

UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO



Facultad de Ingeniería Eléctrica

División de Estudios de Posgrado

**SIMULACIÓN Y ANÁLISIS DE LA INTERFERENCIA  
ELECTROMAGNÉTICA APLICADA A LA DESAGREGACIÓN DE CARGAS  
EN SISTEMAS DE MONITOREO NO INTRUSIVO DE CARGAS**

**TESIS**

Que para obtener el grado de

**MAESTRO EN CIENCIAS EN INGENIERÍA ELÉCTRICA**

Opción en Sistemas Computacionales

Presenta

**Julio Armando Raya Paniagua**

Doctor en Ingeniería Eléctrica y Electrónica

**Dr. Jaime Cerda Jacobo**

**Director de Tesis**

Morelia Michoacán Agosto 2015

# Lista de Publicaciones

Raya, Julio, and Jaime Cerda. "A low frequency power-based event alignment method for trained NILM systems." In Power, Electronics and Computing (ROPEC), 2014 IEEE International Autumn Meeting on, pp. 1-6. IEEE, 2014.

Valencia-Ramirez, Jose Maria, Julio Raya, Jose R. Cedeno, Ranyart R. Suarez, Hugo Jair Escalante, and Mario Graff. "Comparison between Genetic Programming and full model selection on classification problems." In Power, Electronics and Computing (ROPEC), 2014 IEEE International Autumn Meeting on, pp. 1-6. IEEE, 2014.



# Resumen

El uso de la electricidad es imprescindible en la sociedad actual. La mayoría de dispositivos que se utilizan en industrias, hogares, hospitales y transportes, entre otras, requieren de ésta para su correcto funcionamiento. Ésta dependencia hacia los dispositivos eléctricos incrementa la demanda energética en la mayoría de los países.

Recientemente se está trabajando en crear un sistema de monitoreo no intrusivo de cargas (*NILM*), que proporcione informes y estadísticas del consumo de energía eléctrica de cada dispositivo conectado a la red eléctrica en un hogar. A partir de un sistema que proporcione éste tipo de información se puede realizar una toma de decisiones, por ejemplo, que dispositivo eléctrico usar menos, en que horario es conveniente se encuentren funcionamiento, o hasta si es apropiado comprar un electrodoméstico novedoso con un menor consumo de energía que ahorraría dinero en las cuentas de energía a un mediano plazo.

El problema principal de los sistemas *NILM* es separar la potencia total consumida en un instante de tiempo, en la suma de potencias de cada dispositivo eléctrico conectado a la red. Éste trabajo implementa un esquema con un enfoque de mínimos cuadrados que pueda reportar cual o cuales dispositivos eléctricos se encuentran en funcionamiento en cada instante de tiempo, en base a la interferencia electromagnética que se producen los distintos dispositivos en un hogar y por consecuencia poder separar las cargas de los dispositivos.

La metodología propuesta se probó con varios conjuntos de datos generados artificialmente, donde se simuló capturar la interferencia electromagnética de 4 hogares distintos, además los datos fueron contaminados con ruido gaussiano para probar si el método propuesto es robusto o no.

La precisión de la metodología propuesta al clasificar con los distintos conjuntos de datos es en promedio del 94 %, donde conforme aumenta el ruido la precisión baja.

Un propósito de ésta investigación es resaltar la importancia de que las firmas electromagnéticas de los dispositivos sean proporcionadas de manera pública por los fabricantes, con el fin de que un sistema con la metodología propuesta pueda clasificar correctamente sin importar el lugar donde se implemente.

**Palabras Clave:** Sistemas *NILM*, Mínimos Cuadrados, Desagregación de Cargas, Interferencia Electromagnética, Clasificación



# Abstract

The use of electricity is essential in today's society. Most devices that are used in industries, homes, hospitals and transport, among others, require from it for the correct operation. This dependence on electrical devices increases the energy demand in most of the countries.

Recently researchers are working on creating a non-intrusive load monitoring system (*NILM*), which provides reports and statistics power consumption of each device connected to the grid in a home. From a system that provides this type of information can make a decision, for example, which electrical device using less, that schedule is convenient operation are, or even whether it is appropriate to buy a new appliance with lower consumption energy would save money on energy bills in the medium term.

The main problem of the *NILM* systems is to separate the total power consumed in a moment of time, the total power of each network connected to the electrical device. This work implements a scheme with a focus on least squares which can report or electrical devices which are in operation at each instant of time, based on electromagnetic interference that different devices are produced in a home and therefore able to separate loads devices.

The proposed methodology was tested with several datasets, where he pretended to capture electromagnetic interference generated 4 different homes, plus data were contaminated with Gaussian noise to test the proposed method is robust or not.

The accuracy of the proposed methodology to classify different data sets is on average 94 %, which with increasing precision low noise.

One purpose of this research is to highlight the importance of electromagnetic signatures of the devices are provided so published by the manufacturers, in order that a system with the proposed methodology to classify correctly no matter where it is implemented.

**Palabras Clave:** Sistemas *NILM*, Mínimos Cuadrados, Desagregación de Cargas, Interferencia Electromagnética, Clasificación



# Contenido

Lista de Publicaciones . . . . .	III
Resumen . . . . .	V
Abstract . . . . .	VII
Contenido . . . . .	IX
Lista de Figuras . . . . .	XI
Lista de Tablas . . . . .	XIII
Lista de Acrónimos . . . . .	XV
Lista de Símbolos . . . . .	XVII
1. Introducción . . . . .	1
1.1. Sistemas NILM . . . . .	3
1.2. Planteamiento del Problema . . . . .	7
1.2.1. Detección de Eventos . . . . .	7
1.2.2. Extracción de Características . . . . .	8
1.2.3. Clasificación . . . . .	9
1.3. Justificación y Motivación . . . . .	10
1.4. Objetivos y Contribución de la Tesis . . . . .	10
1.4.1. Objetivo General . . . . .	10
1.4.2. Objetivos Particulares . . . . .	10
1.4.3. Contribución de la Tesis . . . . .	11
1.5. Descripción de Capítulos . . . . .	11
2. Trabajo Relacionado . . . . .	13
2.1. Análisis en Estado Estable . . . . .	14
2.1.1. Cambios de Potencia . . . . .	14
2.1.2. Características del Voltaje y Corriente . . . . .	15
2.1.3. Trayectoria del Voltaje y Corriente . . . . .	18
2.1.4. Ruido en la señal de Voltaje . . . . .	19
2.2. Análisis en Estado Transitorio . . . . .	21
2.2.1. Voltaje y Corriente Transitoria . . . . .	21
2.2.2. Ruido en la señal de Voltaje Transitoria . . . . .	22
2.3. Resumen del Capítulo . . . . .	23

---

3.	Alineación de eventos basado en las componentes de baja frecuencia	27
3.1.	Descripción de los conjuntos de datos	28
3.2.	Metodología propuesta	29
3.2.1.	Suavizado con kernel gaussiano	30
3.3.	Resultados	32
3.4.	Resumen de Capítulo	35
4.	Desagregación de cargas basadas en proyecciones	37
4.1.	Creación del Conjunto de datos	38
4.2.	Desagregación	40
4.2.1.	Ejemplo Numérico	44
4.3.	Resumen de Capítulo	48
5.	Pruebas y Resultados	51
5.1.	Parámetros para generar los conjuntos de datos	51
5.2.	Pruebas	55
5.3.	Resultados	56
5.4.	Resumen de Capítulo	59
6.	Conclusiones y trabajo futuro	65
6.1.	Conclusiones Generales	66
6.2.	Trabajos Futuros	67
	Referencias	69

# Lista de Figuras

1.1.	(a) Monitoreo multipunto en red, cada carga cuenta con uno o mas sensores, estos entregan las mediciones al sistema <i>ILM</i> . (b) Monitoreo en un solo punto dentro de la red eléctrica, éste sensor entrega las muestras al sistema <i>NILM</i> .	3
1.2.	Diagrama de bloques de las etapas: detección de eventos, extracción de características y clasificación que se requieren en los sistemas <i>NILM</i> .	6
1.3.	Fragmento del conjunto de datos $H^T$ , donde una secadora de pelo es utilizada y se ejemplifica claramente la desalineación que existe en los eventos al realizar el registro de forma manual.	8
2.1.	Espacio bidimensional $\Delta P - \Delta Q$ utilizado para la desagregación de cargas.	15
2.2.	Variaciones en la potencia consumida a lo largo del tiempo al momento de conectar aparatos a la red eléctrica.	16
2.3.	(a) Formas de onda del voltaje y corriente en el dominio del tiempo. (b) Representación de la corriente en el dominio de la frecuencia.	17
2.4.	Ejemplo de la trayectoria Voltaje-Corriente.	19
2.5.	Fragmento del espectrograma de la interferencia electromagnética, donde dentro de éste fragmento las luces de la cochera son encendidas.	20
2.6.	Ejemplo de las perturbaciones (transitorios) que puede producir un dispositivo antes de llagar a estado estable.	22
3.1.	Efecto que tiene el suavizado mediante un kernel gaussiano. La primer señal de arriba hacia abajo es la señal pura sin suavizar y a las señales 2,3,4,5 y 6 se aplicó el suavizado con un kernel variando $\sigma = 1, 2, 3, 4, 5$ respectivamente.	31
3.2.	Etapas mas significativas de la alineación. En la parte superior es el fragmento de la potencia sin modificación, en el centro la potencia suavizada con un kernel gaussiano generado con $\sigma = 2$ y en la parte inferior ya procesada la señal que contiene picos donde hay variaciones en la potencia.	32
3.3.	Ejemplo de una falla en la alineación con dispositivos eléctricos que consumen potencia de manera variada a traves del tiempo. En la parte superior es el fragmento de la potencia variando en el tiempo, en el centro la potencia suavizada con un kernel gaussiano generado con $\sigma = 2$ y en la parte inferior la señal con picos donde existen cambios en la potencia.	34

---

4.1.	Cubo que forma el espacio de búsqueda cuando se cuenta con solo 3 dispositivos eléctricos. El espacio de búsqueda esta formado por las $2^3$ combinaciones que existen. . . . .	43
4.2.	Gráfica que ejemplifica donde se ubica la solución aproximada que arroja el método y con un redondeo se acerca al vértice mas cercano en el espacio de búsqueda. . . . .	44
4.3.	Ubicación de la solución en el espacio de busqueda. . . . .	47
5.1.	(a) Firma o huella electromagnética que produce las lámparas que se encuentran fuera de la cochera de la casa 1 (H1). (b) Firma o huella electromagnética que produce las luces de la cocina de la casa 2 (H2). . . . .	54
5.2.	(a) Firma o huella electromagnética que produce las luces de la entrada de la casa 3 (H3). (b) Firma o huella electromagnética que produce las luces del armario de la casa 4 (H4). . . . .	54
5.3.	Gráfica que ejemplifica el ruido generado para cada una de las componentes, entre mayor sea la componente mayor sera el ruido generado. . . . .	56
5.4.	Gráfica que ejemplifica el ruido generado tomando como desviación estándar la media de toda la muestra a contaminar. . . . .	57

# Lista de Tablas

2.1. Resumen de los métodos y características utilizados en el análisis en estado transitorio. . . . .	24
2.2. Resumen de los métodos y características utilizados en el análisis en estado estable. . . . .	25
3.1. Lista de parámetros utilizados . . . . .	33
3.2. Índice de desalineación por casa . . . . .	34
4.1. Firmas electromagnéticas de el refrigerador, televisor y lampara, así como el ruido electromagnético base que existe en el hogar cuando no se encuentra ningún dispositivo conectado. . . . .	45
4.2. Muestra que se producen al combinar las interferencias que producen los distintos dispositivos eléctricos. . . . .	45
4.3. Resultado de aplicar la metodología a las tres muestras de ejemplo. . . . .	48
5.1. Lista de dispositivos utilizados en el hogar 1 (H1). . . . .	52
5.2. Lista de dispositivos utilizados en el hogar 2 (H2). . . . .	52
5.3. Lista de dispositivos utilizados en el hogar 3 (H3). . . . .	53
5.4. Lista de dispositivos utilizados en el hogar 4 (H4). . . . .	53
5.5. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 1 (H1) contaminados con ruido gaussiano del tipo que denominamos <i>Proporcional</i> . . . . .	58
5.6. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 2 (H2) contaminados con ruido gaussiano del tipo que denominamos <i>Proporcional</i> . . . . .	59
5.7. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 3 (H3) contaminados con ruido gaussiano del tipo que denominamos <i>Proporcional</i> . . . . .	60
5.8. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 4 (H4) contaminados con ruido gaussiano del tipo que denominamos <i>Proporcional</i> . . . . .	60
5.9. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 1 (H1) contaminados con ruido gaussiano del tipo que denominamos <i>Promedio</i> . . . . .	61

