de corriente y voltaje de la operación de los dispositivos en diferentes escenarios. Se adquirieron señales con eventos de los regímenes transitorio y estacionario de dispositivos individuales y algunas combinaciones de los mismos. El voltaje de alimentación de las cargas se generó con la fuente programable CHROMA 61511 y la adquisición de las señales de voltaje se realizó a una frecuencia de  $50[kHz]$  con los módulos NI9225 y NI9239, y las señales de corriente se adquirieron con el módulo NI9227



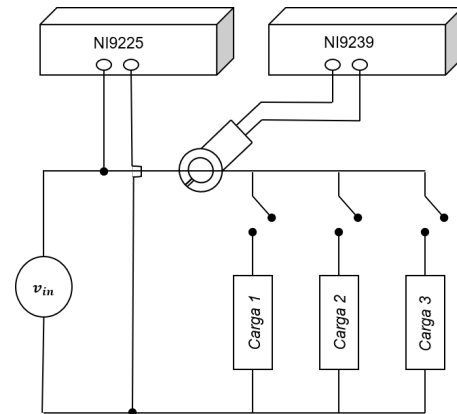
Figura 1. Diagrama de flujo del proceso de diseño del sistema de identificación de cargas con extracción de características basada en la FRFT.

de *national Instruments*. El módulo NI9227 admite una corriente máxima de 5A, por lo que dos diagramas de medición tuvieron que ser implementados considerando las cargas y combinaciones que superen éste umbral. Los circuitos implementados para la adquisición de datos se muestran en la figura 2, el circuito de la figura 2.a se utilizaba cuando el valor máximo de la corriente era menor a 5A. Asimismo, cuando se superaba éste umbral, las mediciones se realizan de forma indirecta utilizando pinzas amperimétricas de referencia Fluke i30s como se muestra en la figura 2.b. Las características de los instrumentos utilizados se pueden encontrar en el anexo A.

La base de datos de eventos es la principal unidad de información para el diseño y evaluación de los algoritmos. Un conjunto de los datos es utilizado en el cálculo de los parámetros del modelo de identificación (entrenamiento) y otro conjunto se emplea para la validación del sistema. En lo



(a) Medición directa de corriente y voltaje de los dispositivos y sus combinaciones cuando la corriente máxima de las cargas es menor a 5A. Con los módulos NI9225 y NI9227 de *national Instruments* se adquieren las señales de voltaje de entrada y de corriente suministrada por la fuente, respectivamente.



(b) Medición directa de voltaje e indirecta de corriente de los dispositivos y sus combinaciones cuando la corriente máxima es superior a 5A. Con los módulos NI9225 y NI9239 de *national Instruments* se adquiere el voltaje de entrada y el voltaje de salida de la pinza amperimétrica Fluke i30s, respectivamente.

Figura 2. Diagramas de medición directa e indirecta de voltaje y corriente de las combinaciones de cargas con la fuente de alimentación de voltaje,  $v_{in}$ , CHROMA 61511.

posible, se busca incorporar en la base de datos la mayor cantidad de información de diferentes escenarios de operación de las cargas. En el grupo de investigación GISEL se cuenta con datos de combinaciones de hasta dos cargas y en este trabajo se expandió el análisis y el tamaño de la base de datos hasta combinaciones de tres cargas.

A cada dispositivo o carga que se desea identificar se le asignó un número de etiqueta como se presenta en la Tabla 2. Es posible que se necesite identificar un dispositivo o carga en diferentes estados de operación, en tal caso, a cada estado de operación de un mismo dispositivo se le asignó una etiqueta diferente. Estas etiquetas se utilizaron para identificar diferentes combinaciones de dos (2) o tres (3) dispositivos como se presenta en la Tabla 3. Para cada combinación de uno (1), dos (2) o tres (3) dispositivos se asignó una etiqueta y se tomaron registros de corriente en estado estacionario utilizando la fuente de alimentación (CHROMA) con una tensión sinusoidal de 120 V y 60 Hz. El número total de señales registradas en estado estacionario para cada combinación

se presenta también en la tabla 3. De otra parte, para cada combinación de uno, dos o tres dispositivos se registró la señal de corriente transitoria que se induce cuando se conecta o desconecta un dispositivo a la vez en cada combinación, mientras los demás dispositivos de la combinación se mantienen conectados. El número total de eventos transitorios (de conexión o desconexión) registrados por cada dispositivo se presenta en la tabla 2. En el Anexo C se especifica la extracción de eventos de las señales sin procesar para obtener las formas de onda que conforman la base de datos para entrenar y validar los sistemas de identificación diseñados en este trabajo de investigación, así como el número de eventos transitorios por cada combinación.

## **2.2 Operadores de extracción de características basados en la transformada fraccionaria de Fourier**

En esta sección se presentan las definiciones y propiedades utilizadas de la FRFT para el diseño de los operadores de extracción de características. Inicialmente, se introduce la definición de la FRFT para señales continuas y discretas, así como el algoritmo de cómputo utilizado. Posteriormente, se realiza la descripción de los operadores de extracción de características, las propiedades utilizadas en cada operador y el método de selección de parámetros en estos. Las características se seleccionan con el propósito de representar información con un reducido número de parámetros. En esta investigación, se utilizan parámetros de información promedia de la distribución de los datos así como de información de la locación de las señales en los dominios fraccionarios.

### **2.2.1 Transformada de Fourier fraccionaria.**

El análisis de Fourier es usado en ingeniería y en muchas ramas de la ciencia debido a que permite la representación de funciones en el dominio de la frecuencia. La FRFT es una generalización de la transformada de Fourier ordinaria en donde una función  $x(t)$  es mapeada a un dominio intermedio

Tabla 2.

*Etiquetas asignadas a dispositivos en uno o varios estados de operación y número total de eventos transitorios registrados por cada etiqueta. Este número incluye los transitorios registrados también cuando el dispositivo se encuentra funcionando en combinación con otros dispositivos como se muestra en la tabla 3. Un evento transitorio corresponde a la conexión o desconexión de un dispositivo a la vez, mientras los demás dispositivos de una combinación dada permanecen conectados.*

<b>Etiqu.</b>	<b>Descripción</b>	<b>Marca</b>	<b>Potencia Nominal [W]</b>	<b>estados</b>	<b>eventos transitorios</b>
1	Lámpara halógena	Opalux	50	2	198
2	Licuada velocidad 1	Oster	450	2	360
3	Licuada velocidad 2	Oster	450	2	320
4	Bombilla	Sylvania	75	2	394
5	Refrigerador ciclo 1	Haceb	1150	2	359
6	Secador de cabello temperatura 1	Conair	1875	2	200
7	Secador de cabello temperatura 2	Conair	1875	2	200
8	Plancha	Recco	1200	2	400
9	Sandwichera	Simplify	750	2	360
10	Cargado de celular	Motorola	24	2	340
11	Computador de escritorio (secuencia 1: Ver anexo C)	Dell	216	2	200
12	Refrigerador ciclo 2	Haceb	1150	2	200
13	Televisor	Sony	90	2	290
14	Computador de escritorio (secuencia 2: Ver anexo C)	Dell	216	4	400
15	Computador de escritorio (secuencia 3: Ver anexo C)	Dell	216	2	200
16	CFL	Daiku	20	2	402
17	CFL	Interlight	9	2	466
18	Ventilador	Silver Cool	48	2	458
19	Lámpara halógena	Blulighting	70	2	628
20	Lámpara LED	Nextlight	7	2	380
21	Cargador de PC	Asus	24	2	120
<b>Total</b>					<b>6875</b>

Tabla 3.

*Etiquetas asignadas para las combinaciones de uno, dos o tres dispositivos y número de señales registradas en estado estacionario por cada combinación. Las descripción de las etiquetas utilizadas para cada combinación se encuentra en la tabla 2.*

Dispositivo según tabla 2			Combinación de dos cargas				Combinación de tres cargas				
Etiqu.	Etiqu. tabla 2	Eventos	Etiqu.	Etiqu. tabla 2	Eventos	Etiqu.	Etiqu. tabla 2			Eventos	
1	1	99	20	17	10	100	48	19	16	20	375
2	2	100	21	17	18	200	49	4	19	17	375
3	3	100	22	17	19	200	50	4	17	9	375
4	4	97	23	17	13	200	51	20	16	8	375
5	5	100	24	10	18	100	52	17	4	2	375
6	6	100	25	10	19	200	53	19	20	18	375
7	7	100	26	10	13	100	54	19	20	13	375
8	8	100	27	18	19	200	55	16	4	21	375
9	9	100	28	18	13	200	56	19	18	9	375
10	10	100	29	19	13	200	57	16	2	8	375
11	11	100	30	17	3	100	58	4	9	13	375
12	12	100	31	17	8	200	59	19	5	21	375
13	14	400	32	17	5	100	60	19	18	13	375
14	15	100	33	10	3	100	61	16	2	21	375
15	16	101	34	10	8	100	62	2	9	13	375
16	17	108	35	10	5	100					
17	18	104	36	18	3	200					
18	19	104	37	18	8	200					
19	20	110	38	18	5	200					
			39	19	3	200					
			40	19	8	200					
			41	19	5	200					
			42	13	3	200					
			43	13	8	200					
			44	13	5	200					
			45	3	8	200					
			46	3	5	200					
			47	8	5	200					
<b>Total</b>		<b>2223</b>	<b>Total</b>		<b>4800</b>		<b>Total</b>				<b>5625</b>

entre el tiempo y la frecuencia (Almeida, 1994). La FRFT en el dominio  $u$  con orden  $\alpha$  de  $x(t)$  se define en (McBride y Kerr, 1987) como

$$\mathcal{F}_\alpha[x(t)](u) = C_\alpha \int_{-\infty}^{\infty} K_\alpha(t, u)x(t)dt, \quad (2.1)$$

$$K_\alpha(t, u) = e^{-i\pi u^2 \cot\alpha} e^{2i\pi tu/\sin\alpha} e^{-i\pi t^2 \cot\alpha}. \quad (2.2)$$

Con  $C_\alpha = \frac{e^{i(s(\alpha)\frac{\pi}{4}-\frac{\alpha}{2})}}{\sqrt{|\sin\alpha|}}$  y  $s(\cdot)$  como la función signo. Cuando  $\alpha = \pi/2$  la FRFT es la transformada de Fourier ordinaria, para  $\alpha = 0$  la FRFT es la función  $x(u)$  y se tiene que para  $\alpha = \pi$  es  $x(-u)$ . La FRFT cumple propiedades de conmutatividad, asociatividad y aditividad de índices, la cual es,  $\mathcal{F}_{\alpha_1}[\mathcal{F}_{\alpha_2}[x](u)] = \mathcal{F}_{\alpha_1+\alpha_2}[x](u)$  (Ozaktas, Kutay, y Mendlovic, 1999).

Debido a que las señales de corriente registradas en la base de datos son señales discretas, es necesario utilizar una definición de transformada de Fourier fraccionaria discreta. En (R. Á. T. Amarís, 2004) se propone una discretización del dominio original y del dominio transformado para la FRFT. Con esto, la transformada fraccionaria de Fourier discreta de una señal  $x[n\Delta t]$ ,  $n \in Z$ , con periodo de muestreo  $\Delta t$  se define en (R. Á. T. Amarís, 2004) como

$$x_\alpha[m] = \mathcal{F}_\alpha(x[n\Delta t])[m\Delta u_\alpha] = e^{i\pi(\frac{m}{M\Delta t})^2 \frac{\sin 2\alpha}{2}} \sum_{n=-M/2}^{M/2-1} x(n\Delta t) e^{i\pi(n\Delta t)^2 \cot\alpha} e^{-\frac{2i\pi mn}{M}}, \quad (2.3)$$

con  $m \in Z$ .  $M$  es el número de muestras a utilizar en el cálculo de la transformada de Fourier fraccionaria. Para un orden fraccionario dado,  $\alpha$ , este número de muestras se calcula junto con los valores  $\Delta t$  y  $\Delta u_\alpha$  a partir de la duración de la señal en el tiempo,  $T$ , y el ancho de banda del

espectro de Fourier, B, como:

$$\Delta u_\alpha = \frac{|\sin\alpha|}{T}, \quad (2.4)$$

$$\Delta t = \frac{|\sin\alpha|}{B}, \quad (2.5)$$

$$M = \frac{BT}{|\sin\alpha|}. \quad (2.6)$$

Por lo tanto, el cómputo de la FRFT en los operadores de extracción de características fue realizado con el diagrama de flujo del algoritmo se muestra en la figura 3. Es necesario interpolar al vector de entrada por un factor de dos (2) para evitar traslapes en las frecuencias debido al efecto de multiplicar el vector por una función tipo chirp ( $e^{i\gamma t^2}$ ) (Ozaktas, Arikan, Kutay, y Bozdogt, 1996). El diezmado se realiza para que el vector de salida tenga el mismo tamaño que el vector de entrada.

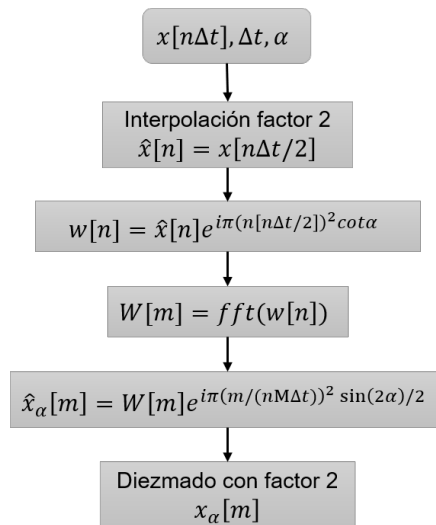


Figura 3. Algoritmo de cómputo de la FRFT basado en la definición de FRFT para señales discretas de la ecuación 2.3.

### 2.2.2 Extracción de características utilizando la relación entre la distribución de Wigner-Ville y la FRFT

La distribución de Wigner-Ville puede ser interpretada como una operación que indica la distribución de energía de la señal en el tiempo  $t$  y en la frecuencia  $\mu$ , y es por esto que también

recibe el nombre de distribución de energía. En (Ozaktas y cols., 1999), se define la distribución de Wigner-Ville  $W[x](t, \mu)$  de una señal  $x(t)$  como

$$W[x](t, \mu) = \int_{-\infty}^{\infty} x(t + y/2)x^*(t - y/2)e^{-2i\pi\mu y} dy. \quad (2.7)$$

Asimismo, se ha mostrado que la relación entre  $W[x](t, \mu)$  y la distribución de Wigner-Ville de la FRFT de orden  $\alpha$  de  $x(t)$ ,  $W[\mathcal{F}_\alpha[x]](t, \mu)$ , está dada por (Almeida, 1994)

$$W[\mathcal{F}_\alpha[x]](t, \mu) = W[x](t\cos\alpha - \mu\sin\alpha, t\sin\alpha + \mu\cos\alpha). \quad (2.8)$$

Así entonces,  $W[\mathcal{F}_\alpha[x]](t, \mu)$  se puede obtener a partir de  $W[x](t, \mu)$  al realizar una rotación en sentido antihorario con un ángulo  $\alpha$  de la distribución de energía de la señal, por lo que la FRFT corresponde a realizar una rotación a la distribución de energía de  $x(t)$ . La figura 4 ilustra esta relación utilizando una función con soporte compacto, donde  $\Omega$  es el soporte de  $x(t)$  y  $\Omega_\alpha$  es el soporte de  $\mathcal{F}_\alpha[x]$ , entonces  $\Omega_\alpha$  es una versión rotada de  $\Omega$  con un ángulo  $\alpha$ .

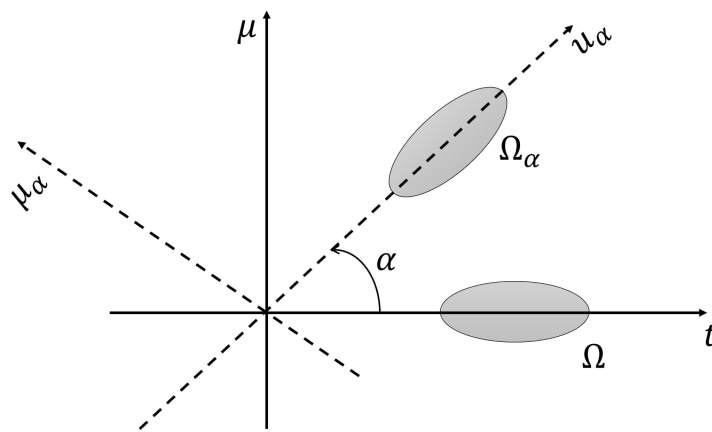


Figura 4. Relación entre las distribuciones de energía de  $x(t)$  y  $\mathcal{F}_\alpha[x(t)]$  con un orden fraccionario  $\alpha$ .

La magnitud de la FRFT de orden  $\alpha$  se relaciona con la distribución de Wigner-Ville  $W[\mathcal{F}_\alpha(x)](t, \mu)$

por la integral de proyección (Ozaktas y cols., 1999)

$$\int_{-\infty}^{\infty} W[\mathcal{F}_\alpha(x)](t, \mu) d\mu = |\mathcal{F}_\alpha[x(t)]|^2. \quad (2.9)$$

De esta forma se pueden obtener proyecciones a diferentes ángulos de la distribución de energía de  $x(t)$  con la magnitud de la FRFT a diferentes órdenes. La extracción de características consiste en calcular proyecciones a diferentes ángulos de la distribución de energía de las señales de corriente y posteriormente extraer parámetros de dichas proyecciones. Uno de los operadores de extracción de características propuestos en este trabajo de investigación,  $V_1$ , está conformado por la media, la mediana, la desviación estándar (std), la curtosis, el factor de asimetría ("skewness"), el valor máximo y los cuartiles  $q1$  y  $q3$  de cada proyección. Con  $|x_{\alpha_i}[m]|^2 = |\mathcal{F}_{\alpha_i}(x[n\Delta t])|^2$  y  $O = \{\alpha_1, \alpha_2, \dots, \alpha_o\}$ , se define el operador de extracción de características  $V_1(\cdot)$  como

$$V_1(x[n\Delta t], O) = [v_1^T(x[n\Delta t], \alpha_1), v_1^T(x[n\Delta t], \alpha_2), \dots, v_1^T(x[n\Delta t], \alpha_o)]^T, \quad (2.10)$$

con

$$v_1(x[n\Delta t], \alpha_i) = [media(|x_{\alpha_i}[m]|^2), mediana(|x_{\alpha_i}[m]|^2), std(|x_{\alpha_i}[m]|^2), curtosis(|x_{\alpha_i}[m]|^2), skewness(|x_{\alpha_i}[m]|^2), max(|x_{\alpha_i}[m]|^2), q1(|x_{\alpha_i}[m]|^2), q3(|x_{\alpha_i}[m]|^2)]^T. \quad (2.11)$$

En el operador de extracción de características en 2.10 es necesario definir el conjunto de órdenes fraccionarios, por lo que este conjunto es una variable de diseño en el operador. Para un conjunto finito de órdenes fraccionarios dado por  $S = \{\beta_i \mid i = 1, 2, \dots, s\}$  y con  $O = \{\alpha_i \mid i = 1, 2, \dots, o \wedge \alpha_i \in S \wedge \alpha_i \neq \alpha_j \wedge j = i + 1, i + 2, \dots, o\}$ , se requiere encontrar el subconjunto  $O$  que produzca la mayor separabilidad entre las clases de la base de datos. Como criterio

de separabilidad, se puede definir una función objetivo que maximice la distancia entre medias de las clases y minimice la desviación de los eventos pertenecientes a la misma clase. El anterior criterio, es utilizado como una pre-selección de características para definir los órdenes fraccionarios en el operador de extracción de características. Considerando  $x_{kz}$  como el evento  $k$  asociado a la etiqueta  $z$  y  $\overline{V}_1^{(z)}(O)$  como el vector medio de características de la clase  $z$  para el conjunto  $O$ , se define el problema de optimización para encontrar el conjunto de órdenes óptimos  $O^*$ :

$$O^* = \underset{O}{\operatorname{argmax}} \frac{\sum_{z=1}^{N_c-1} \sum_{i=z+1}^{N_c} \|\overline{V}_1^{(z)}(O) - \overline{V}_1^{(i)}(O)\|^2}{\sum_{z=1}^{N_c} \sum_{k=1}^{N_z} \|V_1(x_{kz}, O) - \overline{V}_1^{(z)}(O)\|^2}. \quad (2.12)$$

Donde  $N_c$  es el número total de etiquetas y  $N_z$  es el número de eventos de la  $z$ -ésima etiqueta. La resolución de 2.12 es realizada utilizando algoritmos genéticos los cuales se describen en el anexo B.