5.10. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 2 (H2) contaminados con ruido gaussiano del tipo que denominamos <i>Promedio</i> . . . . .	61
5.11. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 3 (H3) contaminados con ruido gaussiano del tipo que denominamos <i>Promedio</i> . . . . .	62
5.12. Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 4 (H4) contaminados con ruido gaussiano del tipo que denominamos <i>Promedio</i> . . . . .	62

# Lista de Acrónimos

<i>SENER</i>	Secretaría de Energía.
<i>THw</i>	Teravatio-hora. <i>Terawatt-hour</i> .
<i>PIB</i>	Producto Interno Bruto.
<i>ALM</i>	Monitoreo de carga del dispositivo. <i>Appliance Load Monitoring</i> .
<i>ILM</i>	Monitoreo de carga intrusivo. <i>Intrusive Load Monitoring</i> .
<i>NILM</i>	Monitoreo de carga no intrusivo. <i>Non-Intrusive Load Monitoring</i> .
<i>FSM</i>	Maquina de estado finito. <i>Finite State Machine</i> .
<i>EMI</i>	Interferencia Electromagnética. <i>Electromagnetic Interference</i> .
<i>SMPS</i>	Fuente de alimentación conmutada. <i>Switch Mode Power Supplies</i> .
<i>SVM</i>	Máquina de Soporte Vectorial. <i>Support Vector Machine</i> .
<i>K-NN</i>	K Vecinos mas Cercanos. <i>K-Nearest Neighbors</i> .
<i>RECAP</i>	Perfilado y reconocimiento de dispositivos eléctricos en tiempo real. <i>RECognition of electrical Appliances and Profiling in real-time</i> .
<i>ANN</i>	Red Neuronal Artificial. <i>Artificial Neural Network</i> .
<i>RMS</i>	Raíz Media Cuadrada. <i>Root Mean Square</i> .
<i>CP</i>	Potencia Constante. <i>Constant Power</i> .
<i>CI</i>	Impedancia Constante. <i>Constant Impedance</i> .
<i>BP</i>	Propagación hacia atrás. <i>Back Propagation</i> .
<i>LVQ</i>	Red de Aprendizaje de Cuantificación Vectorial. <i>Learning Vector Quantization</i> .
<i>WTC</i>	Coefficientes de la Transformada Wavelet. <i>Wavelet Transformation Coefficients</i> .



# Lista de Símbolos

$d \in \mathbb{N}$	Número de dispositivos eléctricos.
$l \in \mathbb{N}$	Número de muestras en baja frecuencia (60Hz).
$h \in \mathbb{N}$	Número de muestras capturadas del espectrograma $H$ .
$P \in \mathbb{R}^l$	Potencia real consumida.
$p_i \in \mathbb{R}$	Potencia real consumida por el $i$ -ésimo dispositivo eléctrico.
$H \in \mathbb{R}^{n \times h}$	Espectrograma del ruido electromagnético generado en la red eléctrica.
$\Delta P \in \mathbb{R}$	Cambio en la potencia real cuando un dispositivo entra en funcionamiento.
$\Delta Q \in \mathbb{R}$	Cambio en la potencia reactiva cuando un dispositivo entra en funcionamiento.



## Capítulo 1

# Introducción

El uso de la electricidad es imprescindible en la sociedad actual. La mayoría de dispositivos que se utilizan en las industrias, hogares, hospitales y transportes, entre otras, requieren de ésta para su correcto funcionamiento. Existen un sinnúmero de aplicaciones donde se utiliza la electricidad como son las telecomunicaciones, iluminación, electrodomésticos, electrónica y robótica, por mencionar algunas.

Debido a la creciente dependencia de utilizar cada vez más dispositivos eléctricos, así como el crecimiento exponencial de la población, surge un problema en materia energética que es la creciente demanda de energía eléctrica para el consumo industrial y doméstico. Este es un problema que a la mayoría de países en el mundo afecta, sobre todo a los países que se encuentran en desarrollo como México.

En México la demanda de energía eléctrica crece relativamente rápido. De acuerdo a la prospectiva del sector eléctrico nacional 2010-2025 emitida por la Secretaría de Energía (SENER) [sener10], en el año 2000 el consumo fue alrededor de 160 *TWh* (*Terawatt-hora*), para el 2009 se incrementó a 206.3 *TWh* aproximadamente y para el año 2025, estimando un crecimiento promedio del producto interno bruto (*PIB*) del 3.5%, se pronostica un consumo de 404.7 *TWh*. El incremento pronosticado del consumo nacional de energía eléctrica representa un gran problema, ya que se debe generar casi el doble de energía eléctrica que la generada en el año 2009, por lo que se deben buscar alternativas para prevenir y satisfacer esta posible demanda.

Actualmente se está trabajando en incrementar la producción de energía eléctrica con nuevas fuentes de energía y que además contaminan menos al momento de su producción, como lo son las fuentes de energía renovables, ejemplos de ellas son la energía solar,

eólica, hidráulica, mareomotriz y geotérmica, entre otras. Lamentablemente, las energías renovables aún se encuentran en desarrollo y proporcionan poca cantidad de energía eléctrica, en proporción con la producción total que se realiza en el país. Aunque la cantidad de energía que aportan las fuentes de energía renovables ayudan en parte a solucionar el problema de la demanda energética, puede que no sea suficiente en un futuro.

Recientemente se está investigando la forma de crear un sistema que proporcione informes y estadísticas del consumo de energía eléctrica de cada dispositivo conectado a la red eléctrica en un hogar. A partir de un sistema que proporcione este tipo de información se puede realizar una toma de decisiones, por ejemplo, que dispositivo eléctrico usar menos, en que horario es conveniente se encuentren funcionamiento, o hasta si es apropiado comprar un electrodoméstico novedoso con un menor consumo de energía que ahorraría dinero en las cuentas de energía a un mediano plazo. Con este tipo de enfoque se pretende realizar un ahorro en el consumo de energía de los hogares e industrias y por consiguiente reducir sustancialmente el consumo pronosticado para los siguientes años.

Para que sea factible el ahorro en el consumo de energía mediante un sistema de información, existen métodos de monitoreo de carga de los dispositivos (*ALM*, del inglés *Appliance Load Monitoring*), donde lo que se pretende es medir de manera precisa el consumo eléctrico por cada dispositivo y producir reportes para la toma de decisiones.

Los dos principales enfoques de *ALM* son monitoreo de carga intrusivo (*ILM*, del inglés *Intrusive Load Monitoring*) y monitoreo de carga no intrusivo (*NILM* del inglés *Non-Intrusive Load Monitoring*) [Zoha12]. *ILM* realiza un monitoreo multipunto en red, donde cada aparato consta de uno o mas sensores que miden el momento de encendido y apagado del dispositivo, así como la potencia consumida, esto se muestra de manera gráfica en la Figura 1.1(a). A partir de estas mediciones directamente se pueden generar distintos reportes como los horarios de funcionamiento y potencia consumida de los distintos dispositivos. Aunque *ILM* resulta ser de gran precisión, es poco factible implementarlo debido al costo excesivo en equipo, difícil instalación y que requiere de un experto para su mantenimiento.

*NILM* a diferencia de *ILM* como se observa en la Figura 1.1(b), se basa en realizar el monitoreo en un solo punto de la red eléctrica, donde éste puede situarse en cualquier parte de la construcción a supervisar, incluso cerca del medidor que se encuentra fuera de la edificación. De esto se deriva parte de su nombre "*No Intrusivo*" ya que evita cualquier instalación dentro de la propiedad del consumidor.

La problemática principal de los sistemas *NILM* es que a partir de un solo punto

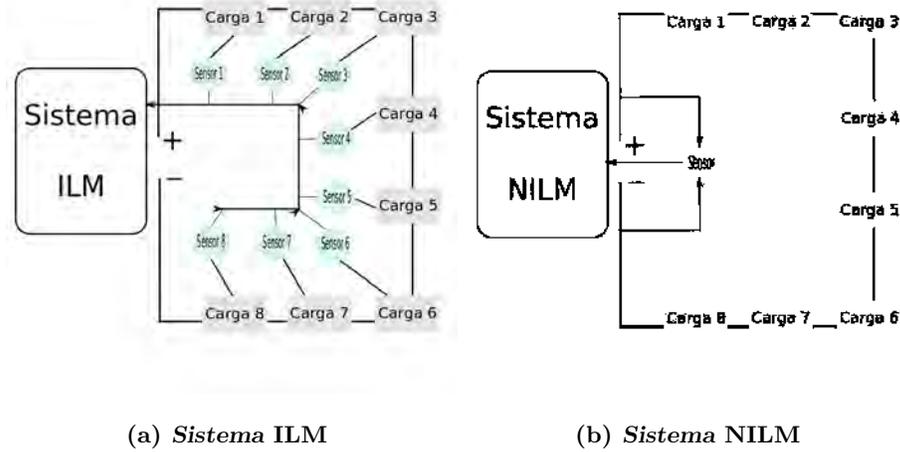


Figura 1.1: (a) Monitoreo multipunto en red, cada carga cuenta con uno o mas sensores, estos entregan las mediciones al sistema *ILM*. (b) Monitoreo en un solo punto dentro de la red eléctrica, éste sensor entrega las muestras al sistema *NILM*.

de medición se debe identificar que dispositivos eléctricos se encuentran en funcionamiento dentro de la red eléctrica, para separar las cargas y generar los reportes que nos auxilien en la reducción en el consumo eléctrico. Debido al bajo costo y fácil instalación con respecto al enfoque *ILM* varios investigadores han puesto su atención en solucionar éste problema para que sea posible su implementación. Los sistemas *NILM* actualmente siguen siendo un tema de investigación, ya que a la fecha no existe ninguna implementación que logre realizar la desagregación de cargas de los dispositivos eléctricos de manera precisa y pueda ser comercializada.

Éste trabajo se centra en uno de los problemas que existe en la implementación de los sistemas *NILM* para la reducción del consumo eléctrico en industrias y hogares, donde a partir de un conjunto de muestras capturadas en un solo punto dentro de la red eléctrica, se debe identificar que dispositivos eléctricos se encuentran en funcionamiento.

## 1.1. Sistemas NILM

El objetivo principal de los sistemas *NILM* es separar la potencia total consumida en un instante de tiempo, en la suma de potencias de cada dispositivo eléctrico conectado a la red, como se muestra en la Ecuación 1.1.

$$P(t) = \sum_{i=1}^d p_i(t) \quad (1.1)$$

Donde  $P(t)$  es la potencia total consumida en el instante de tiempo  $t$ ,  $p_i(t)$  es la potencia consumida por el  $i$ -ésimo dispositivo en el instante de tiempo  $t$  y  $d$  es el número de dispositivos conectados a la red.

A pesar de que el objetivo se puede plantear de forma sencilla, cumplir con él resulta una tarea complicada debido a la gran cantidad de dispositivos que se comercializan actualmente y su variado consumo de potencia, por lo que el dominio de búsqueda (tratándose de un problema de clasificación) crece exponencialmente conforme se crean nuevos aparatos eléctricos con el paso del tiempo.