### 2.2.3 Extracción de características utilizando el centroide espectral fraccionario

El centroide espectral es considerado una estimación del 'centro de masa' de las componentes frecuenciales presentes en una señal. Se define como el promedio ponderado de las componentes frecuenciales utilizando como pesos la amplitud de cada frecuencia en la transformada de Fourier. El centroide espectral  $SC$  de una señal  $x[n\Delta t]$  está dado por

$$SC(x_{\frac{\pi}{2}}[m]) = \frac{\sum_{m=0}^{M-1} |x_{\frac{\pi}{2}}[m]| u_{\frac{\pi}{2}}[m]}{\sum_{m=0}^{M-1} |x_{\frac{\pi}{2}}[m]|}, \quad (2.13)$$

donde  $x_{\frac{\pi}{2}}[m] = \mathcal{F}_{\frac{\pi}{2}}(x[n\Delta t])[m\Delta u_{\frac{\pi}{2}}]$  es equivalente a la transformada de Fourier de  $x[n\Delta t]$  y  $u_{\frac{\pi}{2}}[m]$  es la frecuencia dada por la posición  $m$ .

La anterior definición se puede utilizar para definir un centro de masa para la representación de

la señal  $x[n\Delta t]$  en diferentes dominios fraccionarios como

$$SC_{\alpha}(x_{\alpha}[m]) = \frac{\sum_{m=0}^{M-1} |x_{\alpha}[m]| u_{\alpha}[m]}{\sum_{m=0}^{M-1} |x_{\alpha}[m]|}, \quad (2.14)$$

donde  $u_{\alpha}[m]$  es el dominio transformado y  $x_{\alpha}[m] = \mathcal{F}_{\alpha}(x[n\Delta t])[m\Delta u_{\alpha}]$ . Con esto, la extracción de características propuesta consiste en calcular el centroide espectral fraccionario de las señales para un conjunto de dominios fraccionarios. En cada dominio se segmenta la señal en 10 partes y se calcula el centroide espectral fraccionario de cada segmento, luego se concatenan los valores obtenidos en todos los dominios para formar el vector de características.

Para un conjunto de órdenes fraccionarios  $O = \{\alpha_1, \alpha_2, \dots, \alpha_o\}$  se define el operador de extracción de características  $V_2(\cdot)$  como

$$V_2(x[n\Delta t], O) = [v_2^T(x[n\Delta t], \alpha_1), v_2^T(x[n\Delta t], \alpha_2), \dots, v_2^T(x[n\Delta t], \alpha_o)]^T, \quad (2.15)$$

con

$$v_2(x[n\Delta t], \alpha_j) = [SC_{\alpha_j}(r_1[m]), SC_{\alpha_j}(r_2[m]), \dots, SC_{\alpha_j}(r_{10}[m])]^T.$$

Donde  $r_k[m]$  representa el segmento  $k$  de la señal  $x_{\alpha_j}[m]$ .

La selección de órdenes fraccionarios se realiza de la misma forma presentada en la sección

2.2.2. Dado los conjuntos de órdenes fraccionarios  $S$  y  $O$ , se desea encontrar  $O^*$  como

$$O^* = \operatorname{argmax}_O \frac{\sum_{z=1}^{N_c-1} \sum_{i=z+1}^{N_c} \|\overline{V}_2^{(z)}(O) - \overline{V}_2^{(i)}(O)\|^2}{\sum_{z=1}^{N_c} \sum_{k=1}^{N_z} \|V_2(x_{kz}, O) - \overline{V}_2^{(z)}(O)\|^2}. \quad (2.16)$$

### 2.3 Modelos de estimación

El proceso de estimación en los sistemas de identificación de cargas consiste en utilizar un modelo para asignar etiquetas a los vectores de características que ingresen al modelo. Con esto, la etiqueta

asignada se interpreta como un dispositivo o una combinación de cargas en particular. En los sistemas implementados, se utilizaron modelos de SVMs y ANNs que corresponden a algoritmos de aprendizaje de máquina supervisado y su diseño se realizó utilizando la base de datos descrita en la sección 2.1. A continuación, se describen los conceptos básicos y las variables de decisión en los modelos de SVMs y ANNs, así como el proceso de diseño y evaluación de los algoritmos de estimación.

### 2.3.1 Máquinas de Soporte Vectorial

Las máquinas de soporte vectorial (SVMs) son un modelo de discriminante lineal inicialmente diseñado para problemas de identificación de dos clases linealmente separables. El modelo de las SVMs se define como (Christopher, 2016)

$$\Psi_{\Theta}(v) = W^T \phi(v) + b \quad (2.17)$$

Donde el vector  $W$  y el escalar  $b$  son los parámetros del modelo de estimación,  $\Theta = [W^T, b]^T$ , y la función  $\phi(\cdot)$  mapea el vector  $v$  a un vector de dimensión mayor. Los parámetros del modelo se determinan bajo el criterio de optimización de máximo margen, siendo el margen la distancia del hiperplano  $\Psi_{\Theta}(v)$  a los vectores de características más cercanos. Una vez obtenido el modelo, los vectores de entrada se clasifican con base al signo de  $\Psi_{\Theta}$  (Christopher, 2016), (Duda, Hart, y Stork, 2012).

Para un conjunto de vectores de características  $v_1, v_2, \dots, v_N$  con etiquetas  $l_k = \{-1, 1\}$ , donde  $l_k = -1$  si  $v_k$  pertenece a la clase A y  $l_k = 1$  si pertenece a la clase B, se define el problema de optimización con criterio de máximo margen (Christopher, 2016)

$$W^*, b^* = \operatorname{argmax}_{W, b} \left\{ \frac{1}{\|W\|} \min_k (\{l_k(W^T \phi(v_k) + b) \mid k = 1, 2, \dots, N\}) \right\}. \quad (2.18)$$

Utilizando la función dual, el problema de optimización en (2.18) se puede reescribir como (Christopher, 2016)

$$W^*, b^* = \operatorname{argmin}_{W, b} \frac{1}{2} \|W\|^2, \quad (2.19)$$

$$s.t. \quad l_k(W^T \phi(v_k) + b) \geq 1, \quad k = 1, \dots, N.$$

Cuando el problema es linealmente no separable se modifica el problema óptimo para aceptar algunos datos mal clasificados en el entrenamiento. La modificación que se realiza consiste en incorporar una función de penalización regulada con una constante  $C$  para los datos mal clasificados. Con esta modificación, el problema óptimo es reescrito como (Christopher, 2016), (Duda y cols., 2012)

$$W^*, b^* = \operatorname{argmin}_{W, b} \frac{1}{2} \|W\|^2 + C \sum_{k=1}^N \xi_k, \quad (2.20)$$

$$s.t. \quad l_k(W^T \phi(v_k) + b) \geq 1, \quad k = 1, \dots, N.$$

Donde  $\xi_k = 0$  cuando el vector de características está en la frontera del margen o dentro de la zona correcta de su clase y  $\xi_k = |l_k - \Psi_{\Theta}(v_k)|$  para los demás casos.

En (Christopher, 2016) y (Duda y cols., 2012) el problema de optimización 2.20 se resuelve utilizando las condiciones de Karush-Kunh-Tucker, y se obtiene que

$$W^* = \sum_{k=1}^N a_k l_k \phi(v_k), \quad (2.21)$$

donde  $0 \leq a_k \leq C$  son los multiplicadores de Lagrange. Los vectores de características con  $a_k \neq 0$  se denominan vectores de soporte, los cuales contribuyen en la predicción de una nueva muestra. Siendo  $IVS$  el conjunto de los índices de los vectores de soporte y  $IC$  el conjunto de datos para los cuales  $0 < a_n < C$ , el cálculo de  $b^*$  está dado por (Christopher, 2016)

$$b^* = \frac{1}{N_{IC}} \sum_{k \in IC} \left( l_k - \sum_{n \in IVS} a_n l_n \phi(v_n)^T \phi(v_k) \right). \quad (2.22)$$

Con estos resultados, el modelo de la SVMs para clasificar una nueva muestra es

$$\Psi_{\Theta^*}(v) = \sum_{n=1}^N a_n t_n \Phi(x, x_n) + b^*, \quad (2.23)$$

con  $\Phi(x, x_n) = \phi(v_n)^T \phi(v)$  como la función Kernel.

Dado que el problema de optimización 2.20 es convexo, las soluciones del algoritmo de optimización  $b^*$  y  $W^*$  son minimizadores globales con cualquier inicialización. Con esto, el diseño del algoritmo de SVMs se centra en encontrar el valor de la constante  $C$ , definir el kernel y sus parámetros.

En este trabajo de investigación se utiliza como kernel la función de base radial  $\Phi(v, v_k) = \exp(-\|v - v_n\|^2)$  debido a que es frecuentemente utilizada con éxito en los sistemas de identificación propuestos en la literatura revisada. El modelo binario de las SVMs se extrapola al problema multiclase de identificación de cargas con la estrategia uno contra todos y la validación del modelo se realiza con la exactitud en identificación en validación cruzada para 10 particiones. El proceso de selección del parámetro  $C$  consiste en diseñar un conjunto de modelos de SVMs utilizando diferentes valores de  $C$  y seleccionar el modelo con mayor exactitud en validación cruzada, los

valores de  $C$  que se utilizan en este trabajo son  $\{2^0, 2^1, \dots, 2^{20}\}$ .

### 2.3.2 Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (ANNs) son un modelo basado en capas de neuronas con discriminantes lineales y funciones de activación diferenciables. La salida de la  $j$ -ésima neurona ubicada en la  $n$ -ésima capa es (Christopher, 2016), (Duda y cols., 2012):

$$z_j = h^{(n)} \left( \sum_i w_{ji}^{(n)} q_i + w_{j0}^{(n)} \right). \quad (2.24)$$

Donde  $w_{ji}$  es el parámetro que relaciona la entrada  $q_i$  con la neurona  $j$  y  $w_{j0}$  es un término constante de la neurona  $j$  en la capa  $n$ . La función  $h^{(n)}(\cdot)$  es una función de activación la cual generalmente se selecciona de las funciones sigmoide, tangente hiperbólica o identidad (Christopher, 2016), (Duda y cols., 2012). Debido al tamaño de la base de datos de vectores de características, en el diseño de los algoritmos de ANNs se contemplan únicamente topologías de una y dos capas ocultas las cuales se ilustran en la figura 5. Para una capa oculta (capa (1)), la  $n$ -ésima salida, en la capa de salida (capa (2)), para un vector de entrada  $v$  con  $N_v$  características  $q_1, q_2, \dots, q_{N_v}$  y un vector  $W$ , correspondiente a los pesos de la capa oculta (1) y la capa de salida (2), si  $h^{(2)}$  es la función identidad, está dada por (Christopher, 2016):

$$y_n(v, W) = \sum_{z=1}^Z w_{nz}^{(2)} h^{(z)} \left( \sum_{i=1}^{N_v} w_{zi}^{(1)} q_i + w_{z0}^{(1)} \right) + w_{n0}^{(2)}, \quad (2.25)$$

y para una red con 2 capas ocultas es

$$y_n(v, W) = \sum_{j=1}^J w_{nj}^{(3)} h^{(j)} \left( \sum_{z=1}^Z w_{jz}^{(2)} h^{(z)} \left( \sum_{i=1}^{N_v} w_{zi}^{(1)} q_i + w_{z0}^{(1)} \right) + w_{j0}^{(2)} \right) + w_{n0}^{(3)}. \quad (2.26)$$

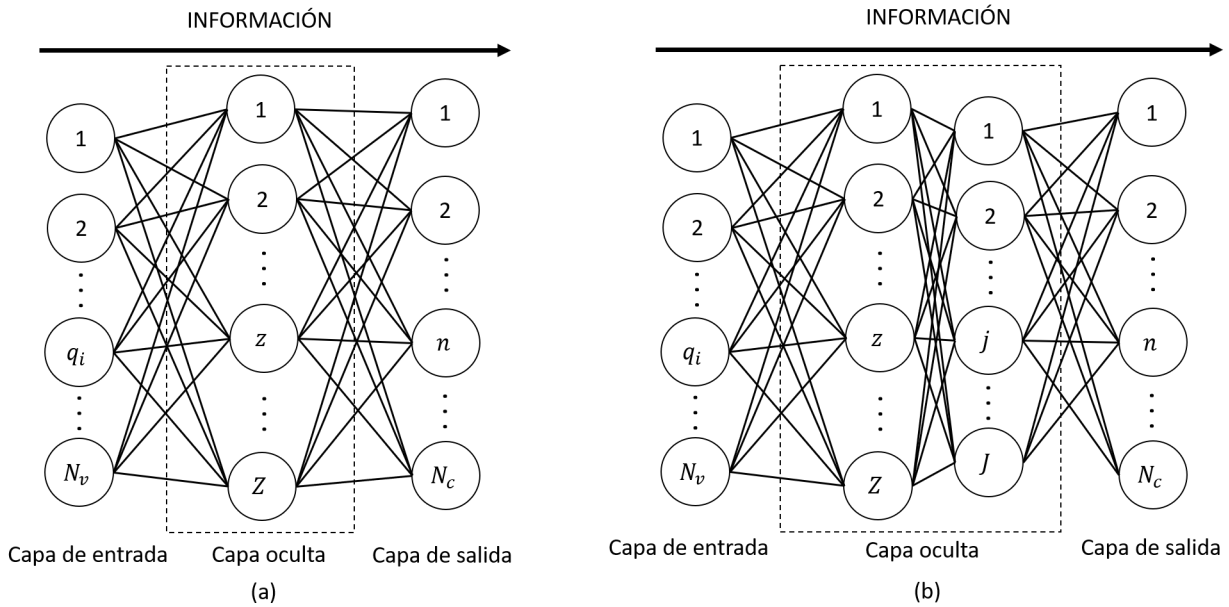


Figura 5. a) Topología de una Red Neuronal Artificial con una capa oculta y  $Z$  neuronas. b) Topología de una Red Neuronal Artificial con dos capas ocultas con  $Z$  y  $J$  neuronas respectivamente.

Donde  $Z$  y  $J$  son el número de neuronas de cada capa oculta. Estos modelos se pueden interpretar como modelos no lineales con entrada  $v$  y salidas  $y_k$  controlados por el vector de parámetros  $W$ . El entrenamiento del modelo consiste en optimizar un criterio y encontrar los valores de los parámetros utilizando una base de datos con etiquetas conocida. Como criterio de optimización para el entrenamiento del sistema, se minimiza el error cuadrático medio entre la predicción y la etiqueta real de los datos de entrenamiento, por lo tanto el problema de optimización está dado por (Christopher, 2016)

$$W^* = \underset{W}{\operatorname{argmin}} \frac{1}{2} \sum_{k=1}^N \sum_{n=1}^{N_c} \|y_n(v_k, W) - l_{kn}\|^2. \quad (2.27)$$

Para la solución de 2.27 se utiliza el método de gradiente descendiente evaluando las derivadas con el método de propagación hacia atrás.

En el diseño de ANNs es necesario seleccionar la arquitectura y la función de activación en cada neurona. Diferentes topologías de red, parámetros iniciales y funciones de activación se consideran en el diseño del sistema, seleccionando la combinación con mayor exactitud en los datos de prueba. Posteriormente, con la topología y funciones de activación ya definidas se entrena un conjunto de modelos utilizando diferentes parámetros iniciales y se selecciona el modelo de mayor exactitud en validación cruzada con 10 particiones. En el diseño se probaron 33 topologías con neuronas  $[Z]$  para una capa oculta y  $[Z \ J]$  para dos capas ocultas, las topologías contempladas (Hagan, Demuth, Beale, y cols., 1996) son  $[4]$ ,  $[6]$ ,  $[8]$ ,  $[10]$ ,  $[12]$ ,  $[14]$ ,  $[16]$ ,  $[18]$ ,  $[20]$ ,  $[25]$ ,  $[30]$ ,  $[2 \ 2]$ ,  $[2 \ 3]$ ,  $[2 \ 4]$ ,  $[2 \ 5]$ ,  $[5 \ 3]$ ,  $[5 \ 2]$ ,  $[5 \ 4]$ ,  $[5 \ 5]$ ,  $[3 \ 2]$ ,  $[2 \ 10]$ ,  $[2 \ 15]$ ,  $[2 \ 20]$ ,  $[5 \ 10]$ ,  $[5 \ 15]$ ,  $[5 \ 20]$ ,  $[10 \ 10]$ ,  $[10 \ 5]$ ,  $[10 \ 2]$ ,  $[10 \ 15]$ ,  $[15 \ 10]$ ,  $[15 \ 15]$ ,  $[20 \ 20]$ . La figura 6 muestra el pseudo-código del algoritmo de diseño del modelo final en ANNs.

#### 2.4 Selección del conjunto de características

El proceso de selección de características se utiliza para determinar los subconjuntos de características con los mejores resultados en exactitud en la identificación. Generalmente, el subconjunto seleccionado es de menor dimensionalidad que el conjunto original, por lo que se reduce el tamaño del vector de entrada en los algoritmos de identificación y por ende su complejidad computacional. Para la selección de características se implementa un proceso de eliminación de características recursivo utilizando como métrica de eliminación el error del modelo entrenado en la identificación deseada. El proceso consiste en calcular  $N_v$  características para  $N$  señales de la base de datos de experimentos y así construir una matriz de características  $MC$  con dimensiones  $N \times N_v$ . Con esta matriz de características se entrena un modelo de estimación cuyo error de identificación se utiliza para determinar la importancia de una característica dada. A continuación, los

**Selección de topología y función de activación en ANN**

**Entradas** : Matriz de características, etiquetas

**Inicializar** :

33 Topologías

2 funciones de activación: tan-sigmoide y log-sigmoide

**Para**  $j = 1$  hasta 2 hacer

**Para**  $i = 1$  hasta 33 hacer

**Para**  $k = 1$  hasta 10 hacer

construir ANN con topología  $i$ .

Usar función de activación  $j$  en todas las neuronas.

Inicializar parámetros aleatoriamente.

dividir datos aleatoriamente con 70% entrenamiento, 15% validación y 15% de prueba.

entrenar red.

Calcular error de identificación en los datos de prueba.

**Salidas** : Topología y función de activación con el mínimo error.

**Modelo final**

**Entradas** : Matriz de características, etiquetas, topología y función de activación.

**Para**  $j = 1$  hasta 20 hacer

construir ANN con topología y función de activación

Usar función de activación  $j$  en todas las neuronas.