Para simplificar un poco su estudio G.W. Hart [Hart92], propone clasificar en distintas categorías los dispositivos eléctricos de acuerdo a sus estados de operación, siendo estos:

- ON-OFF: Éste tipo de dispositivos cuenta solo con 2 estados de operación (encendido ó apagado), entre ellos se encuentran las lámparas y bombas de agua, por mencionar algunos.
- Máquina de estado finito (*FSM*, del inglés *Finite State Machine*): En esta categoría entran una gran cantidad de dispositivos que cuentan con varios modos de operación, como son los refrigeradores, lavadoras y microondas, entre otros.
- Continuamente variable: Son aquellos dispositivos donde no se puede calcular un número finito de estados de operación, como las lámparas con un regulador de intensidad, el taladro, etc.

Además de estas tres categorías Zaifman y Roth [Zeifman11] y Baranski et al. [Baranski03] plantean una cuarta para los dispositivos que siempre se encuentran conectados.

- Dispositivos de consumo permanente: En esta categoría se encuentran aquellos aparatos que se encuentran en funcionamiento por días o semanas, como el decodificador de TV por cable, el teléfono y alarma, entre otros.

La idea es utilizar el perfil de consumo de potencia dependiendo la categoría a la que pertenecen como huella o firma del dispositivo eléctrico y a partir de esto identificar los dispositivos que se encuentran conectados, esto se explica mas a detalle en el Capítulo 2.

Para éste problema se han realizado dos tipos de análisis de las señales eléctricas para identificar las cargas que se encuentran conectadas a la red en los sistemas *NILM*, que son el análisis en estado estable y análisis en estado transitorio.

En el análisis en estado estable se han probado diferentes métodos en el estudio de la señal para poder diferenciar los dispositivos, como lo son el cambio en la potencia real y reactiva, características en el dominio del tiempo y de la frecuencia, la trayectoria que forma el voltaje contra la corriente, ruido en la señal eléctrica, etc. Los métodos que se han probado en el análisis de estado transitorio son la potencia y corriente transitoria principalmente. Los trabajos y resultados obtenidos haciendo este tipo de análisis se explican mas a detalle en el Capítulo 2.

Recientemente Sidhant Gupta et al. [Gupta10] investigaron que la mayoría de los aparatos eléctricos actuales que cuentan con una fuente de alimentación conmutada (*SMPS*, del inglés *Switch Mode Power Supplies*) producen interferencia electromagnética (*EMI*, del inglés *Electromagnetic Interference*) que es repetible y que son específicas para cada dispositivo eléctrico, por lo que la *EMI* puede ser usada como firma o huella de cada dispositivo y a partir de esto identificar que aparato eléctrico se encuentra en funcionamiento. Éste trabajo se centra en la identificación de dispositivos eléctricos basándose en la interferencia electromagnética que producen.

Para realizar el estudio de los sistemas *NILM* como en muchos otros sistemas se debe recabar una gran cantidad de datos para formar conjuntos de entrenamiento, validación y prueba, y a partir de estos conjuntos probar o desarrollar diferentes técnicas que solucionen los problemas del sistema.

Para generar los conjuntos de datos en los sistemas *NILM* generalmente se esta muestreando constantemente el voltaje, corriente y la *EMI* en la red eléctrica a supervisar. Esto con el fin de observar las variaciones que se producen en estas mediciones al momento de encender un dispositivo eléctrico.

Con los conjuntos de datos obtenidos, las fases de entrenamiento y validación se puede dividir en tres etapas principalmente, siendo éstas la detección de los eventos (donde evento se refiere al momento de encendido o apagado de un dispositivo eléctrico), la extracción de características de los dispositivos eléctricos y por último la clasificación

para reportar cual o cuales dispositivos se encuentran en funcionamiento en cada instante de tiempo.

Para realizar la clasificación se necesita que la extracción de características sea de manera precisa, y a su vez para generar la extracción de características es de suma importancia que la detección de los eventos sea de manera exacta. Éste proceso se puede visualizar a través de un diagrama de bloques, como se muestra en la Figura 1.2, donde a partir de la detección de eventos se procede a la extracción de características y después se puede hacer la clasificación.



Figura 1.2: Diagrama de bloques de las etapas: detección de eventos, extracción de características y clasificación que se requieren en los sistemas *NILM*.

Uno de los inconvenientes cuando se están muestreando los diferentes parámetros en la red eléctrica, es que se deben etiquetar los momentos de encendido y apagado de los dispositivos eléctricos. Generalmente éste etiquetado se realiza de forma manual (al no contar con sensores como en los sistemas *ILM*) y genera una desalineación entre el tiempo de etiquetado y el tiempo exacto en el que sucedió el evento. Esta desalineación genera un problema en las etapas de detección de los eventos y extracción de características, por lo que en el Capítulo 3 se propone una metodología para realizar la alineación de eventos basado en las componentes de baja frecuencia.

## 1.2. Planteamiento del Problema

El problema principal en el estudio de los sistemas *NILM*, que es en el que se centra ésta tesis, es identificar qué dispositivos eléctricos se encuentran en funcionamiento en cada instante de tiempo, en base a mediciones hechas en un solo punto dentro de la red eléctrica. El enfoque utilizado en esta investigación es separar las cargas en base a la interferencia electromagnética que producen los dispositivos eléctricos al momento que entran en funcionamiento.

Para lograr la separación de cargas eléctricas es necesario implementar las etapas de detección de eventos, extracción de características y clasificación. Debido a la carencia de equipo para recabar datos de la interferencia electromagnética que se produce en un hogar, se hará uso de varios conjuntos de datos que se encuentran de manera publica en internet [Kaggle], estos conjuntos de datos contienen información de distintas mediciones en la red eléctrica a lo largo de varios días y de 4 casas distintas.

Uno de los conjuntos de datos y tal vez el mas significativo para esta investigación es el espectrograma del ruido electromagnético en altas frecuencias, generado en la red eléctrica. A éste conjunto de datos lo denominaremos con la letra  $H$ , donde  $H \in \mathbb{R}^{n \times h}$ ,  $n$  es el número de frecuencias discretas que conforman el espectrograma y  $h$  es el número de muestras capturadas.

Para realizar la clasificación correctamente y poder reportar que dispositivos se encuentran en funcionamiento en un instante de tiempo, es necesario realizar la extracción de características de los dispositivos y utilizar estas como huellas o firmas electromagnéticas que producen. Para realizar una buena extracción de características es necesario que los eventos estén alineados de forma correcta, por lo que la alineación de los eventos es una de las partes a resolver en esta investigación.

### 1.2.1. Detección de Eventos

Para que el registro del tiempo en que suceden los eventos fuera de manera exacta se necesita que cada dispositivo cuente con sensores que detecten y registren el momento de encendido y apagado (como en los hogares inteligentes), pero resultaría muy costosa la implementación, por lo que en lugar de esto el registro de los tiempos se hace de forma manual y como resultado solo se obtienen tiempos de registro aproximados.

La desalineación que existe entre los tiempos registrados de forma manual y los

tiempos exactos cuando suceden los eventos, genera un problema en las etapas de detección de eventos y extracción de características, ya que se necesita el tiempo exacto para poder identificar la interferencia electromagnética que produce el dispositivo en la red eléctrica.

En la Figura 1.3 se muestra a través de un ejemplo, como se encuentran desalineados los tiempos de encendido de los dispositivos eléctricos. En este ejemplo se muestra un fragmento del conjunto de datos  $H$ , donde  $H^T$  representa a la transpuesta de la matriz  $H$ . Dentro del fragmento mostrado una secadora de pelo es utilizada, se observa claramente que al momento de encendido del dispositivo este comienza a excitar las distintas frecuencias y el tiempo registrado en el conjunto de datos no corresponde a este momento, es esta diferencia la que se tiene que ajustar para que estén alineados los datos y poder realizar la extracción de características de manera precisa.

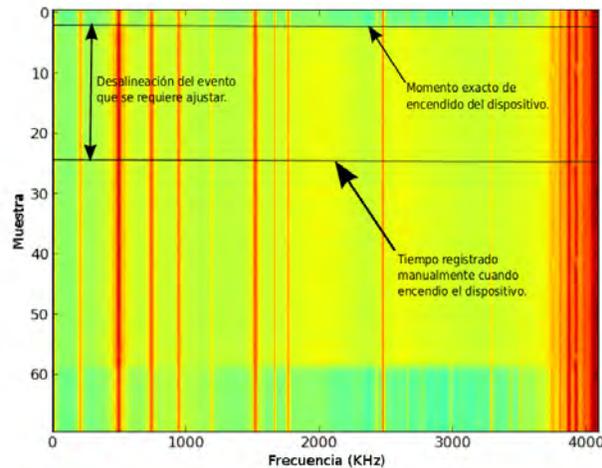


Figura 1.3: Fragmento del conjunto de datos  $H^T$ , donde una secadora de pelo es utilizada y se ejemplifica claramente la desalineación que existe en los eventos al realizar el registro de forma manual.

### 1.2.2. Extracción de Características

Una vez alineados los eventos solo resta extraer las características de cada uno de los distintos dispositivos eléctricos que se encuentran en el conjunto de datos. Para esto basta con observar un instante de tiempo antes y uno después de que suceda el evento, para así poder determinar las alteraciones que produce el dispositivo al momento de encendido.

La problemática principal en la extracción de características es determinar cuales

valores son los mas representativos de los dispositivos, es decir, cuales características ayudan a diferenciar con claridad y facilidad entre los distintos dispositivos eléctricos que existen.

### 1.2.3. Clasificación

En éste trabajo la identificación de que dispositivos se encuentran en funcionamiento en un instante de tiempo, se plantea como un problema de clasificación, por ejemplo, supongamos que contamos con 3 dispositivos eléctricos distintos ( $x_1, x_2$  y  $x_3$ ) y el vector  $[1, 0, 1]$  representa que los dispositivos  $x_1$  y  $x_3$  se encuentran en funcionamiento en un instante de tiempo, entonces, el problema consiste en clasificar cada muestra a una de las  $2^3$  posibles clases.

Resolver el problema de clasificación es lo principal en esta investigación y en la mayoría de estudios que se han realizado utilizando sistemas *NILM*. La clasificación es la que realiza la desagregación de cargas eléctricas reportando cual o cuales dispositivos eléctricos se encuentran en funcionamiento en determinado instante de tiempo. La mayoría de investigaciones actuales en los sistemas *NILM* realizan la clasificación utilizando métodos estocásticos como lo son los algoritmos genéticos, redes neuronales artificiales, algoritmo de recocido simulado y programación genética, entre otros.

Utilizar métodos o algoritmos estocásticos tiene la desventaja de que en problemas de clasificación (como se plantea en este caso) crean un modelo que se ajusta exclusivamente para los datos con los que fue entrenado, por lo que si se utilizan para realizar la desagregación en un hogar. Este se tiene que someter a un proceso de entrenamiento en dicho hogar y se ajustara únicamente para clasificar los dispositivos eléctricos que se utilizaron en la etapa de entrenamiento.

El principal problema en la etapa de clasificación es encontrar o desarrollar una metodología que realice la identificación de aparatos eléctricos de forma general, sin importar el número de dispositivos y el hogar donde se pruebe, esto basándose en la premisa que reporta Sidhant Gupta [Gupta10] en su investigación, donde las firmas o huellas electromagnéticas que producen los dispositivos eléctricos son repetibles (la interferencia electromagnética que genera un dispositivo es similar cada que se enciende), específicas (dos dispositivos distintos generan distinta interferencia electromagnética) e independientes de la casa donde se encuentren.

Cabe destacar como se reporta mas a detalle en el Capítulo 6 que si los fabricantes

proporcionarán por norma la firma o huella electromagnética que producen los dispositivos eléctricos, las etapas de detección de eventos y extracción de características no serían necesarias y se haría la etapa de clasificación directamente.