Inicializar parámetros aleatoriamente.

definir 10 bloques de validación cruzada.

**para**  $k = 1$  hasta 10 hacer validación cruzada

entrenar red con 90% de los datos.

Calcular error de identificación con el 10% restante.

Promediar error de las 10 iteraciones.

**Salidas** : Parámetros del modelo con el mínimo error promedio en validación cruzada.

Figura 6. algoritmos de diseño del modelo final de ANNs para los sistemas de identificación desarrollados.

elementos de la  $j$ -ésima columna de  $MC$  (o valores de la  $j$ -ésima característica para  $N$  señales) se permutan de posición aleatoriamente, generando una matriz permutada  $MC^{(j)}$ . El modelo entrenado con  $MC$  se utiliza para estimar las etiquetas de  $MC^{(j)}$ . El porcentaje de error de esta estimación corresponde a la importancia de la  $j$ -ésima característica. De esta manera, se calcula la importancia de cada característica. Las características con menor importancia, entre el valor mínimo de importancia y 102 % del valor mínimo, son eliminadas de la matriz  $MC$ . Con esta nueva matriz se entrena un nuevo modelo y se repite el proceso de cálculo de importancia y eliminación de características hasta que el número de características seleccionadas se haya reducido logrando que el error de validación cruzada del modelo entrenado presente una baja tendencia a disminuir o un aumento en un valor considerable (Sadeghianpourhamami y cols., 2017). La figura 7 muestra el proceso general del método de selección de características.

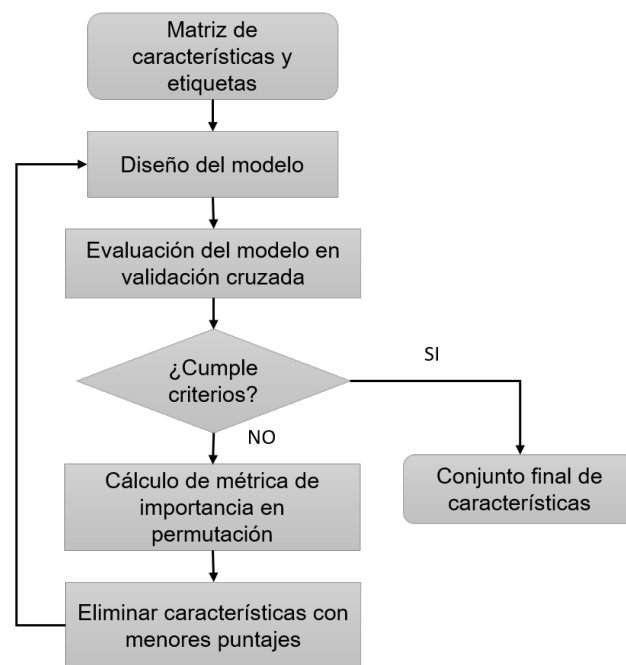


Figura 7. Diagrama de flujo del algoritmo de eliminación de características recursivas hacia atrás

## 2.5 Sistemas propuestos para la identificación de cargas eléctricas basados en la transformada de Fourier fraccionaria

A continuación, se muestran los resultados obtenidos en el proceso de diseño de los sistemas de identificación con extracción de características a partir de la FRFT para el régimen estacionario y transitorio de los dispositivos a identificar. Se muestran los resultados obtenidos en el proceso de selección de órdenes fraccionarios en los operadores de extracción de características, selección del conjunto final de características y los parámetros de los modelos de estimación.

### 2.5.1 Modelos de estimación para información de régimen estacionario

Con la base de datos de eventos en régimen estacionario se resolvieron los problemas óptimos 2.12 y 2.16, obteniendo dos subconjuntos de 10 órdenes fraccionarios cada uno. De un conjunto  $S = \{0.250\pi, 0.265\pi, 0.280\pi, \dots, 1.750\pi\}$  los 20 órdenes seleccionados fueron  $\{0.685\pi, 0.835\pi, 0.910\pi, 0.925\pi, 0.955\pi, 0.970\pi, 0.985\pi, 1.000\pi, 1.015\pi, 1.030\pi, 1.045\pi, 1.090\pi, 1.105\pi, 1.135\pi, 1.150\pi, 1.210\pi, 1.225\pi, 1.285\pi, 1.360\pi, 1.510\pi\}$ . Además, se incluyeron 20 órdenes equidistantes  $\{0.250\pi, 0.325\pi, 0.400\pi, \dots, 1.750\pi\}$  y 10 órdenes aleatorios para un total de 50 órdenes fraccionarios para iniciar el proceso de extracción de características. El conjunto de características por cada orden fraccionario consiste de 18 características. Como se consideran 50 órdenes fraccionarios, se evalúan 900 características en total. La matriz de datos para entrenamiento y validación consta de 12648 experimentos con 900 características por cada experimento. Con esta información se realiza el proceso de selección de características con los modelos de identificación.

El proceso de selección de características se realiza para encontrar los conjuntos de característi-

cas finales en los modelos de SVMs y ANNs, las figuras 8 y 9 muestran el número de características y el error de validación cruzada durante cada iteración en este proceso de selección. Con los resultados obtenidos, se definieron los conjuntos finales de características que se listan en las tablas 4 y 5. En el diseño del algoritmo final de SVMs se obtiene  $C = 512$  y en el diseño de ANNs la topología que genera el menor error en la predicción es  $[20 \ 20]$  con función de activación tangente sigmoide.

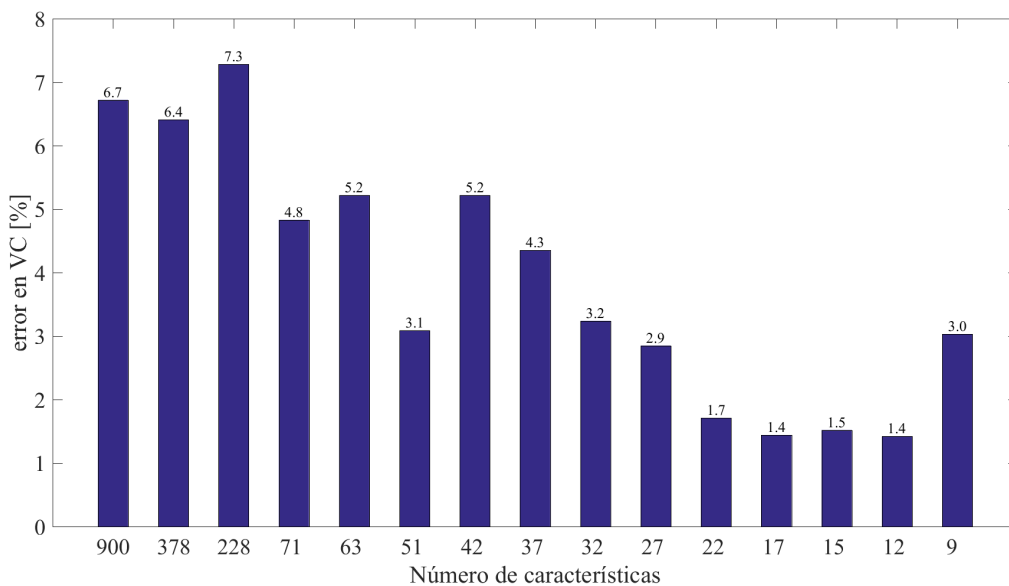


Figura 8. Error en validación cruzada de cada iteración en el proceso de selección de características utilizando SVMs para experimentos en régimen estacionario. Se determina en este caso seleccionar 12 características.

### 2.5.2 Modelos de estimación para información de régimen transitorio

Del conjunto  $S = \{0.250\pi, 0.265\pi, 0.280\pi, \dots, 1.750\pi\}$  se seleccionaron 20 órdenes  $\{0.850\pi, 0.865\pi, 0.925\pi, 0.940\pi, 0.955\pi, 0.970\pi, 0.985\pi, 1.000\pi, 1.015\pi, 1.030\pi, 1.060\pi, 1.135\pi, 1.180\pi, 1.195\pi, 1.210\pi, 1.285\pi, 1.360\pi, 1.435\pi, 1.510\pi, 1.660\pi\}$  de la misma forma que en el régimen estacionario, pero utilizando los eventos de régimen transitorio. Asimismo, fueron incluidos 20

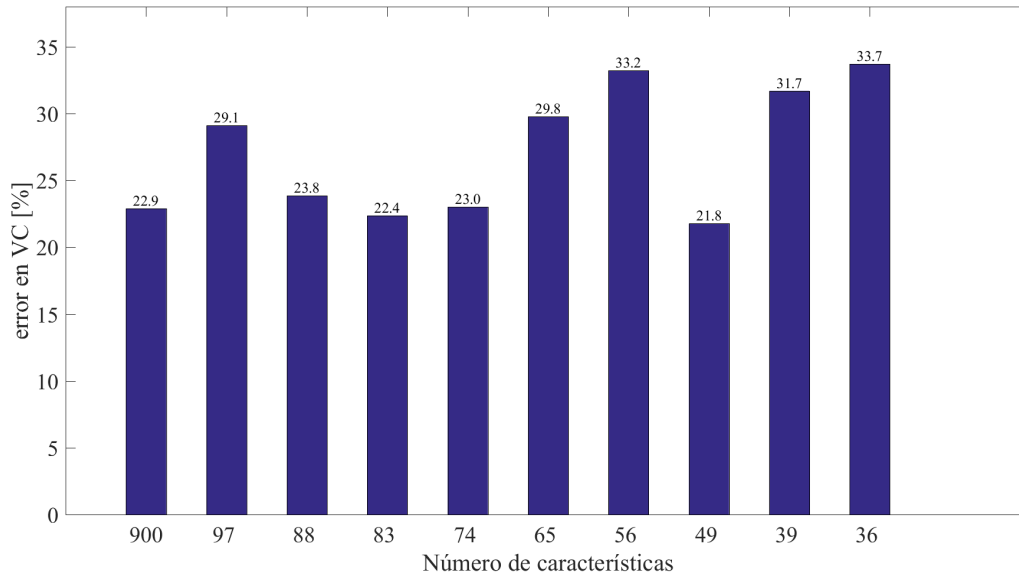


Figura 9. Error en validación cruzada de cada iteración en el proceso de selección de características utilizando ANNs para experimentos en régimen estacionario. Se determina en este caso seleccionar 49 características.

Tabla 4.

*Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen estacionario utilizando SVMs*

Orden	Características
$0.045\pi$	$SC_{\alpha}(r_5)$
$0.486\pi$	$SC_{\alpha}(r_6)$
$0.985\pi$	$SC_{\alpha}(r_6)$
$1.000\pi$	mediana, skewness, max, p1, $SC_{\alpha}(r_1)$ , $SC_{\alpha}(r_3)$ , $SC_{\alpha}(r_4)$ , $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_{10})$

Tabla 5.

*Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen estacionario utilizando ANNs*

<b>Orden</b>	<b>Características</b>
0.045 $\pi$	skewness, $SC_{\alpha}(r_5)$
0.134 $\pi$	$SC_{\alpha}(r_5)$
0.308 $\pi$	$SC_{\alpha}(r_6)$
0.407 $\pi$	skewness, $SC_{\alpha}(r_6)$
0.486 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
0.565 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
0.644 $\pi$	$SC_{\alpha}(r_6)$
0.802 $\pi$	$SC_{\alpha}(r_5)$
0.835 $\pi$	$SC_{\alpha}(r_5)$
0.925 $\pi$	curtosis, $SC_{\alpha}(r_6)$
0.955 $\pi$	$SC_{\alpha}(r_6)$
0.960 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
0.970 $\pi$	skewness, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
0.985 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.000 $\pi$	media, mediana, std, curtosis, skewness, max, p1, p3
1.015 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.030 $\pi$	curtosis, skewness, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.039 $\pi$	$SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.045 $\pi$	$SC_{\alpha}(r_5)$
1.105 $\pi$	skewness
1.118 $\pi$	$SC_{\alpha}(r_5)$
1.135 $\pi$	$SC_{\alpha}(r_5)$
1.150 $\pi$	$SC_{\alpha}(r_5)$
1.434 $\pi$	curtosis, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.510 $\pi$	curtosis, $SC_{\alpha}(r_5)$

órdenes equidistantes  $\{0.250\pi, 0.325\pi, 0.400\pi, \dots, 1.750\pi\}$  y 10 órdenes aleatorios. La matriz de datos para entrenamiento y validación de los sistemas está formada por 6875 experimentos con 900 características por cada experimento. Con esta matriz de información se inicia el proceso de selección de características en cada algoritmo de identificación de eventos transitorios.

El proceso de selección de características en cada iteración se presenta en las figuras 10 y 11 y los conjuntos finales de características en las tablas 6 y 7 para los sistemas de eventos transitorios con SVMs y ANNs respectivamente. En el diseño del algoritmo de SVMs se tiene  $C = 16$  y en el de ANNs la topología que se selecciona es [30] con función de activación tangente sigmoide.

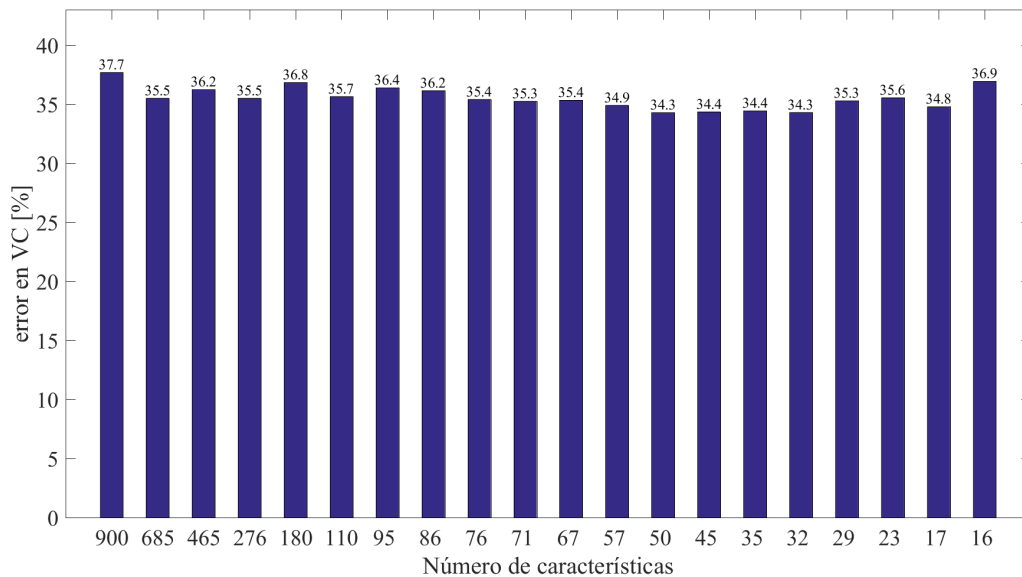


Figura 10. Error en validación cruzada de cada iteración en el proceso de selección de características utilizando SVMs para experimentos en régimen transitorio. Se determina en este caso seleccionar 17 características.

El proceso de diseño propuesto permite encontrar conjuntos de características y modelos de estimación de forma semiautomática, en donde el usuario debe definir un conjunto reducido de

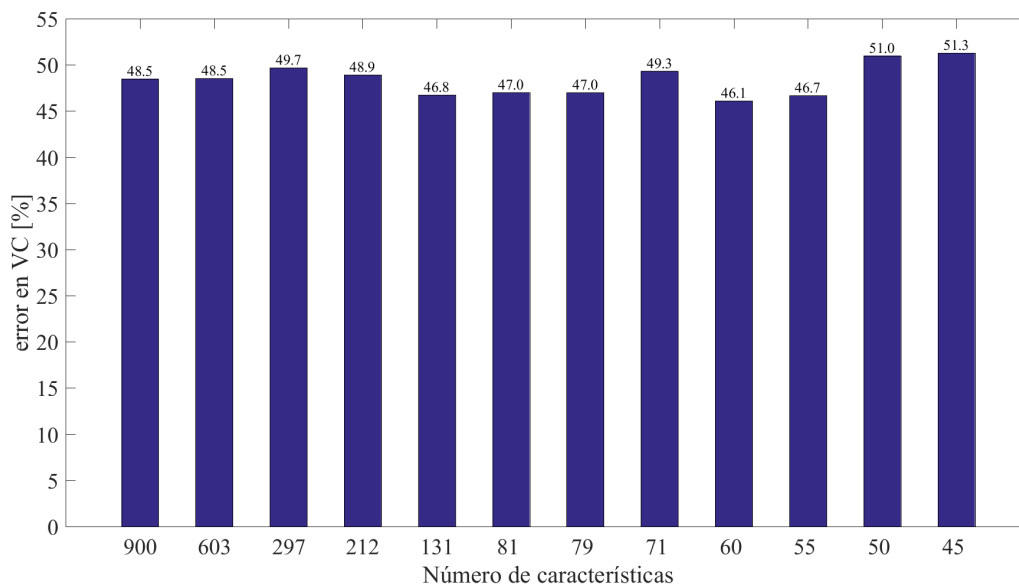


Figura 11. Error en validación cruzada de cada iteración en el proceso de selección de características utilizando ANNs para experimentos en régimen transitorio. Se determina en este caso seleccionar 55 características.

Tabla 6.

*Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen transitorio utilizando SVMs*

Orden	Características
$0.045\pi$	skewness, $SC_{\alpha}(r_4)$ , $SC_{\alpha}(r_6)$
$0.304\pi$	$SC_{\alpha}(r_2)$
$0.565\pi$	$SC_{\alpha}(r_4)$
$0.970\pi$	skewness
$0.985\pi$	skewness, $SC_{\alpha}(r_6)$
$1.000\pi$	std, skewness, max, p3
$1.015\pi$	skewness, $SC_{\alpha}(r_3)$ , $SC_{\alpha}(r_4)$ , $SC_{\alpha}(r_6)$
$1.045\pi$	skewness

Tabla 7.

*Conjunto de características finales utilizando la FRFT para el sistema de identificación de cargas de eventos en régimen transitorio utilizando ANNs*

<b>Orden</b>	<b>Características</b>
0.045 $\pi$	skewness, $SC_{\alpha}(r_3)$ , $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
0.134 $\pi$	Curtosis, skewness
0.154 $\pi$	$SC_{\alpha}(r_5)$
0.208 $\pi$	skewness
0.291 $\pi$	curtosis
0.407 $\pi$	skewness
0.644 $\pi$	$SC_{\alpha}(r_5)$
0.910 $\pi$	skewness
0.925 $\pi$	curtosis
0.955 $\pi$	$SC_{\alpha}(r_6)$
0.960 $\pi$	curtosis, skewness, $SC_{\alpha}(r_6)$
0.970 $\pi$	skewness, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_8)$
0.985 $\pi$	skewness, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$ , $SC_{\alpha}(r_8)$
1.000 $\pi$	media, mediana, std, skewness, max, p1, p3, $SC_{\alpha}(r_5)$
1.015 $\pi$	skewness, $SC_{\alpha}(r_2)$ , $SC_{\alpha}(r_3)$ , $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$ , $SC_{\alpha}(r_{10})$
1.030 $\pi$	skewness, $SC_{\alpha}(r_5)$ , $SC_{\alpha}(r_6)$
1.039 $\pi$	curtosis, skewness
1.090 $\pi$	$SC_{\alpha}(r_5)$
1.105 $\pi$	$SC_{\alpha}(r_5)$
1.118 $\pi$	curtosis, skewness
1.285 $\pi$	curtosis
1.434 $\pi$	$SC_{\alpha}(r_6)$
1.510 $\pi$	$SC_{\alpha}(r_6)$
1.513 $\pi$	skewness, $SC_{\alpha}(r_5)$
1.592 $\pi$	$SC_{\alpha}(r_5)$
1.750 $\pi$	kurtosis, $SC_{\alpha}(r_6)$

parámetros para su implementación. La evaluación de los sistemas finales en los 4 escenarios contemplados (régimenes estacionarios y transitorios y SVMs y ANNs) se realiza en el siguiente capítulo del documento, en donde además los resultados son comparados con propuestas de extracción de características basados en la transformada S y Wavelet.

### **3. Análisis del desempeño de los sistemas de identificación de cargas eléctricas diseñados**

En este capítulo se presentan los resultados obtenidos con los algoritmos de identificación finales utilizando la FRFT, los cuales son comparados con propuestas basadas en las transformadas Wavelet y S. Dos sistemas independientes para el régimen estacionario y transitorio de las señales son implementados con cada propuesta de extracción de características y modelo de estimación (SVMs o ANNs). Como resultados se muestra el error en la estimación global y por cada clase de los algoritmos en validación cruzada.

#### **3.1 Resultados de los sistemas utilizando la FRFT**

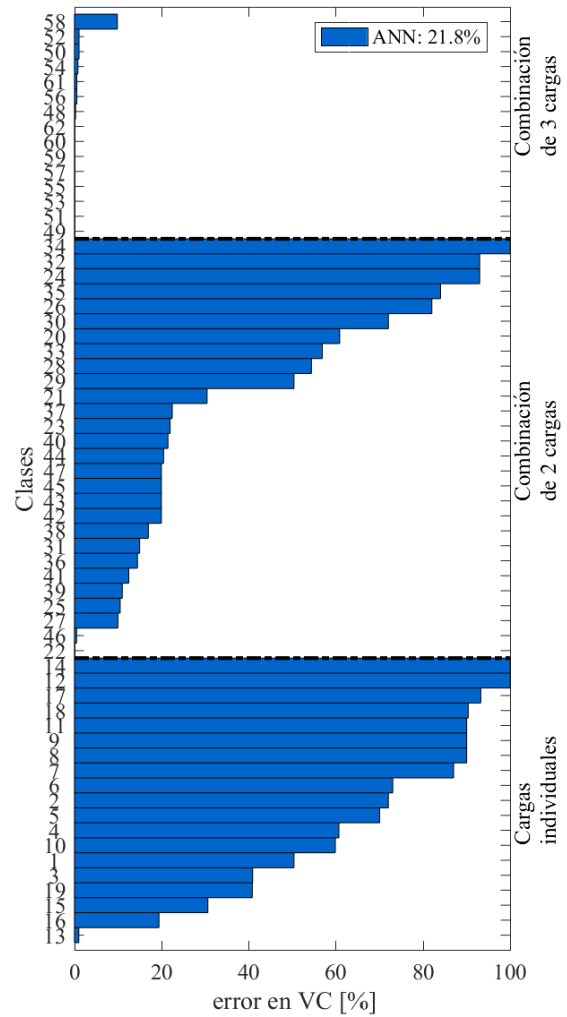
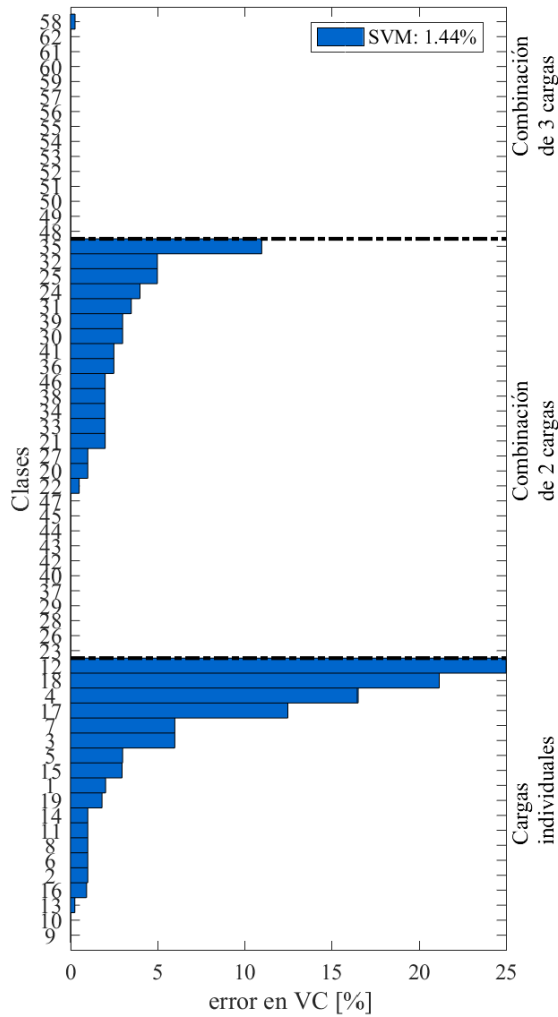
Con los sistemas propuestos en la sección 2.5.1 para el régimen estacionario y en la sección 2.5.2 para el régimen transitorio, se calcula el error de identificación en validación cruzada por clase y global. Las figuras 12 y 13 presentan el error en identificación de los sistemas con SVMs y ANNs para el régimen estacionario y transitorio respectivamente. La etiqueta asignada en los sistemas de eventos de régimen transitorio representan “*dispositivo.estado*”. Por otra parte, las etiquetas en los sistemas de régimen estacionario están asociadas a las combinaciones descritas en la tabla 3. En los resultados obtenidos, se observa que los sistemas con SVMs logran un mejor rendimiento en exactitud en comparación con los de ANNs para la identificación de eventos del régimen estacio-

nario y transitorio. En los resultados de los sistemas de eventos en régimen estacionario, se tiene que el modelo de ANNs se ajusta a las clases con mayor número de eventos, lo que indica que el modelo tiene un sesgo por el número de datos.

### 3.2 Resultados de los sistemas utilizando la transformada Wavelet

Con fines de comparación, se implementa la técnica de extracción de características propuesta en (Duarte y cols., 2012), la cual consiste en calcular la contribución de energía de un conjunto de escalas en la transformada Wavelet. Los parámetros de diseño: Frecuencia central de la función Wavelet madre,  $F_c$ , ancho de banda de 3 dB de la Wavelet madre,  $\Delta F_{3dB}$ , y el número de escalas son seleccionados en base a las propiedades de los eventos estacionarios y transitorios. El diseño de los modelos de estimación se realiza de igual forma que en el sistema basado en la FRFT, el cual es descrito en la sección 2.3. En este sistema no se realiza el proceso selección de características descrito en 2.4, dado que se implementa la propuesta presentada en (Duarte y cols., 2012).

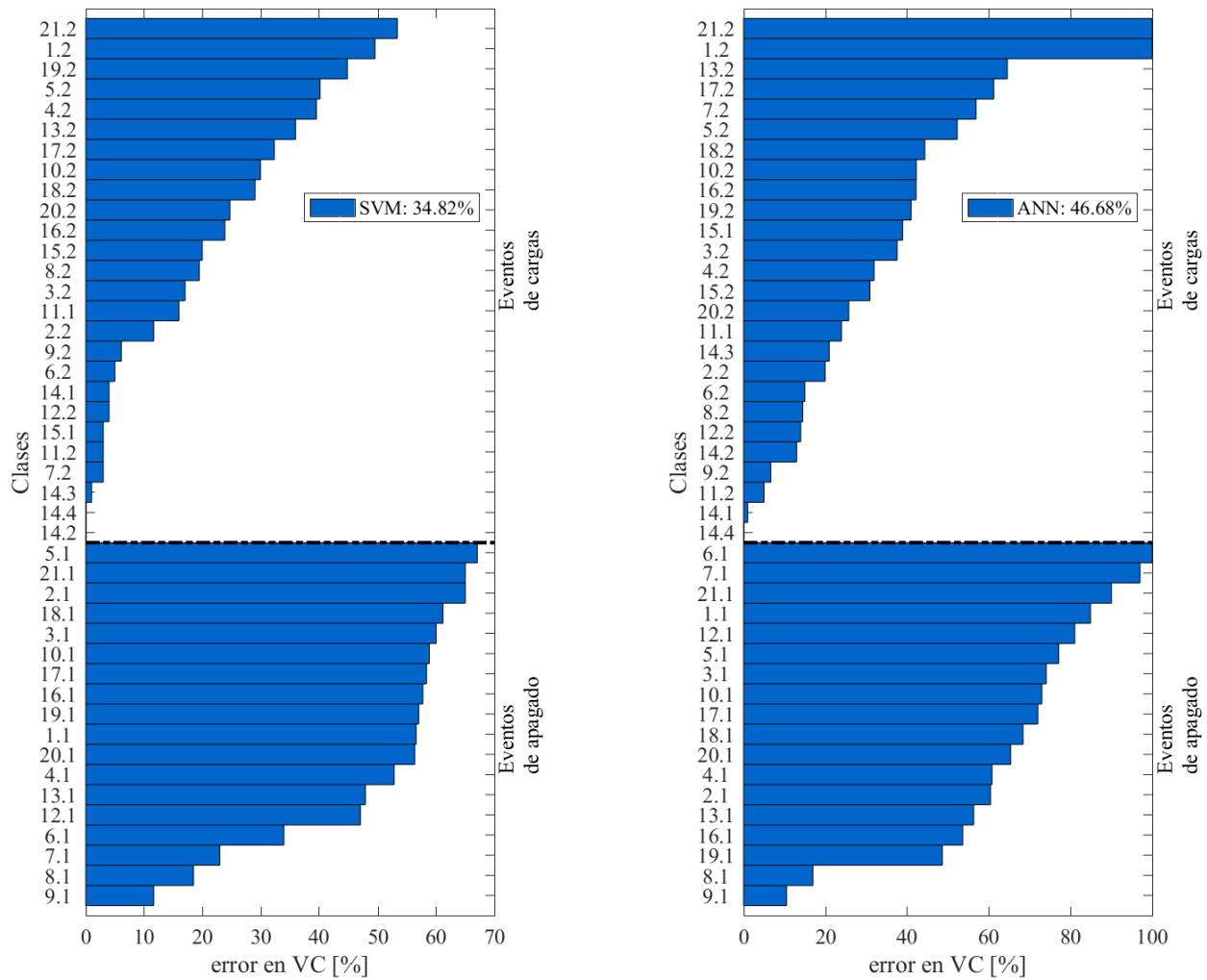
En el régimen estacionario se utilizan como parámetros de diseño de los sistemas  $\Delta F_{3dB} = 1[kHz]$  y  $F_c = 10[kHz]$ . Con estos valores en los parámetros, se tiene el rango de frecuencias de  $40[Hz]$  hasta  $10[kHz]$  con 55 escalas en la transformada. El vector de características contiene elementos que corresponden a la contribución de energía de cada escala, con esto la matriz de datos está formada por 12648 experimentos con 55 características por cada experimento. En el algoritmo de SVMs se utiliza  $C = 256$  y en las ANNs la topología seleccionada corresponde a una capa oculta con 20 neuronas y ( o [20]) y con función de activación tangente sigmoide. La figura 14 muestra el error en identificación global y por clase de los algoritmos con el conjunto de características basado en la transformada Wavelet para eventos del régimen estacionario.



(a) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con las características descritas en la tabla 4 y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con las características descritas en la tabla 5 y estimación realizada con ANNs.

Figura 12. Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando los conjuntos finales de características de la FRFT y estimación realizada con ANNs y SVMs.



(a) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con las características descritas en la tabla 6 y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con las características descritas en la tabla 7 y estimación realizada con ANNs.

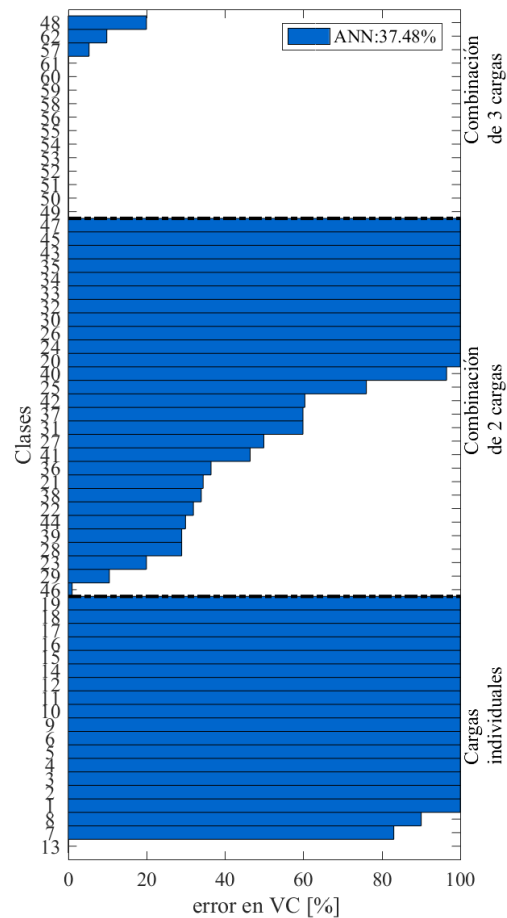
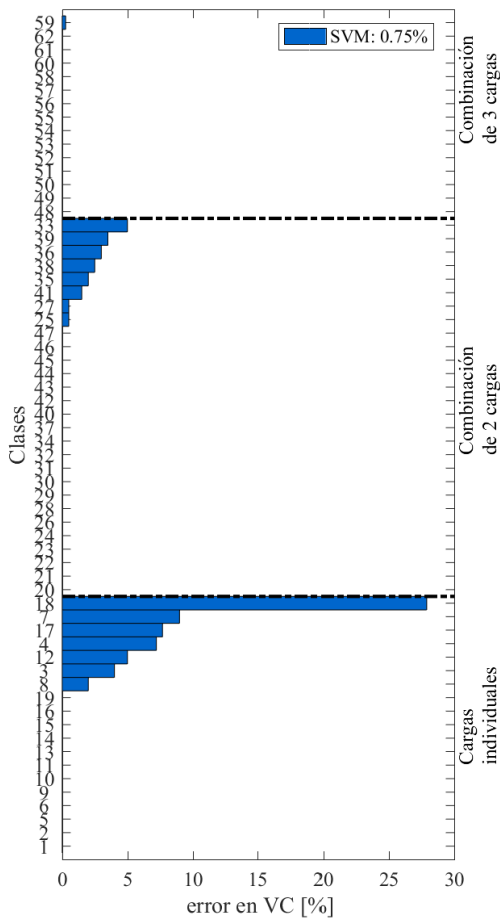
Figura 13. Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando los conjuntos finales de características de la FRFT y estimación realizada con ANNs y SVMs.

Para el régimen transitorio se utiliza  $\Delta F_{3dB} = 2[kHz]$  y  $F_c = 25[kHz]$ , lo cual corresponde a una frecuencia central igual a la mitad de la frecuencia de muestreo. Con estos valores en los parámetros se cubre el rango de frecuencias de  $60[Hz]$  hasta  $25[kHz]$  con 75 escalas en la transformada y la matriz de datos se conforma por 6875 experimentos con 75 características. En el algoritmo de SVMs se tiene  $C = 1024$  y en las ANNs la topología corresponde a una capa oculta con 18 neuronas (o [18]) y con función de activación tangente sigmoide. La figura 15 muestra el error en identificación global y por clase de los algoritmos con el conjunto de características basado en la transformada Wavelet para eventos del régimen transitorio.

### 3.3 Resultados de los sistemas utilizando la transformada S

Con fines de comparación se utiliza también la técnica de extracción de características propuesta en (Jimenez y cols., 2014), la cual consiste en obtener parámetros estadísticos de la representación tiempo-frecuencia en la transformada S. Para la implementación de la transformada S se utiliza el algoritmo propuesto en (Stockwell, Mansinha, y Lowe, 1996) y las frecuencias de análisis son seleccionadas con base en las propiedades de los eventos transitorios y estacionarios. Nuevamente el diseño de los modelos de estimación se realiza siguiendo el proceso descrito en la sección 2.3. En este sistema no se realiza el proceso selección de características descrito en 2.4, dado que se implementa la propuesta realizada en (Jimenez y cols., 2014).

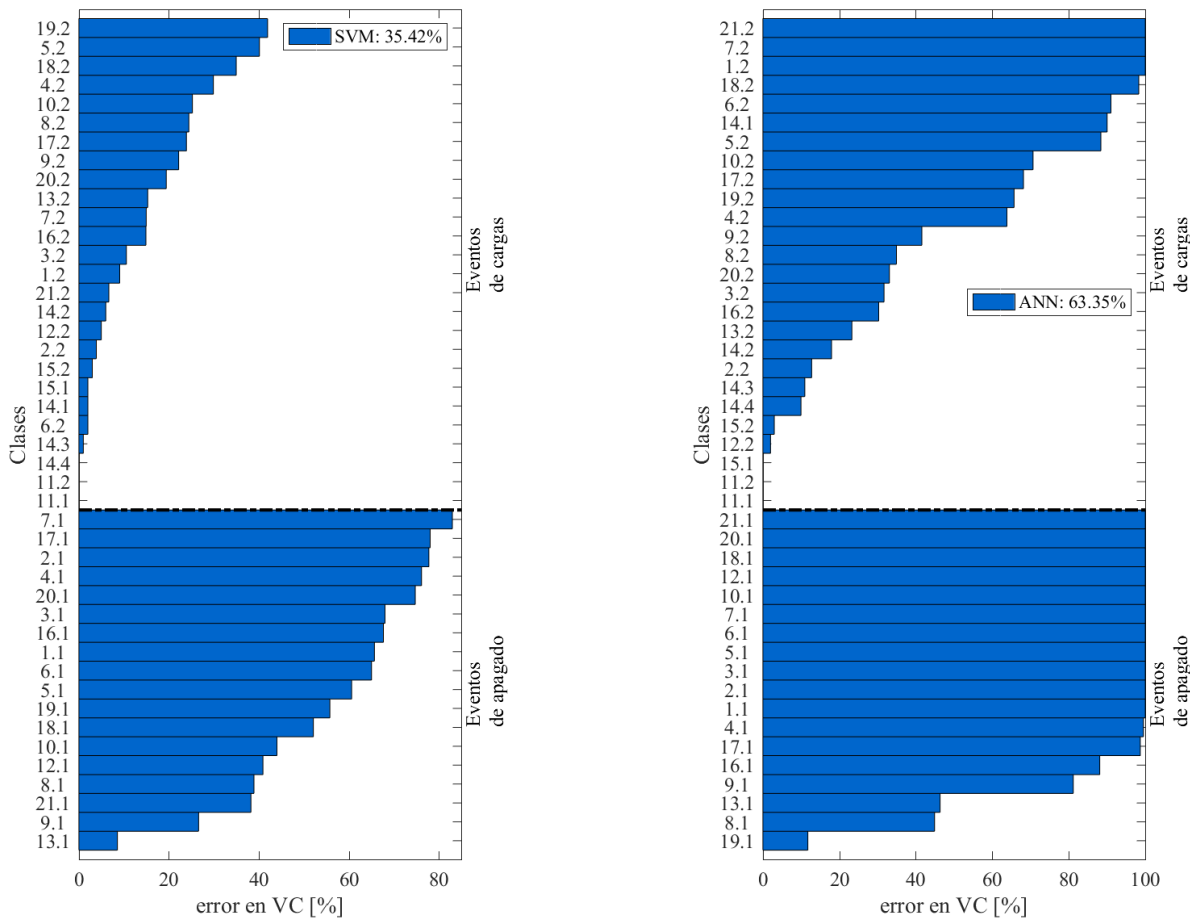
De las 7 características propuestas en (Jimenez y cols., 2014) se utilizan las 6 primeras características para el sistema de régimen estacionario, la séptima característica corresponde a la duración del transitorio y por lo tanto no se utiliza en este caso. El rango de frecuencias en la



(a) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con características extraídas con la transformada Wavelet y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con características extraídas con la transformada Wavelet y estimación realizada con ANNs.