### **1.3. Justificación y Motivación**

La creciente demanda de energía eléctrica es un problema en la mayoría de países en el mundo, sobre todo en los países que se encuentran en desarrollo como México, entonces, existe la necesidad de controlar este problema disminuyendo el consumo de electricidad o creando nuevas fuentes de energía eléctrica que ayuden en la producción requerida para satisfacer dicha demanda energética.

Esta investigación se justifica y está motivada por la idea de reducir el consumo de energía eléctrica en las industrias y hogares, aportando una idea nueva que pueda dar solución al problema de desagregación de cargas eléctricas que existe en los sistemas *NILM*, haciendo uso de la interferencia electromagnética que producen los aparatos eléctricos.

### **1.4. Objetivos y Contribución de la Tesis**

#### **1.4.1. Objetivo General**

Implementar un esquema con un enfoque de mínimos cuadrados que pueda reportar cual o cuales dispositivos eléctricos se encuentran conectados a la red eléctrica en cada instante de tiempo, a partir de un conjunto de mediciones capturadas (voltaje, corriente e interferencia electromagnética).

#### **1.4.2. Objetivos Particulares**

Para lograr el objetivo principal de esta tesis se plantean los siguientes objetivos particulares:

- En base a los conjuntos de datos obtenidos de internet, crear un programa para calcular la potencia real, reactiva y aparente en cada instante de tiempo, a partir de las mediciones de voltaje y corriente que contienen dichos conjuntos de datos.
- Implementar una metodología que realice la detección y alineación de eventos, para

corregir el error que existe entre el tiempo registrado en los conjuntos de datos y el tiempo real cuando sucedió el evento.

- Realizar la extracción de características de cada uno de los dispositivos eléctricos presentes en los conjuntos de datos.
- Implementar y probar una metodología que realice la clasificación de cada una de las mediciones y con esto reportar cual o cuales dispositivos eléctricos se encuentran en funcionamiento.

### 1.4.3. Contribución de la Tesis

Este trabajo pretende aportar al estado del arte de los sistemas *NILM* un enfoque y planteamiento distinto que de solución al problema de identificar cual o cuales dispositivos eléctricos se encuentran en funcionamiento en un determinado instante de tiempo, esto a partir de las mediciones del ruido electromagnético generado en la red eléctrica.

Otro propósito de esta investigación es resaltar la importancia de que los fabricantes de aparatos eléctricos proporcionen la huella o firma electromagnética (espectrograma de la interferencia electromagnética que producen) de cada dispositivo, esto con el fin de crear un dispositivo con un sistema *NILM* que solo necesite de dichas firmas para funcionar sin importar el lugar donde se ponga en funcionamiento.

## 1.5. Descripción de Capítulos

El resto de este trabajo está organizado de la siguiente forma.

En el Capítulo 2 se muestra parte del estado del arte en los sistemas *NILM*, mostrando distintos enfoques, análisis y resultados obtenidos, los dos principales tipos de análisis son estado estable y estado transitorio.

En el Capítulo 3 se expone el procedimiento usado para alinear y detectar los distintos eventos presentes en los conjuntos de datos, también se reportan y discuten los resultados obtenidos.

El Capítulo 4 plantea un modelo lineal de las mediciones y se propone una metodología para realizar clasificación e identificación de cargas simultaneas.

En el Capítulo 5 se presentan distintas pruebas realizadas a la metodología propuesta con distintos conjuntos de datos y se presentan los resultados obtenidos.

Por último en el Capítulo 6 se concluyen los resultados obtenidos al término de este trabajo, así como el aporte que realiza al área y se proponen algunos trabajos futuros que complementan esta investigación.

## Capítulo 2

# Trabajo Relacionado

El estudio de los sistemas *NILM* no es algo nuevo, se ha estudiado desde hace un poco mas de dos décadas, sin tener resultados satisfactorios para su implementación en la vida diaria de la humanidad [Hart85], pero debido al crecimiento en ciencia y tecnología en recientes años varios investigadores en el área de ingeniería eléctrica han retomado el interés en el estudio de estos sistemas, con el objetivo de producir informes del consumo de potencia individual de los dispositivos eléctricos.

A lo largo de la historia en el estudio de los sistemas *NILM* se han realizado dos tipos de análisis principalmente, uno de ellos y el mas utilizado por los investigadores es el *análisis en estado estable*, que como su nombre lo dice extrae características de los dispositivos eléctricos para utilizarlos como huellas o firmas cuando estos llegan a su estado estable de operación, estas características son generalmente la potencia real y la potencia reactiva, donde se pretende identificar que aparato se encuentra en funcionamiento a partir de estas dos mediciones.

El segundo es el *análisis en estado transitorio*, donde básicamente se trata de extraer características de los dispositivos eléctricos al momento de ponerlos en funcionamiento, observando las formas de onda o alteraciones que producen en la red eléctrica antes de llegar a un estado estable. Una desventaja de este tipo de análisis es que se necesita de hardware especializado para capturar mediciones en la red a una alta frecuencia de muestreo (MHz) y por consecuencia con mayor costo..

Este capítulo se divide de la siguiente forma, en la Sección 2.1 se presentan algunas investigaciones y resultados obtenidos a lo largo de la historia que se han realizado en el estudio de los sistemas *NILM* utilizando un análisis en estado estable y por su parte en la

Sección 2.2 con un análisis en estado transitorio, por ultimo en la Sección 2.3 se presenta un resumen de lo visto en este capítulo.

## 2.1. Análisis en Estado Estable

De acuerdo a lo reportado por Zoha et al. [Zoha12] el análisis en estado estable se puede dividir de acuerdo a los métodos y características extraídas de los dispositivos que se utilizan para su estudio como se muestra a continuación.

### 2.1.1. Cambios de Potencia

Uno de los primeros enfoques registrados en el estudio de los sistemas *NILM* fue en 1992 por G. W. Hart [Hart92], donde solo se enfoca en identificar dispositivos eléctricos con estado de operación ON-OFF, para esto utiliza los cambios en la potencia real  $\Delta P$  y potencia reactiva  $\Delta Q$  como características de los aparatos a identificar,  $\Delta$  representa el cambio que existe cuando un dispositivo entra en funcionamiento.

Se utilizan generalmente  $\Delta P$  y  $\Delta Q$  debido a que los dispositivos que son puramente resistivos el voltaje y la corriente siempre se encuentran en fase y consumen solamente potencia real. Para que las cargas sean puramente reactivas se necesita que el voltaje y corriente estén desfasados  $90^\circ$ . La mayoría de dispositivos eléctricos contienen elemento inductivos y capacitivos que son los que generan un desfase entre el voltaje y la corriente, por lo que estos consumirán tanto potencia real como potencia reactiva.

Al utilizar solo  $\Delta P$  y  $\Delta Q$  como huellas o firmas de los dispositivos eléctricos, el dominio de búsqueda es un espacio bidimensional, donde utiliza técnicas de agrupamiento (*clustering*), donde desarrollaron un algoritmo que determina el número apropiado de clusters para realizar la desagregación de las cargas y poder identificar que aparatos se encuentran conectados en ese instante de tiempo.

En la Figura 2.1 se muestra de manera gráfica como se sitúan los dispositivos eléctricos en el espacio  $\Delta P$ - $\Delta Q$ , donde los aparatos que son resistivos como el calentador consumen pura potencia real y los dispositivos que cuentan con motores como la bomba de agua consumen en su mayoría potencia reactiva.

Hart [Hart85] reporta una precisión del método utilizado de entre el 75% y 90% después de realizar pruebas en tres casas distintas con dispositivos eléctricos que solo cuentan con dos estados de operación (ON-OFF), donde uno de los aparatos que no logro clasificar de

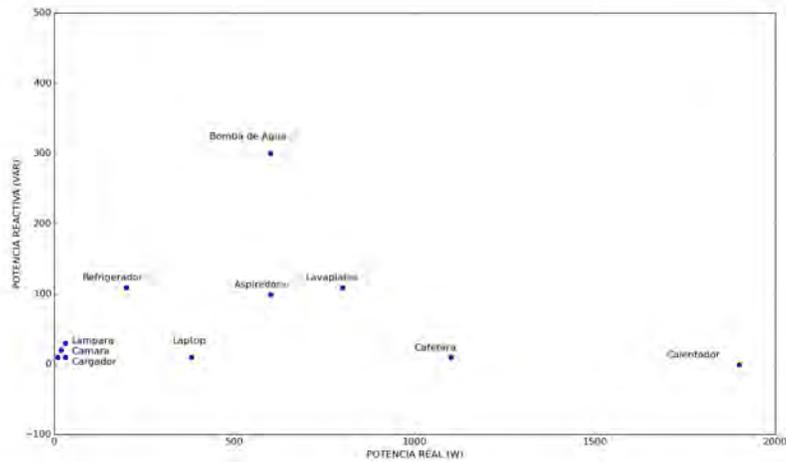


Figura 2.1: Espacio bidimensional  $\Delta P - \Delta Q$  utilizado para la desagregación de cargas.

manera correcta en varias ocasiones es el refrigerador, esto lo adjudica a que el refrigerador presenta un ciclo de descongelación, por lo que estaría en la categoría de maquina de estados finitos (FSM) en lugar de la categoría ON-OFF.

Por otra parte Norford [Norford96], Farinaccio [Farinaccio99] y Marceau [Marceau00] utilizan como única característica de los dispositivos eléctricos la potencia real que consumen, utilizando técnicas en base a reglas y reconocimiento de patrones para la desagregación de cargas. Se obtuvieron buenos resultados para aparatos que registran un consumo de potencia elevado como los calentadores, refrigeradores y bombas hidráulicas, pero persiste el problema de identificar cargas de similares magnitudes que se superponen. La Figura 2.2 muestra como ejemplo y con fines ilustrativos como se observarían los cambios en la potencia real a lo largo del tiempo al momento de conectar y desconectar distintos dispositivos eléctricos.

### 2.1.2. Características del Voltaje y Corriente

El objetivo de superar las limitaciones que se tienen al utilizar como características los cambios en la potencia real  $\Delta P$  y potencia reactiva  $\Delta Q$  se han realizado otras investigaciones agregando como características, parámetros que se pueden obtener de las formas de onda que forman el voltaje ( $V$ ) y la corriente ( $I$ ). Figueiredo [Figueiredo11] hace uso de un tercer parámetro que es el factor de potencia, ya que este sirve de ayuda para discriminar

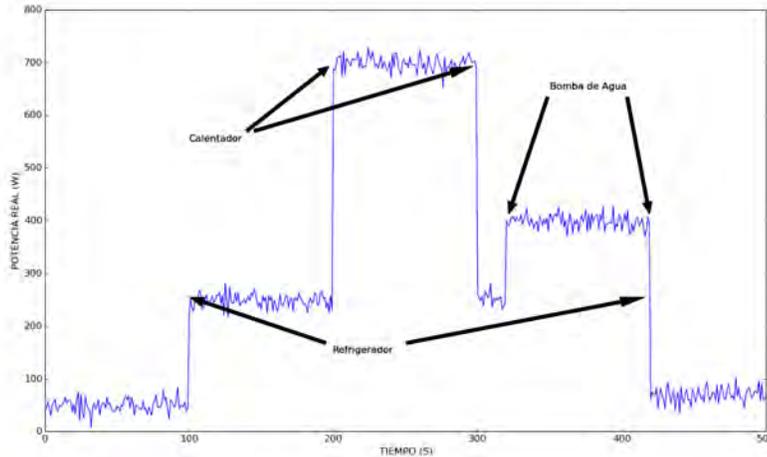


Figura 2.2: Variaciones en la potencia consumida a lo largo del tiempo al momento de conectar aparatos a la red eléctrica.

entre dispositivos de tipo resistivos, inductivos y capacitivos.

Los métodos utilizados como clasificadores son Máquina de Soporte Vectorial (*SVM*, del inglés *Support Vector Machine*) [Cortes95] y *K* Vecinos mas Cercanos (*K-NN*, del inglés *K-Nearest Neighbors*) [Silverman51], se realizaron pruebas con 6 dispositivos eléctricos distintos que son una lámpara incandescente, horno de microondas, tostador, maquina de café y dos televisores LCD de 22 y 23 pulgadas. Los resultados reportados tienen una precisión en la clasificación del  $96.3 \pm 6.4\%$  utilizando *SVM* y  $99.1 \pm 0.9\%$  con *K-NN*.