Figura 14. Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando conjuntos de características extraídos con la transformada Wavelet y estimación realizada con ANNs y SVMs.



(a) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con características extraídas con la transformada Wavelet y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con características extraídas con la transformada Wavelet y estimación realizada con ANNs.

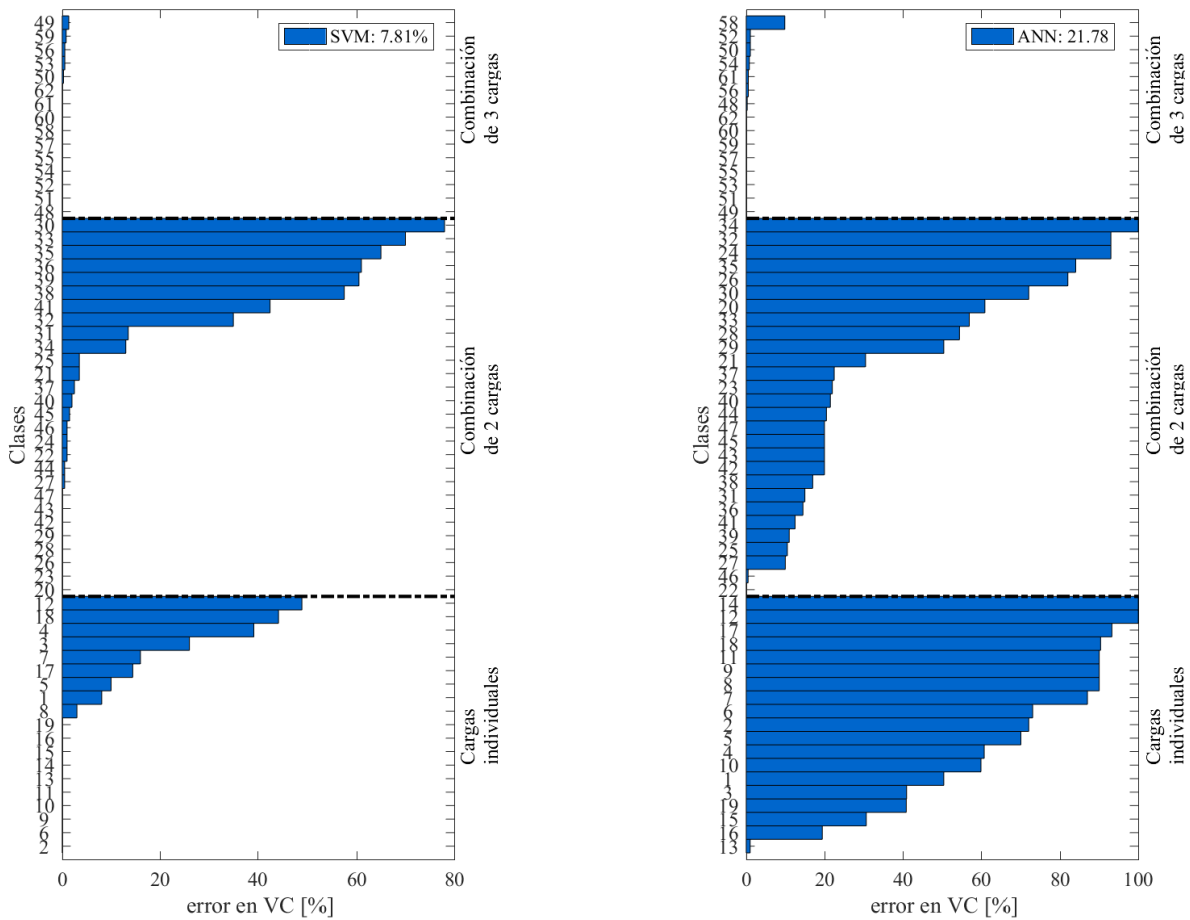
Figura 15. Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando conjuntos de características extraídos con la transformada Wavelet y estimación realizada con ANNs y SVMs.

transformada S es de  $0[Hz]$  hasta  $10[kHz]$  para 100 frecuencias equidistantes y la matriz de datos se conforma por 12648 observaciones con 6 características. En el algoritmo de SVMs se obtiene  $C = 128$  y en las ANNs la topología es  $[20 \ 20]$  con función de activación tangente sigmoide. La figura 16 muestra el error en identificación de los algoritmos con el conjunto de características basado en la transformada S en el régimen estacionario.

En el régimen transitorio se utilizan las 7 características propuestas en (Jimenez y cols., 2014) con rango de frecuencias de  $0[Hz]$  hasta  $25[kHz]$  para 100 frecuencias equidistantes, la matriz de datos consiste de 6875 observaciones con 7 características. En el algoritmo de SVMs se obtiene  $C = 64$  y la topología del algoritmo de ANNs es  $[20 \ 20]$  con función de activación tangente sigmoide. La figura 17 muestra el error en identificación de los algoritmos con el conjunto de características basado en la transformada S en el régimen transitorio.

### **3.4 Análisis de los resultados obtenidos en los sistemas de identificación**

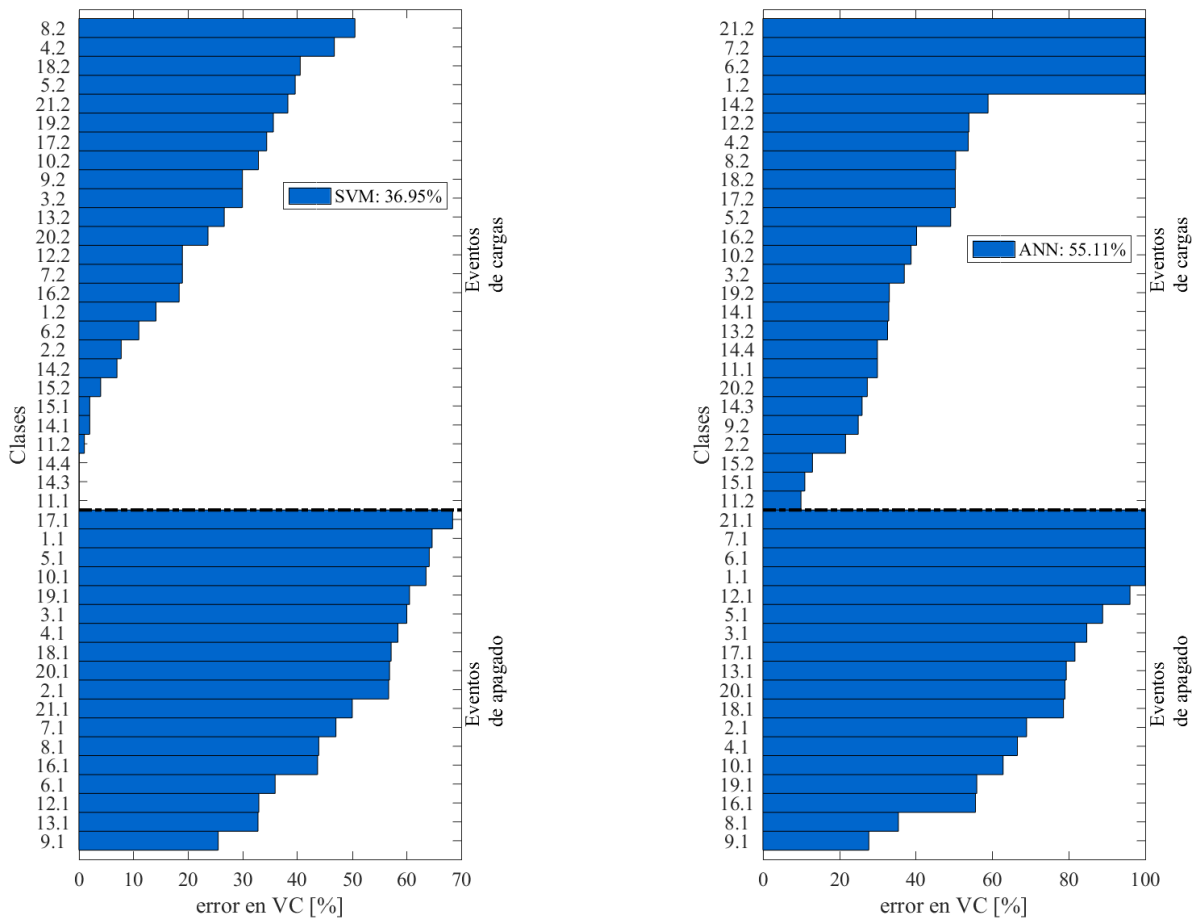
A continuación se presentan las tendencias observadas en los resultados de error en exactitud de los sistemas de identificación desarrollados. Se examinan separadamente los sistemas para el régimen estacionario y transitorio en búsqueda de relaciones en las combinaciones de cargas y categorías de los dispositivos. Los dispositivos son categorizados en: No electrónica-resistivo (para las cargas con etiquetas 1,4,6,7,8,9 y 19), No electrónica-Inductivo (para las cargas con etiquetas 2,3,5,12 y 18), y electrónica ( para las cargas con etiquetas 10,11,13,14,15,16,17,20 y 21) según su tipo (Jimenez y cols., s.f.). Esto con el fin de encontrar relaciones entre el error en la estimación y las posibles categorías basadas en las propiedades físicas de los dispositivos. Además, en la tabla 8 se muestran parámetros estadísticos de la distribución del error en la estimación por clases (o



(a) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con características extraídas con la transformada S y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen estacionario con características extraídas con la transformada S y estimación realizada con ANNs.

Figura 16. Error en validación cruzada de los sistemas de identificación de eventos de régimen estacionario utilizando conjuntos de características extraídos con la transformada S y estimación realizada con ANNs y SVMs.



(a) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con características extraídas con la transformada S y estimación realizada con SVMs.

(b) Error en validación cruzada del sistema de identificación de eventos de régimen transitorio con características extraídas con la transformada S y estimación realizada con ANNs.

Figura 17. Error en validación cruzada de los sistemas de identificación de eventos de régimen transitorio utilizando conjuntos de características extraídos con la transformada S y estimación realizada con ANNs y SVMs.

etiquetas) de los sistemas implementados con ANNs Y SVMs.

Tabla 8.

*Parámetros estadísticos de la distribución del error por clase de los sistemas de identificación con extracción de características basada en FRFT Wavelet y S y estimación con SVMs y ANNs.*

Estimación	Método	$e(\%)$	media	mediana	mín.	máx.	Desviación
SVM	FRFT	34.8	32.2	31.2	0	67.1	22.3
eventos	Wavelet	35.4	31.4	24.9	0	83.0	26.8
transitorios	S	37.0	33.1	33.7	0	68.4	21.2
ANNs	FRFT	46.7	47.6	46.6	0	100	29,56
eventos	Wavelet	63.4	62.7	75.8	0	100	38.7
transitorios	S	55.1	57.7	54.0	10	100	29.1
SVM	FRFT	1.4	2.5	0.9	0	25.0	4.9
eventos	Wavelet	0.8	1.3	0	0	27.8	4.0
estacionarios	S	7.8	11.7	0.5	0	78.0	21.4
ANNs	FRFT	21.8	37.2	21	0	100	35.6
eventos	Wavelet	37.48	58.3	68.3	0	100	43.1
estacionarios	S	21.8	37.2	21.0	0	100	35.6

### 3.4.1 Análisis de los resultados en identificación de los sistemas basados en eventos de régimen estacionario

Para el análisis de los sistemas de identificación con eventos de régimen estacionario, se toman las 5 clases (o etiquetas) con mayor error en la identificación de cada conjunto de combinaciones realizadas, de cada sistema. Esto, con el fin de establecer las clases más confundidas por los sistemas y su incidencia en los diferentes algoritmos de identificación. En la tabla 9 se presenta este ranking de error en la estimación de las cargas y combinaciones de carga para los algoritmos con estimación realizada con SVMs y los diferentes métodos de extracción de características (FRFT, Wavelet y S). En las figuras 18, 19 y 20 se presenta esta información para cada sistema con extracción de características basado en la FRFT, transformada Wavelet y transformada S respectivamente e identificación con SVMs.

En los resultados de los sistemas con estimación realizada con SVMs, se observa que las clases 12 (Refrigerador ciclo 2), 19 (lámpara halógena), 18 (ventilador) y 4 (bombilla) están presentes en el ranking de cargas individuales de todos los métodos de extracción de características. En estas cargas, el sistema con transformada Wavelet obtiene errores relativamente bajos (alrededor del 7 %) en la identificación de los dispositivos 4 y 12. El dispositivo 19 tiene errores de identificación superiores al 25 % en todos los sistemas de identificación implementados, por lo que se puede considerar el dispositivo más difícil de identificar. En los grupos de dos cargas, combinaciones con los dispositivos 10 (cargador de celular), 3 (licuadora velocidad 2), 5 (refrigerador ciclo 1) y 18 (ventilador) son las más frecuentes en el ranking. En el caso de la FRFT, todas las combinaciones de dos cargas en el ranking tienen un elemento de naturaleza electrónica (CFL o cargador de celular) de potencia menor a 24W, y su error en identificación es menor o igual al 11 % en todas las combinaciones. Con esto, el sistema presenta cierta dificultad en la estimación de dispositivos electrónicos de baja potencia cuando es combinado con otra carga. Por otra parte, la mayoría de combinaciones presentes en el ranking de los sistemas con transformadas Wavelet y S tienen dispositivos de categoría no electrónica-Inductivo. Lo anterior sugiere que ambos sistemas presentan alguna dificultad en la estimación de dispositivos de esta categoría cuando son combinados con otra carga. Para los grupos de tres cargas, se obtienen buenos resultados de error en la estimación de eventos de régimen estacionario (menor al 2 %), lo que se puede explicar por el gran número de experimentos utilizados en el diseño de estos.

En la tabla 10 se presenta el mismo estilo de rankings, pero para los sistemas con ANNs. Inicialmente, se observa que un gran conjunto de clases (o etiquetas) presentan un 100 % de error

Tabla 9.

Ranking de mayor a menor en base en el error de estimación para los sistemas con de extracción de características basada la FRFT y las transformadas Wavelet y S con estimación con SVMs. El ranking se realiza para cada conjunto de combinaciones de cargas y método de extracción.

Caract.	Carga individual			Dos cargas				Tres cargas				
	etq	e (%)	dispositivo	etq	e (%)	Combinación		etq	e (%)	Combinación		
FRFT	12	25.0	12	35	11.0	10	5	58	0.3	4	9	13
	18	21.1	19	32	5.0	17	5	-	-	-	-	-
	4	16.5	4	25	5.0	10	19	-	-	-	-	-
	17	12.5	18	24	4.0	10	18	-	-	-	-	-
	7	6.0	7	31	3.5	17	8	-	-	-	-	-
Wavelet	18	27.9	19	33	5.0	10	3	59	0.3	19	5	21
	7	9.0	7	39	3.5	19	3	-	-	-	-	-
	17	7.7	18	36	3.0	18	3	-	-	-	-	-
	4	7.2	4	38	2.5	18	5	-	-	-	-	-
	12	5	12	35	2.0	10	5	-	-	-	-	-
S	12	49.0	12	30	78.0	17	3	49	1.3	4	19	17
	18	44.2	19	33	70.0	10	3	59	0.8	19	5	21
	4	39.1	4	35	65.0	10	5	56	0.5	19	18	9
	3	26.0	3	36	61.0	18	3	53	0.5	19	20	18
	7	16.0	7	39	60.5	19	3	50	0.3	4	17	9

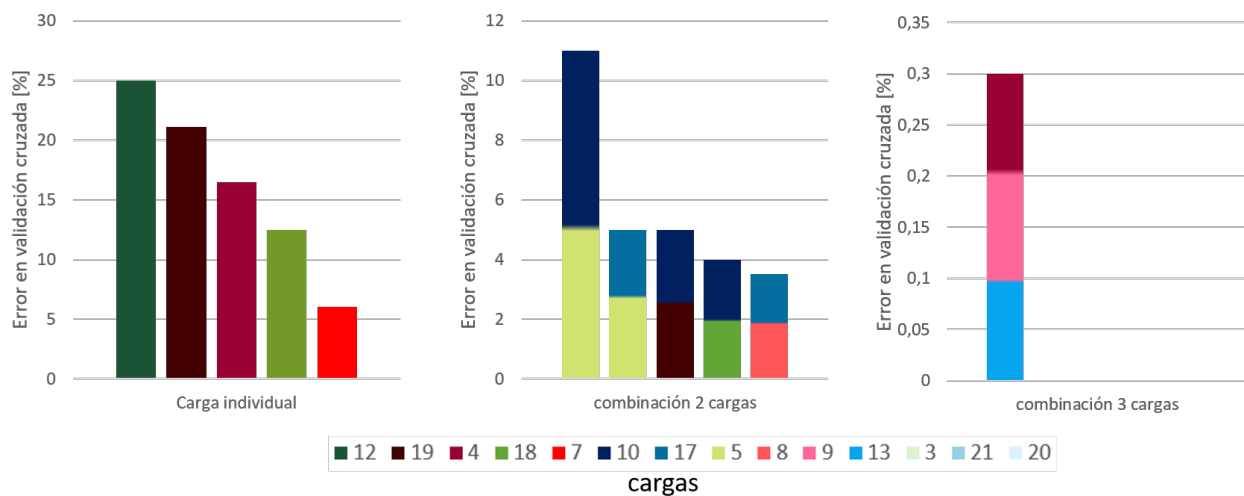


Figura 18. Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la FRFT y estimación con SVM

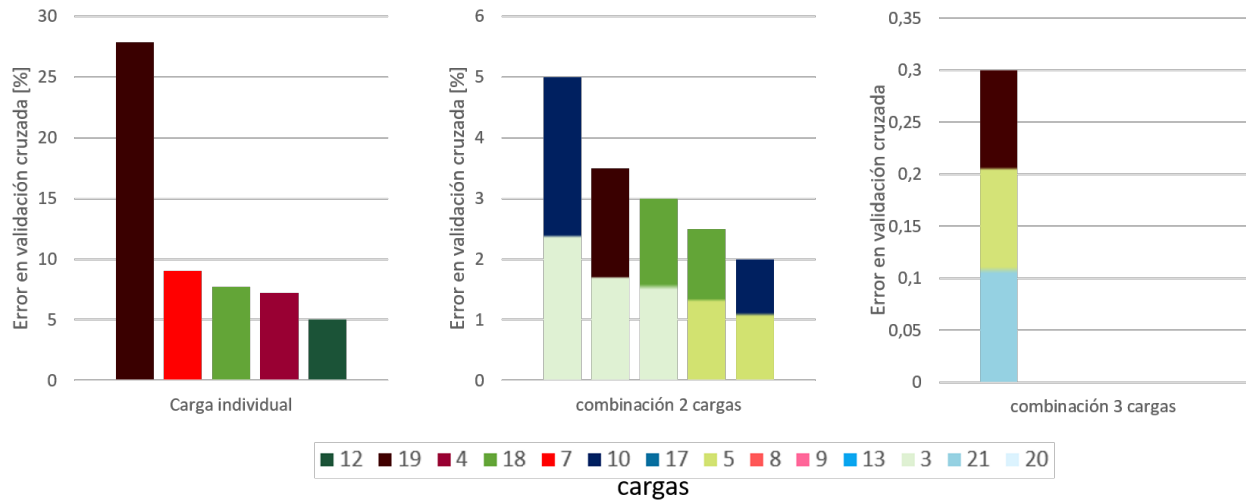


Figura 19. Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la transformada Wavelet y estimación con SVM

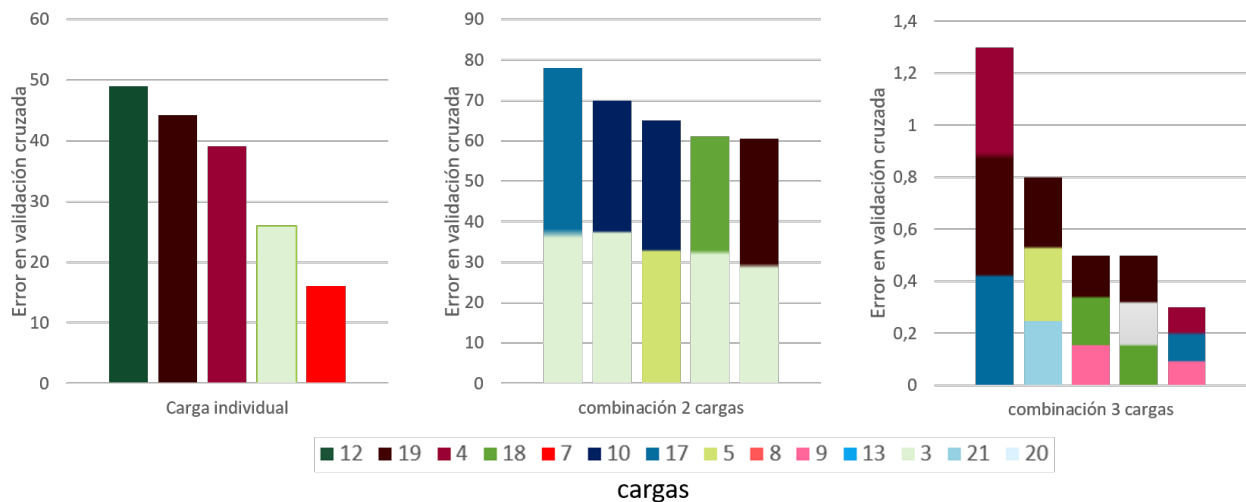


Figura 20. Error en la identificación de las cinco clases más confundidas en cada tipo de combinación de carga (una, dos o tres cargas combinadas) con extracción de características basado en la transformada S y estimación con SVM

en la estimación para todos los métodos de extracción de características. Lo que sugiere, que sin importar el conjunto de características, el modelo de estimación no se ajusta a un gran conjunto de clases en el problema de identificación. Esto puede suceder debido al gran número de clases en la salida, a las topologías de red propuestas, y al número de experimentos en la base de datos. Aun así, se pueden localizar tendencias en los resultados con ANNs como la presencia de los dispositivo 12, 17, 18 y 19 en los rankings individuales de cargas para todas los métodos de extracción. Con estimación con ANNs, en los grupos de dos cargas aparecen los dispositivos 8, 10 y 17 frecuentemente en los rankings de los métodos desarrollados, los cuales también son comunes cuando la estimación se realiza con SVMs. Así mismo, para las combinaciones con 3 cargas, los sistemas con ANNs presentan errores bajos de identificación.

Tabla 10.

*Ranking de 5 elementos de mayor a menor en base en el error de estimación para los sistemas con de extracción de características basada la FRFT y las transformadas Wavelet y S con estimación con ANNs. El ranking se realiza para cada conjunto de combinaciones de cargas y método de extracción.*

Caract.	Carga individual			Dos cargas				Tres cargas				
	etq	e (%)	dispositivo	etq	e (%)	Combinación		etq	e (%)	Combinación		
FRFT	14	100	15	34	100	10	8	58	9.8	4	9	13
	12	100	12	32	93.0	17	5	52	1.1	17	4	2
	17	93.3	18	24	93.0	10	18	50	1.1	4	17	9
	18	90.4	19	35	84.0	10	5	54	0.8	19	20	13
	11	90.0	11	26	82.0	10	13	61	0.5	16	2	21
Wavelet	19	100	20	47	100	8	5	48	20.0	19	16	20
	18	100	19	45	100	3	8	62	9.8	2	9	13
	17	100	18	43	100	13	8	57	5.3	16	2	8
	16	100	17	35	100	10	5	-	-	-	-	-
	15	100	16	34	100	10	8	-	-	-	-	-
S	14	100	15	34	100	10	8	58	9.8	4	9	13
	12	100	12	32	93.0	17	5	52	1.1	17	4	2
	17	93.2	18	24	93.0	10	18	50	1.1	4	17	9
	18	90.3	19	35	84.0	10	5	54	0.8	19	20	13
	11	90.0	11	26	82.0	10	13	61	0.5	16	2	21

### 3.4.2 Análisis de los resultados en identificación de los sistemas basados en eventos de régimen transitorio

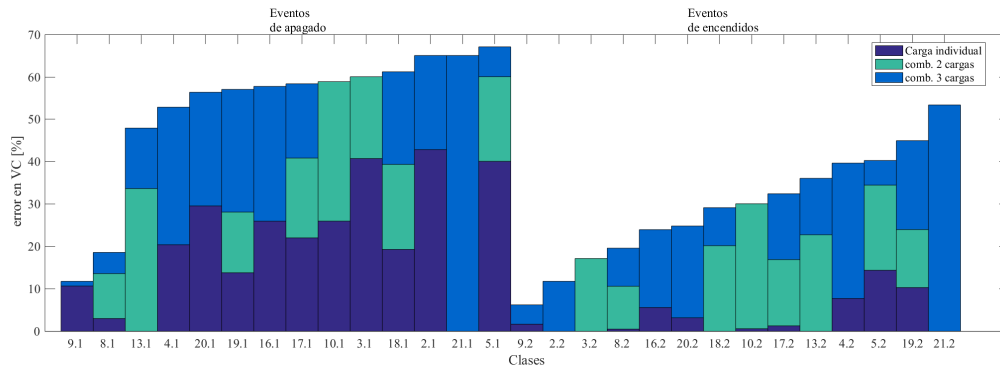
En los sistemas de identificación de eventos transitorios, es de gran interés analizar el rendimiento de los sistemas en la estimación de eventos en condiciones de combinación de carga. Este análisis se realiza a los sistemas con SVMs, los cuales logran los mejores resultados en identificación. En la figura 21 se presenta el error de los sistemas con SVMs en la identificación de eventos de los dispositivos en escenarios de operación simultánea con dos y tres cargas. En la tabla 11 se muestran los errores de estimación de los eventos individuales y eventos combinados con dos y tres cargas. Con los resultados obtenidos, se tiene que el error en identificación de los sistemas para eventos de cargas individuales es de alrededor del 25 %, pero en la identificación de eventos de combinaciones de carga se tienen errores elevados (al rededor del 60 % para grupos de dos cargas y alrededor del 50 % para grupos de tres cargas). Con esto, se evidencia la necesidad de un pre-procesamiento a los eventos transitorios para reducir el efecto cuando el dispositivo se encuentra combinado con otras cargas, lo cual se plantea como un trabajo a futuro.

Tabla 11.

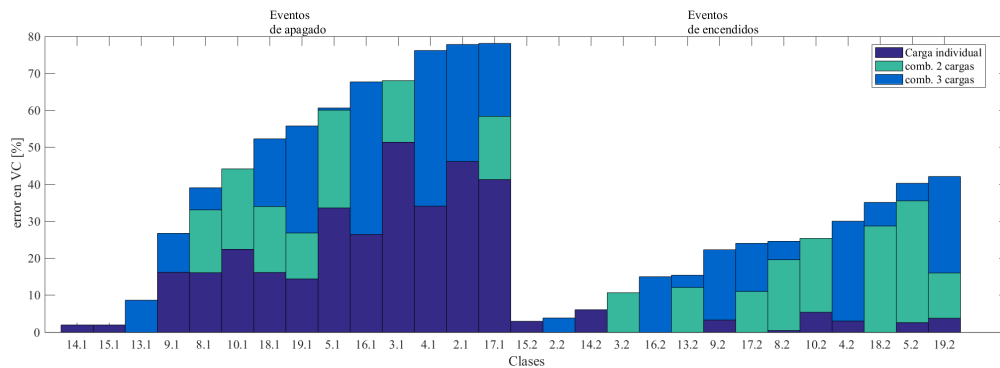
*Error en la identificación de los dispositivos con eventos de combinaciones de una (1), dos (2) y tres (3) cargas de los sistemas con los 3 métodos de extracción de características (FRFT, Wavelet, S) y estimación con SVM.*

<b>Método</b>	<b><math>e(\%)</math> de eventos en Carga Individual</b>	<b><math>e(\%)</math> de eventos en grupo de dos cargas</b>	<b><math>e(\%)</math> de eventos en grupo de tres cargas</b>
FRFT	22.8	60.1	47.6
Wavelet	25.6	54.4	46.7
S	24.0	63.5	50.1

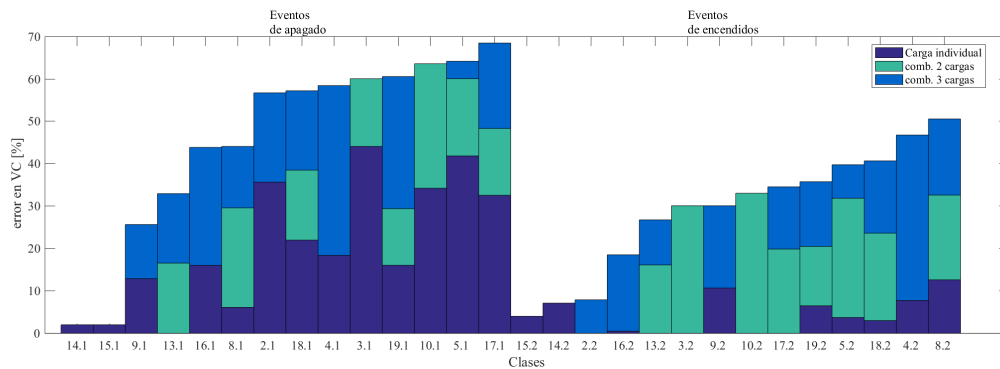
En la figura 22 se presenta el error en identificación de los sistemas cuando las cargas son asociadas a las tres categorías de dispositivos definidas (electrónica, no electrónica-resistivo, no



(a) Error en la identificación de dispositivos con eventos de combinaciones de cargas de los sistemas de identificación de régimen transitorio con estimación realizada con SVM y extracción de características basada en FRFT



(b) Error en la identificación de dispositivos con eventos de combinaciones de cargas de los sistemas de identificación de régimen transitorio con estimación realizada con SVM y extracción de características basada en Wavelet



(c) Error en la identificación de dispositivos con eventos de combinaciones de cargas de los sistemas de identificación de régimen transitorio con estimación realizada con SVM y extracción de características basada en transformada S

Figura 21. Error en la identificación de dispositivos con eventos de combinaciones de cargas de los sistemas de identificación de régimen transitorio con estimación realizada con SVM para los 3 métodos de extracción de características (FRFT, Wavelet, S).

electrónica-inductivo). Con los resultados obtenidos, los sistemas muestran un alta dificultad en la identificación de eventos de apagado (error de alrededor del 60 %). Esta tendencia se mantiene en eventos individuales o de combinación de cargas para todas las propuestas de extracción de características. Además, los sistemas de identificación logran sus errores mas bajos (alrededor del 15 %) en la estimación de eventos de dispositivos de categoría electrónica. Siendo en esta categoría, donde la propuesta con la transformada Wavelet logra su mayor ventaja sobre los demás sistemas. El sistema de identificación con extracción de características basada en la FRFT logra diferenciarse de los demas métodos en la identificación de eventos de apagado de dispositivos resistivos. En este caso, logra una diferencia de alrededor del 17 % sobre el sistema con transformada Wavelet y alrededor del 11 % del sistema con transformada S. Para las demás categorías, los sistemas presentan valores de error muy cercanos entre estas.

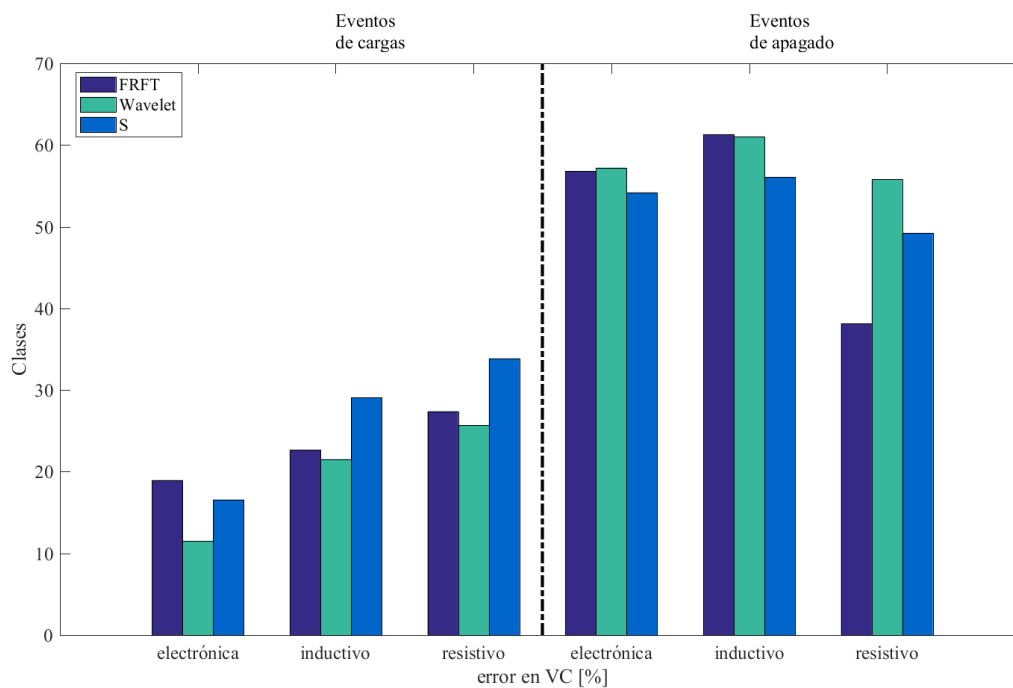


Figura 22. Error en la identificación de eventos transitorios por categoría de dispositivos con los 3 métodos de extracción de características (FRFT, Wavelet, S) y estimación con SVM

### 3.5 Comparación de los sistemas de identificación

La tabla 12 muestra los resultados globales en exactitud de identificación de los algoritmos implementados y su error estándar con un nivel de confianza de 0.95. Para comparar los algoritmos se utiliza la prueba  $t_{est}$  por pares, en la cual se asume la hipótesis nula que no existe una diferencia significativa en la exactitud de los sistemas de identificación. En el cálculo del estadístico, los sistemas se probaron con una base de datos única, y en este caso se asume cada partición de la validación cruzada como una base de datos independiente (Bramer, 2013). Las tablas 13 y 14 presentan los resultados de error en cada partición y la diferencia entre los sistemas implementados del régimen estacionario y transitorio con SVMs. Con esta información, se calcula el estadístico  $t_{est}$  utilizando la expresión 3.1. Para las propuestas con ANNs no se realiza prueba de hipótesis debido a que sus valores en exactitud son considerablemente más bajos en comparación con los de SVMs.

$$t_{est} = \frac{\sum_{i=1}^{n_t} z_{ti}}{n_t} \sqrt{\frac{(n_t - 1)n_t}{\sum_{i=1}^{n_t} z_{ti}^2 - (\sum_{i=1}^{n_t} z_{ti})^2/n_t}} \quad (3.1)$$

Tabla 12.

*Exactitud en identificación en validación cruzada de los sistemas de identificación con SVMs y ANNs para eventos del régimen estacionario y transitorio con las técnicas de extracción de características desarrolladas*

Predictor	Régimen transitorio			Régimen Estacionario		
	Wavelet	S	FRFT	Wavelet	S	FRFT
SVMs	64.58 ± 1.11	63.05 ± 1.14	65.08 ± 1.13	99.25 ± 0.15	92.19 ± 0.47	98.56 ± 0.20
ANNs	36.65 ± 1.14	44.89 ± 1.18	53.32 ± 1.18	62.52 ± 1.07	78.2 ± 0.72	78.2 ± 0.72

Donde  $n_t$  es el número de particiones y  $z_{ti}$  es la diferencia en exactitud de los sistemas en comparación para la partición  $i$ . Con un nivel de confianza de 95 % y 9 grados de libertad se rechaza

Tabla 13.

*Error en la estimación en cada partición del proceso de validación cruzada de los sistemas de identificación de eventos de régimen estacionario con estimación con SVMs. El valor de  $z a_t$  y  $z b_t$  corresponden a la diferencia entre el sistema de FRFT con Wavelet y S respectivamente.*

<b>partición</b>	<b>Wavelet</b>	<b>S</b>	<b>FRFT</b>	$z a_t = \text{FRFT}$ – Wavelet	$z a_t^2$	$z b_t = \text{FRFT}$ – S	$z b_t^2$
1	0.71	7.35	1.74	1.03	1.06	-5.61	31.47
2	0.71	8.06	1.50	0.79	0.62	-6.56	43.03
3	0.71	8.14	1.34	0.63	0.40	-6.80	46.24
4	0.40	7.51	1.34	0.94	0.88	-6.17	38.07
5	0.94	7.59	1.10	0.16	0.03	-6.49	42.12
6	0.87	8.06	1.81	0.94	0.88	-6.25	39.06
7	0.71	8.37	1.58	0.87	0.76	-6.79	46.10
8	0.79	8.53	1.34	0.55	0.30	-7.19	51.70
9	0.87	7.35	1.18	0.31	0.09	-6.17	38.07
10	0.80	7.12	1.42	0.62	0.38	-5.70	32.49

Tabla 14.

*Error en la estimación en cada partición del proceso de validación cruzada de los sistemas de identificación de eventos de régimen transitorio con estimación con SVMs. El valor de  $z a_t$  y  $z b_t$  corresponden a la diferencia entre el sistema de FRFT con Wavelet y S respectivamente.*

<b>partición</b>	<b>Wavelet</b>	<b>S</b>	<b>FRFT</b>	$z a_t = \text{FRFT}$ – Wavelet	$z a_t^2$	$z b_t = \text{FRFT}$ – S	$z b_t^2$
1	34.50	33.33	33.77	-0.73	0.52	0.44	0.19
2	36.48	35.90	36.05	-0.43	0.18	0.15	0.02
3	34.45	37.65	35.76	1.31	1.69	-1.89	3.57
4	34.30	39.39	37.94	3.64	13.18	-1.45	2.10
5	35.03	37.94	37.35	2.32	5.38	-0.59	0.35
6	36.77	36.34	35.47	-1.30	1.69	-0.87	0.76
7	37.70	36.97	34.64	-3.06	9.36	-2.33	5.43
8	34.35	35.66	30.42	-3.93	15.44	-5.24	27.45
9	35.52	38.86	32.17	-3.35	11.16	-6.69	44.76
10	35.08	37.41	35.66	0.58	0.34	-1.75	3.06

la hipótesis nula si  $|t_{est}| > 2.262$ . En la tabla 15 se muestran los valores del estadístico  $t_{est}$  cuando se compara el sistema de FRFT con el de Wavelet y transformada S para el régimen estacionario y transitorio. En el régimen transitorio existe una diferencia significativa a favor del sistema con FRFT sobre el sistema con transformada S, y en la comparación entre FRFT y Wavelet se concluye con la prueba que no existe una diferencia entre estos dos. En el régimen estacionario existe una diferencia significativa en los resultados de identificación a favor del sistema con características basadas en transformada Wavelet y el de FRFT logra una ventaja sobre el sistema con transformada S.