Ruzzelli et al. [Ruzzelli10] presentan *RECAP* (del inglés, *RECOgnition of electrical Appliances and Profiling in real-time*), que fué el primer esfuerzo sistemático para integrar perfiles y reconocimiento de los dispositivos eléctricos en un único marco de trabajo. *RECAP* utiliza un sensor de monitoreo de energía inalámbrico, el sensor transmite de manera inalámbrica los datos muestreados en la red y los envía a una máquina para su almacenamiento y procesamiento. El sistema *RECAP* consiste de tres partes principalmente:

1. Guiar al usuario para crear perfiles de los dispositivos eléctricos y crear una base de datos con firmas únicas de los electrodomésticos.
2. Utilizar esas firmas para entrenar una red neuronal artificial [Rojas96] (*ANN*, del inglés *Artificial Neural Network*) que es empleada para reconocer las actividades de los diferentes electrodomésticos.

3. Proveer de un descriptor de carga para permitir una evaluación comparativa por cada dispositivo eléctrico.

*RECAP* incorpora nuevas características de los electrodomésticos a las ya mencionadas  $\Delta P$ ,  $\Delta Q$  y el factor de potencia, que son extraídas de las formas de onda que producen el voltaje y la corriente. Estas con el voltaje *RMS* ( $V_{RMS}$  del inglés, *Root Mean Square*), la corriente *RMS* ( $I_{RMS}$ ), el voltaje pico ( $V_p$ ) y la corriente pico ( $I_p$ ).

Con la *ANN* previamente entrenada, esta se utiliza como clasificador, por lo que las características serán las entradas del clasificador y este determinará que electrodomésticos se encuentran en funcionamiento. Las pruebas reportadas en esa investigación se realizaron de manera real con aparatos eléctricos que se encuentran comúnmente en una cocina. Los resultados publicados con las pruebas realizadas al sistema *RECAP* reportan una precisión del 84 % para todos los casos de estudio.

Recientemente Jiaming Li, Sam West y Glenn Platt [Li12] presentaron un algoritmo de descomposición de la potencia basado en una maquina de soporte vectorial (*SVM*), para esto la corriente de las cargas se puede representar tanto en el dominio del tiempo como en el dominio de la frecuencia como se muestra en la Figura 2.3(a) y 2.3(b).

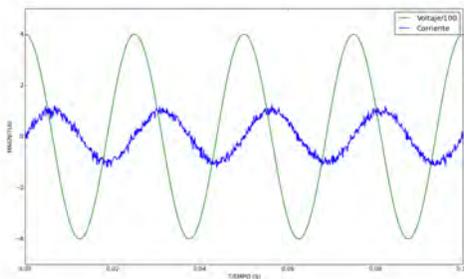
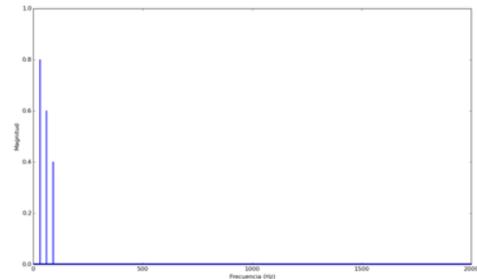
(a) *V-I en el dominio del tiempo*(b) *I en el dominio de la frecuencia*

Figura 2.3: (a) Formas de onda del voltaje y corriente en el dominio del tiempo. (b) Representación de la corriente en el dominio de la frecuencia.

Haciendo uso de los valores de la corriente en el dominio de la frecuencia se introducen al modelo que proponen basado en un *SVM* para estimar la proporción de potencia constante (*CP*, del inglés *Constant Power*) en relación con la impedancia constante (*CI*, del inglés *Constant Impedance*) de las cargas. Se observó que la mayoría de cargas con

$CI$  son cargas resistivas y con  $CP$  son cargas conmutables, por lo que la relación de estos coeficientes  $CP/CI$  es de ayuda para discriminar entre este tipo de electrodomésticos.

### 2.1.3. Trayectoria del Voltaje y Corriente

Lam, Fung y Lee [Lam07] proponen un novedoso método para clasificar las cargas de los electrodomésticos, para esto se sugiere caracterizar los dispositivos eléctricos haciendo uso de la trayectoria que forman en un plano de 2 dimensiones el voltaje y la corriente. A grandes rasgos la metodología que proponen para generar la taxonomía es la siguiente:

1. Deben ser medidas las formas de onda del voltaje y la corriente que producen los dispositivos eléctricos utilizados en la casa que se pretende monitorear.
2. Se construye la trayectoria Voltaje-Corriente para formar las firmas o huellas de cada una de las cargas.
3. Se extraen características de la figura que produce la trayectoria Voltaje-Corriente creada previamente.
4. Se aplican métodos de agrupamiento para asociar los dispositivos eléctricos.
5. La taxonomía de cargas se construye de acuerdo a los resultados que producen los métodos de agrupamiento.
6. La taxonomía basada en la trayectoria Voltaje-Corriente se compara con otras taxonomías basadas en otras métricas, como las que utilizan la potencia.

Son 7 las características extraídas de la figura que forma la trayectoria Voltaje-Corriente de los electrodomésticos, siendo estas la asimetría, área, curvatura de la línea media, auto intersección, pendiente del segmento medio, área de los segmentos izquierdo y derecho, y el pico del segmento medio. Estas características son utilizadas como entrada de los métodos de agrupamiento para asociar los distintos dispositivos eléctricos y poder formar la taxonomía.

En la Figura 2.4 se muestra un ejemplo del tipo de formas que se crean a partir de la trayectoria Voltaje-Corriente, así como la división de la gráfica para extraer las características ya mencionadas. Por ejemplo, el pico es la distancia que existe entre la línea media y la altura de la figura que forma la trayectoria V-I. Los resultados reportados de

este trabajo es una taxonomía de 9 grupos distintos, donde por lo regular se agrupan de acuerdo a la cantidad y tipo de potencia consumida.

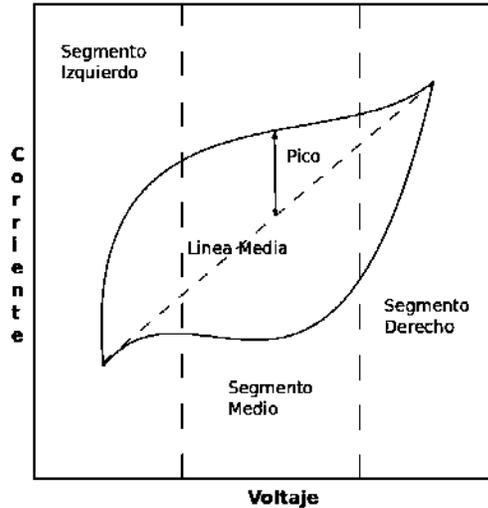


Figura 2.4: Ejemplo de la trayectoria Voltaje-Corriente.

#### 2.1.4. Ruido en la señal de Voltaje

Sidhant Gupta et al. [Gupta10] presentaron el sistema *ElectriSense* como una nueva solución para detectar y clasificar automáticamente el uso de dispositivos eléctricos en los hogares. *ElectriSense* se basa en que los dispositivos que cuentan con fuentes de alimentación conmutadas (*SMPS*, del inglés *Switch Mode Power Supplies*) generan ruido electromagnético en altas frecuencias mientras estos se encuentran en operación y lo propagan a lo largo de red eléctrica.

En la Figura 2.5 se observa el espectrograma del ruido electromagnético registrado en un periodo de tiempo donde las luces de la cochera fueron encendidas, se muestra claramente que en un lapso de tiempo (cuando las luces están encendidas) algunas frecuencias discretas del espectrograma son alterados debido a la EMI producida, por lo que estas alteraciones producidas pueden ser usadas como huellas del dispositivo eléctrico.

Determinaron que esta EMI es repetible y específica para cada aparato eléctrico, por lo que puede ser usada como firma o huella de cada dispositivo y a partir de estas introducir las a un clasificador que pueda reportar que dispositivos se encuentran en funcionamiento. En la siguiente lista se describen de manera básica los pasos seguidos para la

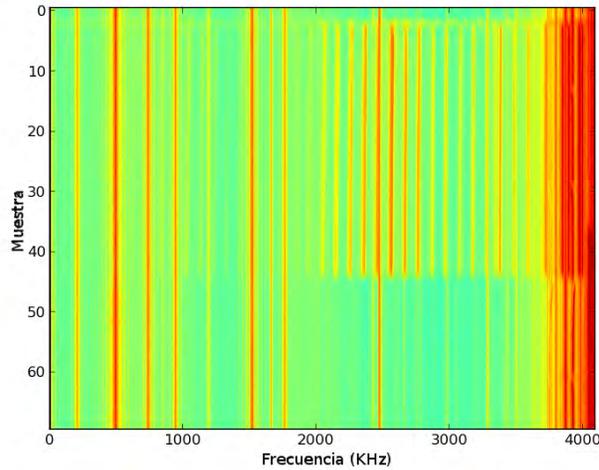


Figura 2.5: Fragmento del espectrograma de la interferencia electromagnética, donde dentro de éste fragmento las luces de la cochera son encendidas.

implementación del sistema.

1. Filtrar la señal de 60Hz de la red eléctrica mediante un filtro pasa altas, debido a que la EMI se encuentra solo en altas frecuencias.
2. La señal filtrada se conecta a un dispositivo de adquisición de datos de alta frecuencia que convierte la señal de analógica a digital con una frecuencia de muestreo de 1 MHz.
3. Los datos son pasados en bloques a una computadora que los transforma del dominio del tiempo al dominio de la frecuencia mediante la transformada rápida de Fourier.
4. Un algoritmo utiliza los datos en el dominio de la frecuencia para detectar los eventos que ocurren. Básicamente realiza la diferencia entre vectores a lo largo del tiempo para detectar cambios en los datos.
5. Cuando un evento ocurre se ejecuta un algoritmo que extrae las características del dispositivo. Este algoritmo ajusta una función gaussiana a los picos detectados, donde de las gaussianas se extraen media, amplitud y varianza que se utilizan como las características de cada aparato.
6. Por último se utiliza como clasificados K-NN y distancia euclidiana como métrica, ya que se ha probado que funciona bien con problemas de baja dimensionalidad.

Realizaron pruebas en 7 casas distintas a lo largo de 6 meses, donde en promedio se probaron 14 dispositivos por casa. Se recolectaron 2,576 eventos donde en mayor parte son encendido de lámparas. Utilizando validación cruzada de 10 pliegues reportan una precisión del 91.75 %.

## 2.2. Análisis en Estado Transitorio

El análisis en estado transitorio se centra en estudiar las perturbaciones que producen dispositivos eléctricos cuando estos se ponen en funcionamiento, los transitorios generalmente ocurren en fracciones de milisegundos o microsegundos, por lo que resulta costoso este tipo de análisis en comparación en el de estado estable.

Se ha demostrado que los transitorios que producen los electrodomésticos brindan mas información para diferenciarlos entre ellos que cuando llegan a su estado estable, sin embargo uno de los mayores inconvenientes es que se necesita de hardware que capture la señal a una alta frecuencia de muestreo para poder detectar estos transitorios [Figueiredo11].

En esta sección se presentan algunos trabajos que existen en el estudio de sistemas NILM usando un análisis en estado transitorio.

### 2.2.1. Voltaje y Corriente Transitoria

PChang et al. [Chang08] analizan la idea de que los métodos que utilizan características en estado estable como la potencia real ( $P$ ) y la potencia reactiva ( $Q$ ) no son capaces de distinguir los dispositivos que se encuentran conectados al mismo tiempo y realizan un determinado consumo de potencia, de otro grupo de aparatos distintos que tienen en conjunto un consumo de potencia similar.

Entonces proponen utilizar como características adicionales el voltaje, corriente y energía transitoria que se presenta cuando estos se ponen en funcionamiento, ya que estos tienen la propiedad de que no son aditivos, por lo que si dos grupos de dispositivos distintos tienen consumos de potencia similares, los transitorios serán distintos para poder diferenciarlos. En la Figura 2.6 se muestra un ejemplo de los transitorios que puede producir un aparato eléctrico al conectarlo a la red eléctrica.

Lo primero que proponen mediante un algoritmo es realizar perfiles de los dispositivos eléctricos de acuerdo a la energía transitoria de los dispositivos cuando estos entran en funcionamiento.

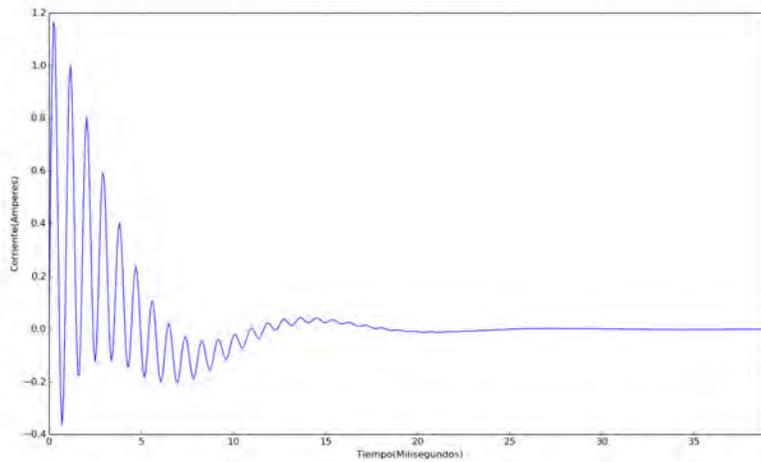


Figura 2.6: Ejemplo de las perturbaciones (transitorios) que puede producir un dispositivo antes de llegar a estado estable.

Para la detección de eventos proponen el análisis de la transformada wavelet para detectar y localizar varios transitorios cuando un dispositivo es encendido. El enfoque también determina cuales de los coeficientes de la transformada wavelet (*WTC*, del inglés *Wavelet Transformation Coefficients*) son únicos para que estos representen el momento que ocurre un evento.

La mayoría de dispositivos con los que se probó tienen perfiles de consumo que son repetibles, por lo que estos son utilizados como características para distinguir que dispositivos se encuentra en funcionamiento.

Las pruebas fueron realizadas con 5 dispositivos en su mayoría motores de una planta industrial, los clasificadores probados son una red neuronal artificial con propagación hacia atrás (*BP*, del inglés *Back Propagation*) y una red de aprendizaje de cuantificación vectorial (*LVQ*, del inglés *Learning Vector Quantization*). Los resultados reportados son de una precisión del 95.3% para *BP* y del 42.94% para *LVQ*.

### 2.2.2. Ruido en la señal de Voltaje Transitoria

Patel et al. [Patel07] presentan un interesante trabajo donde utilizan el ruido que existe en la señal de voltaje, el enfoque se basa en que las cargas eléctricas que cuentan con una fuente de alimentación conmutada (*SMPS*) producen ruido que se propaga a lo largo

de toda la instalación eléctrica tanto de manera continua como de forma transitoria.

En éste caso se enfocan en el ruido transitorio generado, donde se producen impulsos que solo duran algunos microsegundos pero que proporcionan una gran cantidad de información en los componentes de frecuencia de su respectivo espectrograma, en un rango que va de los 10 Hz a los 100 KHz.

También observaron que la mayoría de estos dispositivos que cuentan con una *SMPS* producen ruido transitorio tanto al momento de encendido como al momento de apagado, por lo que estas excitaciones que producen en los distintos componentes de frecuencia al momento de encendido y apagado, pueden ser utilizadas como características para poder identificarlos.

Para la clasificación se empleó una máquina de soporte vectorial (*SVM*), que a demostrado obtener buenos resultados para diferenciar datos que son similares pero no idénticos, siendo adecuado para el conjunto de datos que se genero en la investigación.

Las pruebas se realizaron en 6 hogares distintos, variando el estilo, antigüedad, tamaño y ubicación. En la primer casa se hicieron pruebas a lo largo de 6 semanas para determinar la precisión de la clasificación con distintos electrodomésticos. En las otras 5 casas solo se hicieron pruebas con un lapso de una semana por cada hogar, para corroborar los resultados obtenidos en la primer casa.

Se generaron alrededor de 3000 eventos (encendido o apagado de un dispositivo eléctrico) a lo largo del periodo de prueba, donde los resultados indican que el método puede aprender y clasificar los eventos con un rango de precisión del 85 – 90 %.

## 2.3. Resumen del Capítulo

En éste capítulo se hizo una revisión del estado del arte en el estudio de los sistemas NILM, donde se puede dividir principalmente en dos tipos de análisis, análisis en estado estable y análisis en estado transitorio.

El análisis en estado estable se basa en extraer características de las alteraciones que producen los dispositivos en la red eléctrica de manera continua mientras estos se encuentran en funcionamiento. En la Tabla 2.2 se presenta un resumen de este tipo de análisis de acuerdo a los métodos y características utilizadas.

Por otra parte el análisis en estado transitorio se centra en la idea que las perturbaciones transitorias que produce un dispositivo cuando este se enciende, proporcionan mucha

mas información que cuando llega a modo estable, por lo que sería efectiva la identificación de las cargas. La desventaja es la necesidad de un hardware especializado que capture las muestras de la señal a altas frecuencias para poder detectar los transitorio, ya que estos ocurren en fracciones de segundos (microsegundos). En la Tabla 2.1 se muestra un resumen de los métodos vistos en la Sección 2.2.

En el siguiente capítulo se presenta una metodología que realiza el primer paso para desagregación de cargas eléctricas en el estudio de los sistemas *NILM*, que es la alineación de eventos basándose en las componentes de baja frecuencia como lo son el voltaje y corriente, reportando los resultados obtenidos.

<b>Método</b>	<b>Características</b>	<b>Ventajas</b>	<b>Desventajas</b>
Voltaje y Corriente Transitoria	Transitorios del Voltaje ( $V$ ) Corriente ( $I$ ) y Energía ( $U$ )	Dispositivos con consumos de potencia similares pueden ser diferenciados y clasificados	Hardware con frecuencias altas de muestro, Monitoreo continuo
Ruido en la señal de Voltaje Transitoria	Espectrograma del ruido transitorio generado	Efectivo para identificar y clasificar dispositivos que cuentan con SMPS	No todos los dispositivos cuentan con SMPS.

Tabla 2.1: Resumen de los métodos y características utilizados en el análisis en estado transitorio.

Método	Características	Ventajas	Desventajas
Cambios de Potencia	Potencia real ( $\Delta P$ ) Potencia reactiva ( $\Delta Q$ )	Efectivo para identificar cargas con un elevado consumo de potencia y con estado de operación ON-OFF	Bajo rendimiento con dispositivos con cargas similares y que cuenten con mas de dos estados de operación
Características del Voltaje y Corriente	Voltaje ( $V$ ) Corriente ( $I$ ) $V_{RMS}$ $I_{RMS}$ Voltaje Pico ( $V_p$ ) Corriente Pico ( $I_p$ ) Potencia constante ( $CP$ ) Impedancia constante ( $CI$ )	Buena para identificar y diferencias cargas resistivas e inductivas	Difícil identificar eventos que son muy similares, baja precisión con dispositivos de estado continuamente variable
Trayectoria del Voltaje y Corriente	Asimetría Área Curvatura de la línea media Auto intersección Pendiente del segmento medio Área de los segmentos izquierdo y derecho Pico del segmento medio	Generar una taxonomía para agrupar los dispositivos eléctricos debido a la trayectoria del voltaje y corriente	Computacionalmente costoso, cargas pequeñas no tienen buena trayectoria V-I
Ruido en la señal de Voltaje	EMI Media, amplitud y varianza de los picos en el espectrograma de la EMI	Detección de múltiples eventos, efectivo para identificar dispositivos con SMPS	No todos los dispositivos cuentan con SMPS

Tabla 2.2: Resumen de los métodos y características utilizados en el análisis en estado estable.



## Capítulo 3

# Alineación de eventos basado en las componentes de baja frecuencia

Cuando se implementa un sistema *NILM* se debe recopilar gran cantidad de información para poder generar conjuntos de entrenamiento y validación, y con estos desarrollar, probar y evaluar diferentes métodos que puedan dar solución al problema planteado.

En este caso se cuenta con información del ruido electromagnético en altas frecuencias, generado en la red eléctrica a lo largo de varios días y de diferentes casas. A este conjunto lo denotaremos con la letra  $H$ , así como el registro de tiempo de encendido y apagado de cada dispositivo eléctrico conectado a la red. Este último servirá como etiquetas para los conjuntos de entrenamiento y validación.

Para que el registro del tiempo en que suceden los eventos sea exacta se necesita que cada dispositivo cuente con sensores que detecten y registren el momento de encendido y apagado. Esta implementación resultaría muy costosa, por lo que en lugar de esto el registro de los tiempos se hace de manera manual y como resultado solo se obtienen tiempos de registro aproximados.

Es de suma importancia que la detección de los eventos sea de manera precisa para poder realizar una buena extracción de características de los dispositivos. Por lo tanto el problema consiste en alinear los eventos del tiempo aproximado que se registró al tiempo exacto en que sucedió.

En la Figura 1.3 se muestra de manera mas clara a traves de un ejemplo de como se encuentran desalineados los tiempos de encendido de los dispositivos eléctricos.

En este ejemplo se muestra un fragmento en el tiempo de  $H^T$  donde una secadora de pelo es utilizada. Se observa que al momento de encendido del dispositivo, éste comienza a excitar los distintos armónicos y el tiempo registrado en el conjunto de entrenamiento no corresponde a este momento. Esta diferencia es la que se tiene que ajustar para que estén alineados los datos y poder realizar la extracción de características de manera precisa.

Este capítulo se divide de la siguiente manera, en la Sección 3.1 se describen los conjuntos de datos con los que se cuenta, así como su estructura, en la Sección 3.2 se explica la metodología propuesta para realizar la alineación de eventos. En la Sección 3.3 se presentan los resultados obtenidos en la alineación de eventos. Por último, en la Sección 3.4 se presenta un resumen de lo visto en este capítulo.

### 3.1. Descripción de los conjuntos de datos

Para realizar la alineación se cuentan con varios conjuntos de entrenamiento y validación<sup>1</sup>, los cuales se describen a continuación.

Los conjuntos de datos proporcionados son las medidas de voltaje y corriente en baja frecuencia para cada fase  $\phi$  que representaremos con las letras  $V_\phi$  e  $I_\phi$  respectivamente, donde  $V_\phi, I_\phi \in \mathbb{C}^{m \times h}$ ,  $\phi$  representa el número de fase donde se tomaron las mediciones (en este caso solo existen 2 fases en los hogares),  $m$  es el número de mediciones capturadas en baja frecuencia y  $h$  representa el número de armónicos utilizados (en este caso 6: la fundamental y los primeros 5 armónicos).

Además se cuenta con los tiempos de captura de cada muestra de las mediciones de voltaje y corriente en baja frecuencia. Estos se representan con la letra  $T_\phi$  respectivamente, donde  $T_\phi \in \mathbb{R}^m$ . Además la frecuencia de muestreo es diferente tanto en alta como en baja frecuencia, por lo que no existe una correspondencia directa del tick en baja frecuencia con la de alta.

Por último, se cuenta con un conjunto de datos que contiene las etiquetas de los eventos registrados y denotaremos con la letra  $E$ , cada etiqueta contiene el ID del dispositivo, su nombre, tiempo de encendido y tiempo de apagado.