Tabla 15.

*Valor de estadístico  $|t_{est}|$  para la prueba de hipótesis en la comparación de los sistemas de régimen estacionario y transitorio con SVMs.*

	<b>Régimen transitorio</b>		<b>Régimen estacionario</b>	
	Wavelet	S	Wavelet	S
FRFT	0.62	2.80	7.59	40.63

La comparación de velocidad de cómputo de cada algoritmo se realiza en un computador con procesador intel(R) Core(TM) i7-5500u CPU @ 2.40GHz, memoria ram de 12GB y sistema operativo Windows 10 Home. En las tablas 16 y 17 se muestran los tiempos de cómputo estimados en el proceso de extracción de características y estimación de los algoritmos de eventos transitorios y estacionarios respectivamente. En los sistemas con SVMs los operadores de extracción de características con FRFT logran el tiempo de cómputo más veloz en comparación con las demás propuestas, siendo el sistema con Wavelet el de menor desempeño. Cuando la estimación se realiza con ANNs la propuesta con transformada S logra el tiempo de cómputo más veloz seguido del algoritmo con FRFT. En general el sistema con SVMs y extracción de características utilizando la FRFT obtiene el mejor desempeño en tiempo de cómputo tanto para el régimen transitorio y

estacionario.

Tabla 16.

*Tiempo de cómputo en segundos de la estimación de un evento en los sistema de identificación de eventos transitorios con tamaño de evento  $M = 153965$ .*

Proceso	SVMs			ANNs		
	FRFT	Wavelet	S	FRFT	Wavelet	S
Extracción de características (s)	0.8745	40.0863	1.9057	2.7872	40.5254	1.9243
Estimación (s)	0.2164	0.2731	0.2255	0.3588	0.3505	0.3744
<b>Total (s)</b>	<b>1.0909</b>	<b>40.3594</b>	<b>2.1312</b>	<b>3.1410</b>	<b>40.8759</b>	<b>2.2987</b>

Tabla 17.

*Tiempo de cómputo en segundos de la estimación de un evento en los sistema de identificación de eventos estacionarios con tamaño de evento  $M = 2500$ .*

Proceso	SVMs			ANNs		
	FRFT	Wavelet	S	FRFT	Wavelet	S
Extracción de características (s)	0.0659	0.3249	0.0628	0.1781	0.3322	0.0613
Estimación (s)	0.2322	0.2556	0.2413	0.3674	0.3474	0.3545
<b>Total (S)</b>	<b>0.2981</b>	<b>0.5805</b>	<b>0.3041</b>	<b>0.5455</b>	<b>0.6796</b>	<b>0.4158</b>

En los resultados de exactitud por clase de los sistemas de eventos transitorios se observa que todos los sistemas implementados presentan errores elevados en la identificación de eventos de apagado. Considerando esto, y el hecho que los algoritmos con SVMs tienen los mejores resultados en identificación, se diseñaron sistemas independientes para los eventos de apagado y los demás eventos utilizando modelos de SVMs. La tabla 18 presenta los resultados para cada propuesta de extracción de características, donde el sistema A es el de eventos de apagado y el sistema B es para los demás eventos. Con estos resultados, se concluye que se puede mejorar el rendimiento en identificación al implementar sistemas independientes para los encendidos y apagados de los dispositivos.

Tabla 18.

*Error de identificación (%) en validación cruzada de los sistema A y B de eventos de régimen transitorio con estimación realizada con SVMs. El sistema A corresponde a los eventos transitorios de apagado de los dispositivos y el sistema B a los demás eventos*

Características	Regimen transitorio		
	sistema A	sistema B	total
FRFT	42.99	20.74	31.43
Wavelet	45.81	16.21	30.43
S	44.63	25.84	34.87

Con los resultados obtenidos, se tiene que el modelo de estimación con la mayor exactitud en la identificación es el de SVMs para las tres propuestas de extracción de características comparadas. Con este modelo, la exactitud en identificación para eventos de régimen estacionario es superior al 92 %, llegando a alcanzar un valor de alrededor del 98 % con los sistemas de Wavelet y FRFT. En los sistemas de eventos de régimen transitorio se tienen exactitudes alrededor del 63 % para las tres técnicas comparadas, donde el mayor número de eventos mal clasificados se agrupan en los eventos de apagado de los dispositivos. Lo anterior permite pensar en la posibilidad de desarrollar sistemas independientes para la identificación de eventos de encendido y apagados y optimizar el rendimiento de estos sistemas por separado. La prueba ilustrada en la tabla 18, muestra que se puede mejorar el rendimiento del sistema de esta manera. Por otra parte, cuando la estimación se realiza con los modelos de ANNs se observa que el modelo se ajusta a las clases con mayor número de eventos, pero obtiene bajos resultados de identificación en las demás clases. Esto sugiere que un modelo más complejo y un mayor volumen de datos es necesario para la estimación del gran número de clases en el problema de identificación de cargas eléctricas con ANNs. Finalmente, las conclusiones principales de la investigación desarrollada se presentan en el siguiente capítulo.

#### 4. Conclusiones

Los desarrollos contenidos en la presente investigación de maestría describen el diseño de sistemas de identificación de cargas con extracción de características basada en la FRFT y las comparaciones realizadas con enfoques basados en las transformadas Wavelet y S. Se implementan sistemas independientes para los eventos transitorios y estacionarios de los dispositivos con cada propuesta de extracción de características y la predicción se realiza con algoritmos de SVMs y ANNs. Los resultados en identificación muestran que los algoritmos con SVMs logran exactitudes más elevadas que los sistemas con ANNs en todos los escenarios contemplados. Cuando la estimación se realiza con SVMs, se tiene que los sistemas de Wavelet y FRFT logran exactitudes de alrededor del 98 % en la identificación de eventos de régimen estacionario. En el caso de los sistemas de eventos de régimen transitorio, se tiene una exactitud de alrededor del 63 % para las tres técnicas de extracción de características comparadas. En el caso de los sistemas con estimación realizada con ANNs, se observa que los modelos se ajustan a las clases con mayor número de eventos, lo que puede indicar que las topologías planteadas son muy simples para el problema de identificación de cargas.

Con las comparaciones realizadas, se tiene que el rendimiento en identificación de los sistemas basados en Wavelet y FRFT son muy similares logrando una ventaja sobre el sistema basado en la transformada S. En velocidad de procesamiento, el sistema de FRFT obtiene el mejor rendimiento seguido del sistema con transformada S, y la extracción de características con transformada Wavelet tiene el costo computacional más elevado de las propuestas desarrolladas. Considerando lo

anterior, el sistema con extracción de características basado en la FRFT muestra el mejor balance entre exactitud y velocidad entre las alternativas estudiadas.

A continuación se enuncian los principales aportes contenidos en cada capítulo del documento: En el capítulo 2 se presenta el problema de desagregación de energía y una posible solución con la implementación de sistemas de identificación con algoritmos de reconocimiento de patrones. También, se describe el proceso general de identificación y los aspectos a considerar en el diseño de los sistemas, además, se destacó que el rendimiento del sistema es directamente dependiente del operador de extracción de características, modelo de estimación y de los datos utilizados en el entrenamiento y validación. Considerando esto, se presenta la revisión literaria de algunos sistemas de identificación de cargas, haciendo énfasis en los conjuntos de características extraídas y en los modelos utilizados para la estimación en cada sistema. Se observa que la mayoría de los sistemas revisados utilizan algoritmos de ANNs, SVMs y vecinos cercanos para la predicción, y la extracción de características es realizada en base a las propiedades transitorias o estacionarias de las señales eléctricas. Con esta tendencia se propone el desarrollo y diseño de sistemas independientes para eventos transitorios y estacionarios con modelos de estimación de ANNs y SVMs.

En el capítulo 3 se describe la metodología de diseño propuesta, la cual incluye el proceso de extracción de características, selección de parámetros y conjuntos finales de características, así como el diseño de los modelos de estimación. En la primera parte, se realiza el planteamiento de los operadores de extracción de características, su selección de parámetros y las definiciones de la FRFT utilizadas. Luego, se describe el método recursivo de selección de características finales y

el proceso de diseño de los modelos de SVMs y ANNs. Finalmente, se presentan los resultados obtenidos en cada etapa de diseño y los sistemas finales de identificación de cargas con los conjuntos de características. En el capítulo 4 se exponen los resultados finales en identificación por clase y global de los sistemas propuestos, así como de los sistemas con extracción de características basada en las transformadas Wavelet y S. Además, se comparan los sistemas desarrollados con pruebas de hipótesis donde se concluye que en los sistemas de régimen estacionario la extracción de características con la transformada Wavelet presenta una diferencia significativa sobre el sistema con FRFT y éste último presenta mejores resultados que el sistema de transformada S. En el régimen transitorio no existe diferencia entre los sistemas de FRFT y Wavelet pero si hubo una diferencia entre los sistemas de FRFT y S a favor del primero. En cómputo, el sistema con SVMs y extracción de características realizada con la FRFT logra el menor tiempo para los sistemas de eventos de régimen estacionario y transitorio.

### **Trabajo futuro**

El problema de investigación abordado en el presente trabajo, posibilita una serie de futuras actividades destinadas a la implementación de un sistema de monitorización de cargas eléctricas residenciales, algunas de las actividades son:

A nivel del grupo de investigación

- Prototipado de los sistemas de identificación.
- Desarrollo de aplicaciones web y móvil para el manejo de los datos e información.
- Manejo de datos en servidores dedicados, así como el desarrollo de un sistema automático

de mediciones.

- Selección del conjunto óptimo de características considerando todas las propuestas planteadas en el grupo de investigación.
- Exploración de sistemas de procesamiento de señales que permitan extraer la información de los eventos transitorios de los dispositivos mitigando el efecto de las cargas activas en el instante del evento.
- Exploración de metodologías para disminuir el número de clases en los sistemas de régimen estacionario. Esto puede incluir estudios estadístico y probabilístico de las combinaciones de cargas más comunes en la actividad diaria de la comunidad regional.

A nivel general

- Normatividad para estandarizar las bases de datos y una integración global o regional de los datos adquiridos. Esto con el fin de desarrollar sistemas de identificación bajo los mismos escenarios y facilitar la comparación de los sistemas.
- Con el acceso a bases de datos de tamaño “big data” se pueden explorar algoritmos de estimación con técnicas de aprendizaje profundo. Considerando los avances en reconocimiento de voz y facial con este tipo de modelos de estimación se puede intuir una mejoría en el rendimiento con esta alternativa en la identificación de cargas.
- Una comparación más amplia de los conjuntos de características propuestos en la literatura, para identificar el mejor conjunto y facilitar el contraste de futuras propuestas de características.

- Operadores de extracción de características utilizando las propiedades de filtrado tiempo-frecuencia de la FRFT.

### **Divulgación de resultados**

Se realizó la siguiente ponencia en conferencia internacional:

H. Cala; Y. Jimenez; R. Torres; Cesar Duarte. Efecto de una distorsión de onda achatada sobre un sistema de identificación de cargas basado en características extraídas a través de la transformada de Fourier fraccionaria. IX Simposio Internacional sobre Calidad de la Energía Eléctrica SICEL 2017. Bucaramanga, Colombia, November 3-7, 2017.

### Referencias Bibliográficas

- Abubakar, I., Khalid, S., Mustafa, M., Shareef, H., y Mustapha, M. (2015). An overview of non-intrusive load monitoring methodologies. En *Energy conversion (cencon), 2015 IEEE conference on* (pp. 54–59).
- Almeida, L. B. (1994). The fractional fourier transform and time-frequency representations. *IEEE Transactions on signal processing*, 42(11), 3084–3091.
- Amarís, R. Á. T. (2004). *Holografía numérica por transformación de fourier fraccional* (Tesis de Maestría). Universidad Industrial de Santander.
- Amarís, R. T. (2008). *Tratamiento de senales por transformación de fourier fraccionaria. aplicaciones a la holografía sintética y al filtrado óptimo* (Tesis de doctorado). Universidad Industrial de Santander.
- Berges, M. E., Goldman, E., Matthews, H. S., y Soibelman, L. (2010). Enhancing electricity audits in residential buildings with nonintrusive load monitoring. *Journal of industrial ecology*, 14(5), 844–858.
- Bhattacharjee, S., Kumar, A., y RoyChowdhury, J. (2014). Appliance classification using energy disaggregation in smart homes. En *Computation of power, energy, information and communication (iccpeic), 2014 international conference on* (pp. 1–6).
- Bianson, S., y Plangklang, B. (2016). Nonintrusive load monitoring (nilm) using an artificial neural network in embedded system with low sampling rate. En *Electrical engineering/electronics, computer, telecommunications and information technology (ecti-con), 2016*

*13th international conference on* (pp. 1–4).

Bramer, M. (2013). *Principles of data mining*. Springer Science & Business Media.

Chahine, K., Drissi, K. E. K., Pasquier, C., Kerroum, K., Faure, C., Jouannet, T., y Michou, M. (2011). Electric load disaggregation in smart metering using a novel feature extraction method and supervised classification. *Energy Procedia*, 6, 627–632.

Chang, H.-H. (2012). Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses. *Energies*, 5(11), 4569–4589.

Chang, H.-H., Chien, P.-C., Lin, L.-S., y Chen, N. (2011). Feature extraction of non-intrusive load-monitoring system using genetic algorithm in smart meters. En *e-business engineering (icebe)*, 2011 *IEEE 8th international conference on* (pp. 299–304).

Chang, H.-H., Kuo-Lung, L., Su, Y.-C., y Lee, W.-J. (2014). Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification. *IEEE Transactions on Industry Applications*, 50(3), 2081–2089.

Chang, H.-H., Lian, K.-L., Su, Y.-C., y Lee, W.-J. (2013). Energy spectrum-based wavelet transform for non-intrusive demand monitoring and load identification. En *Industry applications society annual meeting, 2013 IEEE* (pp. 1–9).

Christopher, M. B. (2016). *Pattern recognition and machine learning*. Springer-Verlag New York.

Darby, S., y cols. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, 486.

Du, L., Yang, Y., He, D., Harley, R. G., Habetler, T. G., y Lu, B. (2012). Support vector machine based methods for non-intrusive identification of miscellaneous electric loads. En *Iecon 2012-38th annual conference on IEEE industrial electronics society* (pp. 4866–4871).

- Duarte, C., Delmar, P., Goossen, K. W., Barner, K., y Gomez-Luna, E. (2012). Non-intrusive load monitoring based on switching voltage transients and wavelet transforms. En *Future of instrumentation international workshop (fiw), 2012* (pp. 1–4).
- Duda, R. O., Hart, P. E., y Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.
- Ehrhardt-Martinez, K., Donnelly, K. A., Laitner, S., y cols. (2010). Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities..
- Figueiredo, M., De Almeida, A., y Ribeiro, B. (2012). Home electrical signal disaggregation for non-intrusive load monitoring (nilm) systems. *Neurocomputing*, 96, 66–73.
- Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy efficiency*, 1(1), 79–104.
- Gillis, J. M., y Morsi, W. G. (2017). Non-intrusive load monitoring using semi-supervised machine learning and wavelet design. *IEEE Transactions on Smart Grid*.
- Hagan, M. T., Demuth, H. B., Beale, M. H., y cols. (1996). *Neural network design* (Vol. 20). Pws Pub. Boston.
- Hamid, O., Barbarosou, M., Papageorgas, P., Prekas, K., y Salame, C. (2017). Automatic recognition of electric loads analyzing the characteristic parameters of the consumed electric power through a non-intrusive monitoring methodology. *Energy Procedia*, 119, 742–751.
- Jiang, L., Luo, S., y Li, J. (2012). An approach of household power appliance monitoring based on machine learning. En *Intelligent computation technology and automation (icicta), 2012 fifth international conference on* (pp. 577–580).
- Jimenez, Y., Duarte, C., Petit, J., y Carrillo, G. (2014). Feature extraction for nonintrusive load mo-

- monitoring based on s-transform. En *Power systems conference (psc), 2014 clemson university* (pp. 1–5).
- Jimenez, Y., Duarte, C., Petit, J., Meyer, J., Schegner, P., y Carrillo, G. (s.f.). Characterization of current switching transients for appliance identification.
- Lin, Y.-H., y Tsai, M.-S. (2011). Applications of hierarchical support vector machines for identifying load operation in nonintrusive load monitoring systems. En *Intelligent control and automation (wcica), 2011 9th world congress on* (pp. 688–693).
- McBride, A., y Kerr, F. (1987). On namias's fractional fourier transforms. *IMA Journal of applied mathematics*, 39(2), 159–175.
- Ozaktas, H. M., Arikan, O., Kutay, M. A., y Bozdagt, G. (1996). Digital computation of the fractional fourier transform. *IEEE Transactions on signal processing*, 44(9), 2141–2150.
- Ozaktas, H. M., Kutay, M. A., y Mendlovic, D. (1999). Introduction to the fractional fourier transform and its applications. En *Advances in imaging and electron physics* (Vol. 106, pp. 239–291). Elsevier.
- Sadeghianpourhamami, N., Ruyssinck, J., Deschrijver, D., Dhaene, T., y Develder, C. (2017). Comprehensive feature selection for appliance classification in nilm. *Energy and Buildings*, 151, 98–106.
- Srinivasan, D., Ng, W., y Liew, A. (2006). Neural-network-based signature recognition for harmonic source identification. *IEEE Transactions on Power Delivery*, 21(1), 398–405.
- Stockwell, R. G., Mansinha, L., y Lowe, R. (1996). Localization of the complex spectrum: the s transform. *IEEE transactions on signal processing*, 44(4), 998–1001.
- Su, Y.-C., Lian, K.-L., y Chang, H.-H. (2011). Feature selection of non-intrusive load monitoring

- system using stft and wavelet transform. En *e-business engineering (icebe), 2011 ieee 8th international conference on* (pp. 293–298).
- Tabatabaei, S. M., Dick, S., y Xu, W. (2017). Toward non-intrusive load monitoring via multi-label classification. *IEEE Transactions on Smart Grid*, 8(1), 26–40.
- Tsai, M., y Lin, Y. (2011). Development of a non-intrusive monitoring technique for appliance identification in electricity energy management. En *Advanced power system automation and protection (apap), 2011 international conference on* (Vol. 1, pp. 108–113).
- Wang, T., y Yin, B. (2017). A new method for the nonintrusive load monitoring based on bp neural network. En *2017 2nd international conference on multimedia and image processing (icmip)* (pp. 93–97).
- Yang, H.-T., Chang, H.-H., y Lin, C.-L. (2007). Design a neural network for features selection in non-intrusive monitoring of industrial electrical loads. En *Computer supported cooperative work in design, 2007. cscwd 2007. 11th international conference on* (pp. 1022–1027).
- Zoha, A., Gluhak, A., Imran, M. A., y Rajasegarar, S. (2012). Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12), 16838–16866.

## Apéndices

### A. Especificaciones de los instrumentos utilizados en la adquisición de la base de datos

En la presente sección se anexan las especificaciones de la fuente programable CHROMA 61511, los módulos de adquisición de datos NI9225, NI9227 y NI9239 de *national Instrument* y las pinzas amperimétricas Fluke i30s.

Tabla 19.

*Especificaciones del módulo de adquisición de datos NI9225*

<b>Especificaciones</b>	<b>Módulo NI9225</b>
Número de canales	3 canales de entrada analógica
resolución ADC	24bits
tipo de ADC	Delta-Sigma (con un prefiltrado analógico)
Modo de muestreo	simultáneo
<b>Rango de datos utilizando el reloj interno</b>	
mínimo	1.613 kS/s
Máximo	50 kS/s
<b>Voltajes de operación</b>	
típico	300Vrms
Impedancia en la entrada	1M $\Omega$
Ruido de entrada	70 $\mu$ Vrms

Tabla 20.

*Especificaciones del módulo de adquisición de datos NI9227*

<b>Especificaciones</b>	<b>Módulo NI9227</b>
Número de canales	4 canales de entrada analógica
resolución ADC	24bits
tipo de ADC	Delta-Sigma (con un prefiltrado analógico)
Modo de muestreo	simultáneo
<b>Rango de datos utilizando el reloj interno</b>	
mínimo	1.613 kS/s
Máximo	50 kS/s
Rango de operación seguro	5Arms
Ruido de entrada	400 $\mu$ Arms

Tabla 21.

*Especificaciones del módulo de adquisición de datos NI9239*

<b>Especificaciones</b>	<b>Módulo NI9239</b>
Número de canales	4 canales de entrada analógica
resolución ADC	24bits
tipo de ADC	Delta-Sigma (con un prefiltrado analógico)
Modo de muestreo	simultáneo
<b>Rango de datos utilizando el reloj interno</b>	
mínimo	1.613 kS/s
Máximo	50 kS/s
<b>Voltajes de operación</b>	
Nominal	$\pm 10v$
Impedancia en la entrada	1M $\Omega$
Ruido de entrada	2mVrms
Distorsión armónica total (THD)	-95dB

Tabla 22.

*Especificaciones de la fuente programable CHROMA 61511*

<b>Especificaciones</b>	<b>Fuente CHROMA 61511</b>
<b>Fases en la salida</b>	1 o 3 seleccionable
<b>Potencia de salida-AC</b>	
Potencia	12kVA
Cada Fase	4kVA
<b>Voltaje</b>	
Rango	0 150V/0 300V
Exactitud	0.2 % + 0.2 % <i>F.S.</i>
Resolución	0.1 V
Distorsión	0.3 % @ 50/60 Hz, 1 % @ 15 – 1 kHz, 1.5 % @ > 1 kHz
Regulación de línea	0.1 %
Regulación de carga	0.2 %
<b>Corriente máxima (1-fase modo)</b>	
RMS	96A / 48A
Pico (CF=4)	384A / 192A
<b>Corriente máxima (Cada fase en el modo de 3 fases)</b>	
RMS	32A / 16A
Pico	(CF=4) 128A / 64A
<b>Frecuencia</b>	
Rango	DC, 15-1.5kHz
Exactitud	0.15 %
<b>fase</b>	
Rango	0 - 360°
Resolución	0.3°
Exactitud	< 0.8° @ 50/60 Hz
<b>Salida de DC (1-fase modo)</b>	
Potencia	6kW
Voltaje	212V / 424V
Corriente	48A / 24A
<b>Salida de DC (3-fases modo)</b>	
Potencia	2kW
Voltaje	212V / 424V
Corriente	16A / 8A
Dimensiones (H x W x D)	1163 x 546 x 700 mm / 45.78 x 21.5 x 27.56 inch *5
Peso	229.4 kg / 505.29 lbs

Tabla 23.

*Especificaciones de la pinza amperimétrica Fluke i30s*

Todas las exactitudes definidas para $23^{\circ}C \pm 1^{\circ}C$	
Rango de corriente	20A ACrms o DC
Rango de medida	$\pm 30$ A
sensibilidad en la salida	100 mV/A
Exactitud(a $25^{\circ}C$ )	$\pm 1\%$ de lectura $\pm 2mA$
Resolución	$\pm 1$ mA
Ruido típico en la salida	$200 \mu$ Vrms
Impedancia de carga	$> 100$ kOhms $\leq 100$ pF
Rango de frecuencias	DC a 100 kHz (0.5 dB)
temperatura de operación	0 a $+50^{\circ}C$
Peso	290 g

### B. Algoritmos genéticos usados para resolver el problema de órdenes óptimos

La resolución de los problemas óptimos 2.12 y 2.16 se realiza con el método de algoritmos genéticos. En el método, un individuo es un posible candidato solución del problema óptimo dado, y una población es un conjunto de individuos. Con esto, el proceso del algoritmo consiste en generar una población inicial de soluciones aleatorias y evaluar la función objetivo en esta población. Luego, se seleccionan los individuos con mejores puntuaciones en base al criterio de optimización y se procede a generar una población futura que incluya estos individuos y sus mutaciones. Asimismo, este proceso se repite hasta un número de iteraciones definido por el usuario y la salida consiste en el mejor individuo de la población final. En la resolución planteada, se utilizó las mutaciones de reemplazo, la cual consiste en reemplazar un segmento consecutivo de posiciones del individuo por valores aleatorios. También se implementó la mutación de reemplazar un elemento de la solución por un valor aleatorio y de incorporar individuos nuevos totalmente aleatorios. La figura 23 muestra el diagrama de flujo de los algoritmos genéticos implementados.

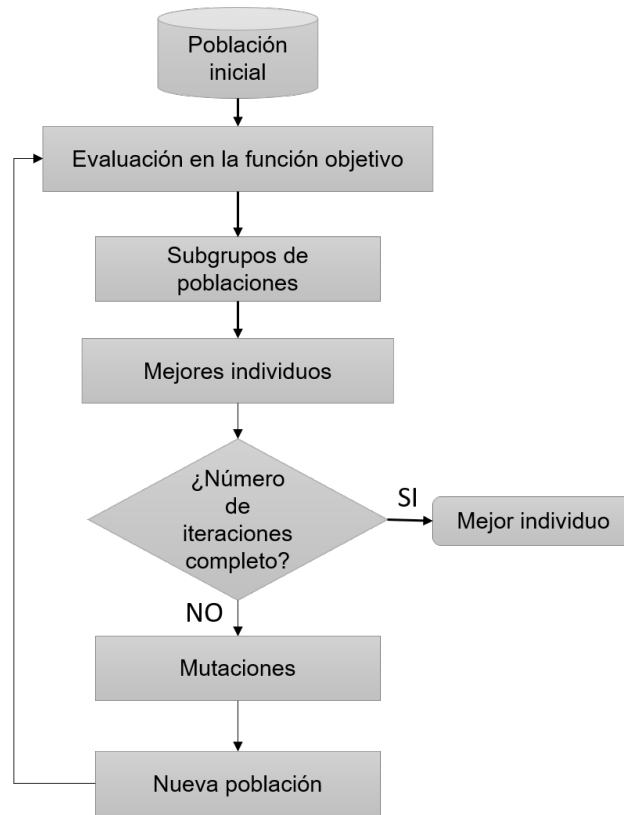


Figura 23. Diagrama de flujo de los algoritmos genéticos para la resolución de los problemas óptimos 2.12 y 2.16

### C. Extracción de eventos en las señales sin procesar y descripción de los eventos de régimen transitorio y secuencias de eventos para el computador de escritorio

Cuando se realiza la adquisición de datos de la operación de los dispositivos se obtienen señales como se ilustran en la figura 24. Estas señales contienen información de eventos de régimen transitorio y estacionario por lo que se realiza una extracción de eventos para cada caso de evento. Para el régimen transitorio, se buscan manualmente los puntos de inicio y fin de los transitorios para recortar la señal y almacenarla, en la etiqueta de la señal se especifica si el estado es de encendido o de apagado. La figura 25 muestra un ejemplo de extracción de evento transitorio de encendido

en la señal de corriente de la licuadora Oster 450w en velocidad 1 sin procesar. La señal dibujada en rojo se almacena y pasa a conformar un experimento en la base de datos de régimen transitorio. La base de datos de experimentos de régimen transitorio se presenta en la tabla 24, la cual esta conformada por eventos de los dispositivos en operación individual y/o en combinación con otras cargas. En el régimen estacionario, se extraen eventos de longitud de tres periodos, separados del punto final del evento transitorio como se muestra en la figura 26. Este evento en rojo se almacena y pasa a conformar un experimento en la base de datos del régimen estacionario.

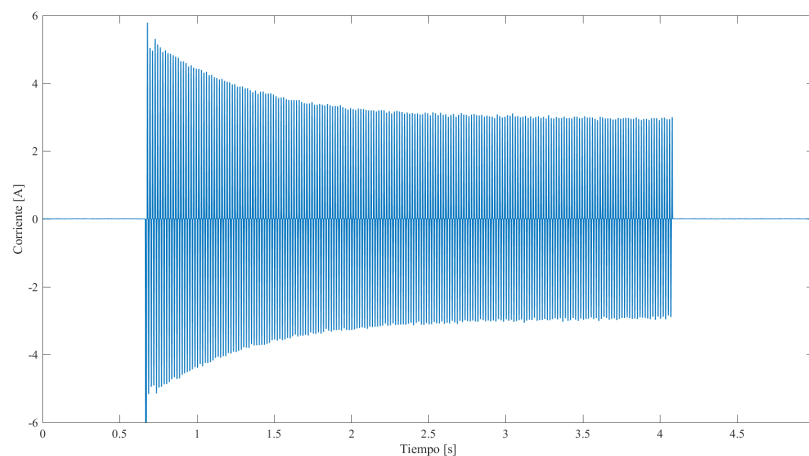


Figura 24. Adquisición de la señal de corriente del encendido y apagado de la licuadora Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz].

En la mayoría de dispositivos se realizaron experimentos de encendidos y apagados, pero en el caso del computador de escritorio (PC) DELL se desarrollaron secuencias de operación para la adquisición de los eventos. La secuencia de operación 1 consiste encender y apagar el monitor del PC cuando este se encuentra en modo de operación normal. La figura 27 presenta la forma de onda que se obtiene y los eventos (en rojo) que se extraen de la señal. La secuencia de operación 2 es la siguiente: se conecta el monitor, luego se conecta el PC, se enciende el monitor y se enciende el

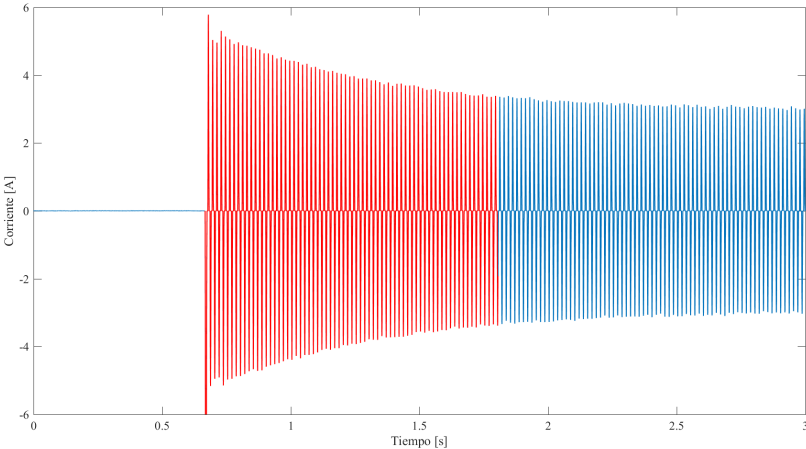


Figura 25. Evento de régimen transitorio de la señal de corriente del encendido de la licuadora Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz].

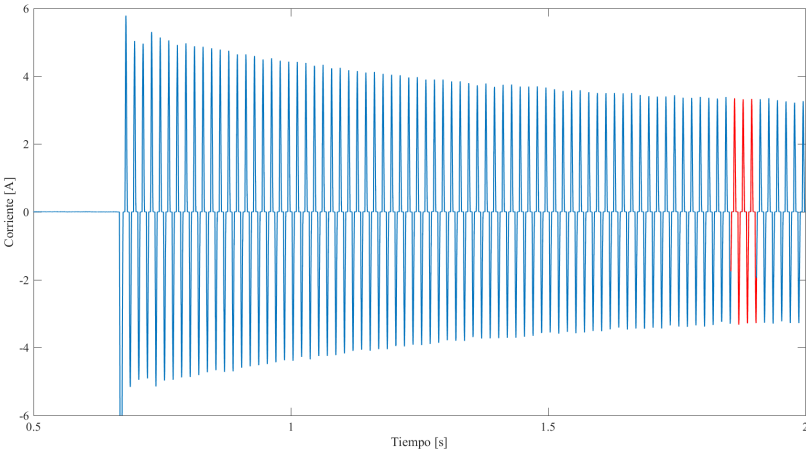


Figura 26. Evento de régimen estacionario de la señal de corriente del encendido de la licuadora Oster 450w en velocidad 1 con frecuencia de muestreo 50[kHz].

Tabla 24.

*Etiquetas asignadas para las combinaciones de uno, dos o tres dispositivos y número de señales registradas en estado transitorio por cada combinación. Las descripción de las etiquetas utilizadas para cada combinación se encuentra en la tabla 2.*

Dispositivo según tabla 2			Combinación de dos cargas				Combinación de tres cargas				
Etiqu.	Etiqu. tabla 2	Eventos	Etiqu.	Etiqu. tabla 2	Eventos	Etiqu.	Etiqu. tabla 2		Eventos		
1	1	198	20	17	10	30	48	19	16	20	120
2	2	200	21	17	18	40	49	4	19	17	120
3	3	200	22	17	19	40	50	4	17	9	120
4	4	194	23	17	13	40	51	20	16	8	120
5	5	200	24	10	18	30	52	17	4	2	120
6	6	200	25	10	19	40	53	19	20	18	120
7	7	200	26	10	13	30	54	19	20	13	120
8	8	200	27	18	19	40	55	16	4	21	120
9	9	200	28	18	13	40	56	19	18	9	120
10	10	200	29	19	13	40	57	16	2	8	120
11	11	200	30	17	3	30	58	4	9	13	120
12	12	200	31	17	8	40	59	19	5	21	120
13	14	400	32	17	5	29	60	19	18	13	120
14	15	200	33	10	3	30	61	16	2	21	120
15	16	202	34	10	8	20	62	2	9	13	120
16	17	216	35	10	5	30					
17	18	208	36	18	3	40					
18	19	208	37	18	8	40					
19	20	220	38	18	5	40					
			39	19	3	40					
			40	19	8	40					
			41	19	5	40					
			42	13	3	40					
			43	13	8	40					
			44	13	5	40					
			45	3	8	40					
			46	3	5	40					
			47	8	5	40					
<b>Total</b>		<b>4046</b>	<b>Total</b>		<b>1029</b>		<b>Total</b>			<b>1800</b>	

PC. Este proceso fue repetido en cada experimento y se ilustra en la figura 28. Para la secuencia de operación 3, el PC esta funcionando en estado normal y se procede a iniciar el software Microsoft Word y posteriormente se cierra el programa como se muestra en la figura 29.

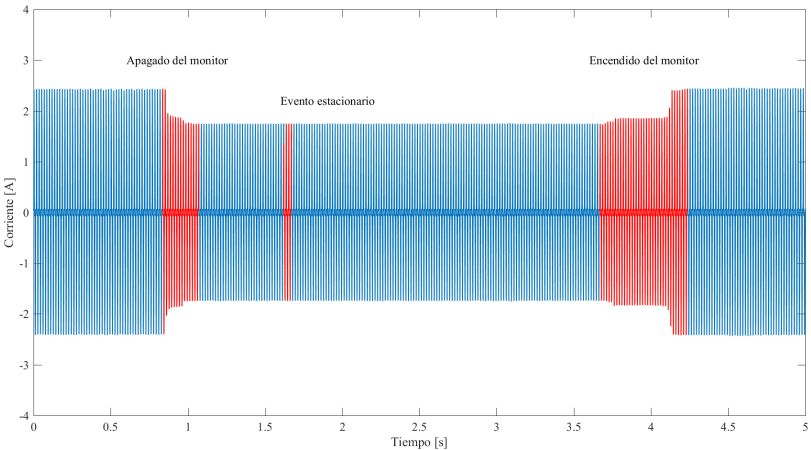


Figura 27. Eventos de régimen transitorio y estacionario de la secuencia 1 para el computador de escritorio Dell con frecuencia de muestreo 50[kHz].

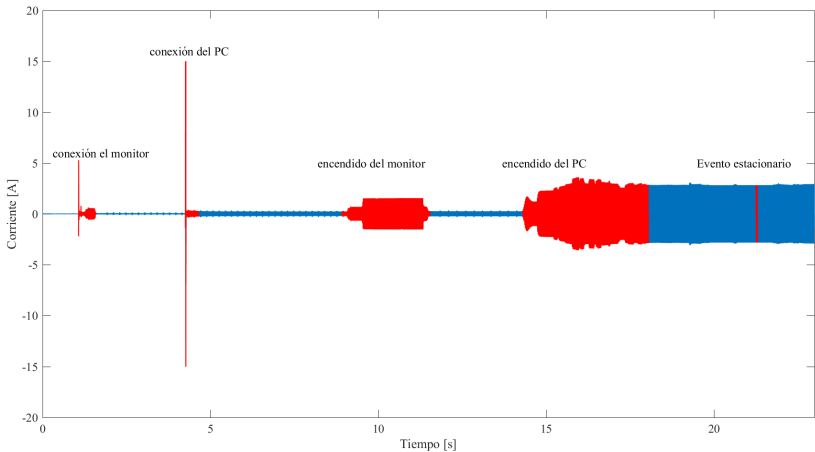


Figura 28. Eventos de régimen transitorio y estacionario de la secuencia 2 para el computador de escritorio Dell con frecuencia de muestreo 50[kHz].

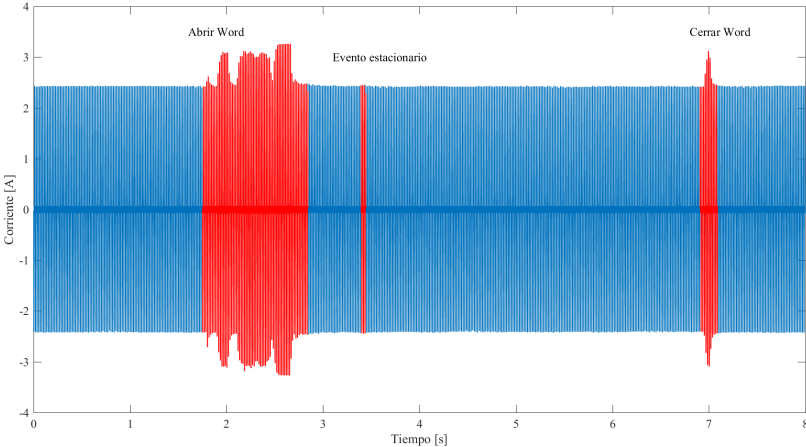


Figura 29. Eventos de régimen transitorio y estacionario de la secuencia 3 para el computador de escritorio Dell con frecuencia de muestreo 50[kHz].