Ya con los conjuntos de datos descritos se procede a la elaboración de un método que pueda realizar correctamente la alineación de los eventos para poder extraer las

---

<sup>1</sup> Los conjuntos de datos se puede encontrar en la siguiente ruta previamente registrándose en el sitio <http://www.kaggle.com/c/belkin-energy-disaggregation-competition/data>

características de los dispositivos de manera exacta.

## 3.2. Metodología propuesta

Lo primero que se planteó fue realizar la alineación directamente en el dominio de alta frecuencia utilizando el conjunto de datos que contiene el espectrograma del ruido electromagnético generado en la red eléctrica del hogar.

El enfoque utilizado es muy sencillo, simplemente se usó la Ecuación 3.1 para tener como resultado otro conjunto de datos que nos muestra como esta variando el ruido electromagnético a lo largo del tiempo.

$$\Delta H[n, m] = H[n, m] - H[n, m - 1] \quad (3.1)$$

Una vez procesada la información solo es necesario buscar un cambio brusco en el ruido electromagnético generado, en un periodo de tiempo cercano al que se registró manualmente. Esto nos diría que en el momento en que sucedió el cambio, este sería el tiempo donde se realizó el evento exactamente y con esto se podría realizar la alineación en tiempos.

Desafortunadamente debido a que existe una gran cantidad de ruido en las mediciones, este método generaba una gran cantidad de falsos positivos que afectaban mucho al momento de extraer las características de los dispositivos.

Es por esta razón que se decidió trabajar en la baja frecuencia, donde existe menor ruido en las mediciones. Para esto se hará uso de la potencia y a partir de ésta identificar en que momento un aparato comienza o deja de consumir dicha potencia. Lo primero es calcular la potencia consumida en ambas fases. Para esto usaremos la Ecuación 3.2.

$$S_\phi[n] = \sum_{i=0}^{h-1} V_\phi[n, i] I_\phi^*[n, i] \quad (3.2)$$

Donde  $h = 6$  en este caso porque el voltaje  $V_\phi$  y la corriente  $I_\phi$  contiene la fundamental y los primeros 5 armónicos,  $I_\phi^*$  es el conjugado de  $I_\phi$  y  $S_\phi$  representa la potencia consumida en la fase  $\phi$ ,  $S_\phi \in \mathbb{C}^m$ . Al pertenecer la potencia al dominio complejo se puede representar como se muestra en la Ecuación 3.3.

$$S_\phi = P_\phi + jQ_\phi \quad (3.3)$$

Donde  $P_\phi$  y  $Q_\phi$  representan la parte real e imaginaria de la potencia por cada una de las fases. A partir de esto se utilizará solo la parte real de la potencia como la potencia real consumida en los hogares.

Por cada evento se extrae un fragmento de la potencia registrada y realizar un análisis para detectar cuando ocurre el evento exactamente. Para esto hay que formar un vector en forma de caja que multiplicaremos por la potencia real registrada. En la Ecuación 3.4 se muestra la forma de construir una caja de tamaño  $b + a + 1$  aunque el tamaño del vector generado es  $n_L$ .

$$X[n] = \begin{cases} 1 & i - b \leq n \leq i + a \\ 0 & \text{en caso contrario} \end{cases} \quad n = [0, 1, 2, \dots, n_L] \quad (3.4)$$

Donde  $i$  es el índice donde supuestamente ocurrió un evento en el registro hecho manualmente,  $b$  y  $a$  son variables para tomar el fragmento un tiempo antes y después respectivamente del índice  $i$ . Una vez construida la caja se necesita calcular la potencia contenida en esa caja. Sea  $\hat{P}_\phi$  los componentes de potencia de esa caja. Entonces, utilizaremos la ecuación 3.5 para extraer de la potencia el fragmento  $\hat{P}_\phi$  que se utilizará para la alineación de los eventos.

$$\hat{P}_\phi[n] = P_\phi[n]X[n] \quad (3.5)$$

Ya con el fragmento de la potencia extraído se observó en algunos casos que en el momento de encendido del dispositivo ocurren picos de consumo en potencia de muy corta duración antes de llegar a modo estable (como se muestra en la señal superior de la Figura 3.1), lo que afectaba la precisión en la detección de eventos al detectarlos cuando ocurre el pico y no cuando llega a modo estable.

### 3.2.1. Suavizado con kernel gaussiano

Para solucionar esto se aplicó un suavizado a la señal mediante la convolución de esta con un kernel, tomando la idea de investigaciones hechas en el área de procesamiento digital de imágenes [Perona90]. El kernel utilizado es el gaussiano, utilizando la Ecuación 3.6.

$$G[n] = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(n-\mu)^2}{2\sigma^2}} \quad (3.6)$$

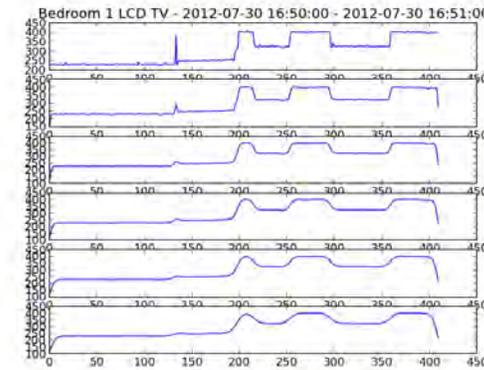


Figura 3.1: Efecto que tiene el suavizado mediante un kernel gaussiano. La primer señal de arriba hacia abajo es la señal pura sin suavizar y a las señales 2,3,4,5 y 6 se aplicó el suavizado con un kernel variando  $\sigma = 1, 2, 3, 4, 5$  respectivamente.

Donde se pueden ajustar los parámetros  $\mu$  que es la media de la distribución y  $\sigma$  es la varianza. Para este caso la media  $\mu$  es igual a cero y el único parámetro que ajusta el tamaño del kernel es la varianza  $\sigma$ .

Una vez generado el kernel se procede a realizar la convolución utilizando la Ecuación 3.7 con el fragmento de la potencia extraído  $\overline{P}_\phi$ . El tamaño del kernel es  $2n_k + 1$ , donde  $n_k = 3 * \sigma$ .

$$Y_\phi[n] = \sum_{k=-n_k}^{n_k} \hat{P}_\phi[k]G[n-k] \quad (3.7)$$

En la Figura 3.1 se muestra el efecto de realizar el suavizado mediante un kernel gaussiano, donde se observa que efectivamente resta amplitud a picos de corta duración. La primer señal (de arriba hacia abajo) es la señal pura sin aplicar el suavizado, y en las siguientes se muestra que conforme hacemos el tamaño del kernel mas grande el suavizado sera mayor utilizando  $\sigma = 1, 2, 3, 4, 5$ .

Ahora con la potencia suavizada mediante la convolución para restar importancia a picos de corta duración se procede a detectar donde ocurren cambios de amplitud en la potencia, para esto se hace uso de la ecuación 3.8. Donde  $\alpha$  es una variable que indica el número de instantes de tiempo atrás con el que se va a comparar con el actual y así determinar como fue cambiando la señal con respecto del tiempo.

$$\overline{Y}_\phi[n] = Y_\phi[n] - Y_\phi[n - \alpha] \quad (3.8)$$

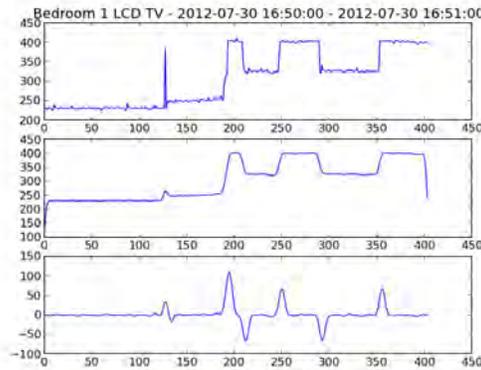


Figura 3.2: Etapas mas significativas de la alineación. En la parte superior es el fragmento de la potencia sin modificación, en el centro la potencia suavizada con un kernel gaussiano generado con  $\sigma = 2$  y en la parte inferior ya procesada la señal que contiene picos donde hay variaciones en la potencia.

Al realizar esta operación la señal resultante es una con picos ubicados en los instantes de tiempo donde ocurre un cambio en la potencia, mientras mas fuerte sea el cambio en la potencia mas amplitud tienen los picos.

Por último solo resta sacar el índice del valor máximo en el arreglo que denominaremos  $t$ , este índice contiene el momento donde existe un cambio brusco positivo en la potencia y quiere decir que en ese instante de tiempo un dispositivo eléctrico fue puesto en funcionamiento. Por lo tanto la hora y fecha exacta la extraemos del vector de tiempos en baja frecuencia; la obtenemos de  $T_\phi[t]$ .

Esta metodología se debe repetir por cada uno de los eventos registrados, al final y almacenando todos los tiempos se obtiene un nuevo conjunto de tiempos que serán los registros de encendido de los dispositivos pero ahora si de forma alineada.

La Figura 3.2 muestra los pasos mas importantes en la alineación, donde en la parte superior se muestra la señal sin modificar, en el centro ya suavizada usando un kernel con  $\sigma = 2$  y en la parte inferior una señal que contiene picos donde existe variaciones en la potencia.

### 3.3. Resultados

Existen varios parámetros que podrían modificar el funcionamiento de la metodología propuesta, en la Tabla 3.1 se muestran los valores utilizados en este trabajo.

Tabla 3.1: Lista de parámetros utilizados

Parámetro	Valor
$b$	100
$a$	300
$\sigma$	2
$n_k$	$3 * \sigma$
$\alpha$	6

Donde  $b$  y  $a$  se utilizan para extraer el fragmento un tiempo antes y un tiempo después respectivamente de donde supuestamente ocurrió el evento. Analizando los datos se observó que en promedio los aparatos duraban encendidos un minuto y que al momento de apagarlos esperaban 30 segundos al menos para encender otro aparato. Para que no existan traslapos entre eventos en un fragmento extraído se optó por los valores  $b = 100$  (alrededor de 17 segundos debido a la frecuencia de muestreo con la que se tomaron los datos en baja frecuencia) y  $a = 300$  (aproximadamente son 50 segundos).

Se observó que  $\sigma = 2$  era suficiente para el suavizado de la potencia, debido a que este generaba un kernel de tamaño 13.  $\alpha = 6$  es para observar como va cambiando la potencia con respecto del tiempo en cada segundo (debido a que 6 muestras en baja frecuencia equivalen a 1 segundo aproximadamente).

La metodología se sometió a 461 eventos registrados en las 4 casas, donde en 444 de las casos pudo alinear el evento de forma correcta y en 17 no, obteniendo una precisión del **96.31 %**.

En la mayoría de los 17 eventos que no se pudieron alinear es porque existen dispositivos eléctricos que consumen muy poca potencia y esta no se alcanza a distinguir debido al ruido que existe en la red eléctrica, o los aparatos eléctricos consumen potencia que está variando constantemente en el tiempo y el método no distingue el momento de encendido entre tantas variaciones. Esto último se puede observar en la Figura 3.3, donde se muestra como un reproductor de video consume potencia de manera variable a lo largo del tiempo (señal superior), por lo que al momento de realizar el suavizado, éste aplana la señal (señal del centro) y al extraer los picos existen muchos de poca amplitud (señal inferior), de manera que no se puede distinguir el momento de encendido del aparato.

Además de la alineación se calculó un índice de desalineación por casa, que es en

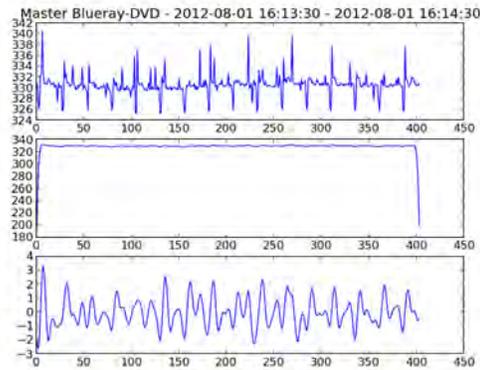


Figura 3.3: Ejemplo de una falla en la alineación con dispositivos eléctricos que consumen potencia de manera variada a través del tiempo. En la parte superior es el fragmento de la potencia variando en el tiempo, en el centro la potencia suavizada con un kernel gaussiano generado con  $\sigma = 2$  y en la parte inferior la señal con picos donde existen cambios en la potencia.

Tabla 3.2: Índice de desalineación por casa

Casa	Índice de desalineación
$C1$	12.711712
$C2$	11.872881
$C3$	16.137405
$C4$	6.689320

promedio el número de muestras que está desalineado un evento. Los índices se muestran en la Tabla 3.2, donde se observa que en la casa  $C3$  es donde están mas desalineados los eventos, y en la casa  $C4$  es donde se realizó el registro de los eventos con mejor precisión.

Por otra parte la verificación de que los eventos se alinearan se tuvo que realizar de forma visual al crear figuras de donde ocurren los eventos y poner una marca donde el método indica qué ocurrió exactamente. Si existiera un método para verificarlo de forma automática, esta metodología no tendría razón de desarrollarse porque el mismo método de verificación podría realizar la alineación.

### 3.4. Resumen de Capítulo

La alineación de los eventos es muy importante debido a que sin ésta no habría una buena extracción de características, y a la vez sin una adecuada extracción de características la clasificación de los eventos no sería posible.

En este capítulo se desarrolló una metodología para alinear los eventos en los conjuntos de entrenamiento para los sistemas NILM donde fueron registrados los tiempos en que ocurrieron de forma manual, debido a que sería muy costoso dotar de sensores a todos los aparatos eléctricos para que registraran el tiempo de manera exacta cuando se ponen en funcionamiento los dispositivos.

Esta metodología ha demostrado tener gran precisión al detectar y alinear al tiempo exacto en que ocurrieron los eventos el **96.31 %** de los casos. Para lograrlo se usó la potencia en baja frecuencia, así como de la convolución de la señal con un kernel gaussiano para realizar el suavizado de esta y con esto detectar cambios en la potencia.

Cabe destacar que esta metodología requiere un tiempo aproximado cuando sucedió el evento para funcionar debido a que los registros se hicieron en hogares habitados. Si en el conjunto de entrenamiento hay una gran variedad de eventos que no fueron etiquetados no habría forma de saber que eventos corresponden con los que si fueron etiquetados.

A partir de los resultados obtenidos en este capítulo se procedió a extraer las características de los dispositivos, tomando como huella o firma electromagnética los 4096 armónicos que excitan al momento de que estos se ponen en funcionamiento.

En el siguiente capítulo se explica la metodología propuesta para realizar la identificación de que dispositivos eléctricos se encuentran en funcionamiento en cada instante de tiempo, a partir del conjunto de huellas electromagnéticas de todos los dispositivos.



## Capítulo 4

# Desagregación de cargas basadas en proyecciones

Desafortunadamente a pesar de realizar una alineación de los eventos y una buena extracción de características la metodología propuesta en este capítulo no puede dar buenos resultados con los conjuntos de datos que se contaba. Esto debido a que en los conjuntos de datos había una gran cantidad de dispositivos que no fueron etiquetados y por consecuencia no se pudo extraer su firma electromagnética.

La metodología propuesta en este capítulo requiere del conjunto de firmas electromagnéticas de todos los dispositivos para su correcto funcionamiento, por lo que los dispositivos que no fueron etiquetados y de los cuales no se contaba con su firma causarían muchos errores en la metodología.

Una de los principales inconvenientes del estudio de los sistemas NILM enfocándose en la interferencia electromagnética que producen los dispositivos eléctricos, es la falta de conjuntos de datos (entrenamiento y validación) que sean accesibles de manera pública a través de internet. El único que se encontró al término de este trabajo es el tratado en el Capítulo 3, pero debido a la gran cantidad de dispositivos sin firma fue imposible implementar la metodología propuesta con ese conjunto de datos.

La necesidad de tener un conjunto de datos con todas las firmas electromagnéticas de los dispositivo eléctricos, nos obliga a crear uno de manera simulada, ya que no se cuenta con el equipo necesario para obtenerlo de manera real monitoreando la interferencia electromagnética producida en un hogar a lo largo de varios días. Sin embargo cabe destacar

que en el conjunto de datos creado se respetan las firmas electromagnéticas de los aparatos obtenidas en el capítulo anterior, esto con el fin de que los conjuntos de datos sea lo mas cercano a lo real.

La creación del conjunto de datos se explica mas a detalle en la sección 4.1, así mismo en la sección 4.2 se explica el procedimiento que se propone para realizar la identificación de que dispositivos eléctricos se encuentran en funcionamiento en cada instante de tiempo, así como un ejemplo numérico.

## 4.1. Creación del Conjunto de datos

Para poder simular el monitoreo de la interferencia electromagnética producida en un hogar y que ésta sea lo mas cercano a la realidad, debemos considerar como base que aunque no se encuentre ningún dispositivo conectado en la red eléctrica, existe por mínima que sea, interferencia electromagnética producida por el exterior (Otro Hogar, un transformador, etc.). Por lo que la interferencia electromagnética base que existe en un hogar la podemos representar con la letra  $b$ , donde  $b \in \mathbb{R}^m$  es un vector y  $m$  es el número de armónicos calculados del espectro en el dominio de la frecuencia.

Además de la interferencia electromagnética base que se produce en los hogares, debemos considerar que la mayoría de dispositivos eléctricos producen interferencia electromagnética que es única y repetible, por lo que podemos utilizar esta como huella o firma electromagnética de cada aparato. Por lo tanto el conjunto de huellas electromagnéticas de los dispositivos lo podemos representar con la letra  $F$ , donde  $F \in \mathbb{R}^{m \times d}$ ,  $d$  es el número de dispositivos eléctricos con los que cuenta el hogar y  $m$  es el número de armónicos del espectro.

Si consideramos que la interferencia electromagnética es aditiva cada que se agrega un dispositivo a la red eléctrica como se menciona en [Gupta10], la creación del conjunto de datos se puede expresar de forma general como se muestra en la Ecuación 4.1.

$$H[t] = b + \sum_{i=1}^d F[i]Z[t, i] \quad (4.1)$$

Donde  $H$  representa el conjunto de mediciones de la interferencia electromagnética producida en un hogar en un periodo de tiempo,  $H \in \mathbb{R}^{n \times m}$ ,  $n$  es el número total de muestras capturadas en el periodo de tiempo,  $Z$  es un conjunto de datos binario, que determina si en el

instante de tiempo  $t$  se encuentra en funcionamiento el dispositivo eléctrico  $i$ ,  $Z \in \{0|1\}^{d \times n}$ .

Como se puede apreciar en la Ecuación 4.1, los elementos necesarios para crear el conjunto de datos son principalmente la interferencia electromagnética producida en el hogar cuando no se encuentra ninguna carga eléctrica conectada a la red  $b$ , sumando a esta el conjunto de huellas o firmas electromagnéticas de todos los posibles dispositivos eléctricos que se pueden conectar a la red  $F$ , donde el conjunto  $Z$  determina si en el instante de tiempo  $t$  se encuentra en funcionamiento el dispositivo  $i$ .

El conjunto  $Z$  es el que determina que dispositivos se encuentran en funcionamiento en cada instante de tiempo, por lo que este mismo conjunto puede ser utilizado para validar y medir la precisión de una metodología que realice la desagregación de  $H$ . Para que los conjuntos de datos generados sean lo mas cercano a la realidad se toma como referencia los resultados obtenidos del Capítulo 3.

El ruido electromagnético base utilizado  $b$  es el extraído de los conjuntos de datos utilizados en el capítulo 3.

El conjunto de firmas electromagnéticas de los dispositivos eléctricos utilizados, son los obtenidos como resultado de la alineación y extracción de características presentado en el capítulo anterior.

El conjunto  $Z$  a pesar de que se genera aleatoriamente, esta condicionado, ya que cada dispositivo cuenta con un perfil de consumo para determinar la probabilidad de que un dispositivo se encuentre en funcionamiento dependiendo el horario del día en que se este generando la muestra.

Por ejemplo, el perfil de consumo de un televisor varía ya que generalmente existe mayor probabilidad de que se encuentre en funcionamiento por las mañanas, a la hora de la comida y por la noche, quedando poca probabilidad en el horario restante. Por lo que existe gran probabilidad de que al momento de generar el conjunto de datos, este se encuentre presente en un horario entre 14:00 y 16:00 horas.

En la siguiente sección se propone una metodología que identifica que dispositivos se encuentran en funcionamiento, a partir de una muestra simulada de la interferencia electromagnética que se produce en un hogar.

## 4.2. Desagregación

Como se presentó en el Capítulo 2 el estudio de los sistemas *NILM* tiene varios años bajo investigación, sin tener hasta la fecha resultados satisfactorios, esto se debe en gran medida a que las investigaciones que se han hecho para realizar la desagregación utilizan métodos estocásticos como lo son los algoritmos genéticos, redes neuronales artificiales, algoritmo de recocido simulado, programación genética, entre otros.

Utilizar métodos o algoritmos estocásticos tiene la gran desventaja de que en problemas de clasificación (como lo es el estudio de los sistemas *NILM*) crean un modelo que se ajusta exclusivamente para los datos con los que fue entrenado, por lo que si se utilizan para realizar la desagregación en un hogar, este se tiene que someter a un proceso de entrenamiento en dicho hogar y se ajustara únicamente para clasificar los dispositivos eléctricos que se utilizaron en la etapa de entrenamiento.

Si por alguna razón se compra un nuevo electrodoméstico y se agrega la carga a la red eléctrica, el modelo previamente generado ya no servirá para realizar la desagregación y se tendrá que someter nuevamente a una etapa de entrenamiento para crear un modelo que se ajuste al nuevo conjunto de dispositivos eléctricos. Como es evidente no es factible la implementación de un sistema *NILM* mediante este tipo de análisis, ya que el usuario tendría que estar entrenando el modelo constantemente.

En esta sección se explica una técnica que realiza la desagregación de dispositivos eléctricos que no se ha utilizado en el estudio de los sistemas *NILM*, donde se pretende que se pueda implementar sin importar el lugar donde se aplique o el numero de dispositivos eléctricos con los que cuenta la construcción a supervisar. El único requisito es contar con las huellas o firmas electromagnéticas de cada uno de los dispositivos, esto sería posible si los fabricantes proporcionaran de forma pública la firma electromagnética de cada uno de los dispositivos eléctricos que producen.

La idea que dá origen a la metodología propuesta en esta capítulo se obtiene de [Chvatal83], allí tratan un problema donde a partir de la cantidad de luz que deja pasar una mezcla de sustancias, se debe determinar la cantidad de cada sustancia presente en la mezcla, ya que cada sustancia tiene distinto coeficiente de absorción. Esta idea se puede aplicar a este problema al intentar determinar que dispositivos se encuentran presentes a partir de una medición de la mezcla de interferencias electromagnéticas que se producen en un hogar, tomando en cuenta que cada dispositivo produce una interferencia distinta.

El método que se propone se debe realizar por cada instante de tiempo, es decir, se debe aplicar cada que se toma una muestra de la interferencia electromagnética producida en el hogar a supervisar. Para el correcto funcionamiento de la metodología propuesta es indispensable contar con todas las firmas electromagnéticas de los dispositivos eléctricos que se encuentran en la construcción a monitorear.

Entonces definamos que  $F$  es el conjunto de huellas o firmas electromagnéticas de los dispositivos eléctricos con los que cuenta el hogar, donde  $F \in \mathbb{R}^{m \times d}$ ,  $m$  es la cantidad de armónicos de las firmas de los dispositivos y  $d$  es la cantidad dispositivos eléctricos. Además  $h$  es una muestra capturada de la interferencia electromagnética producida en el hogar en un instante de tiempo y  $b$  es el ruido electromagnético "base" que se mide cuando ningún dispositivo se encuentra conectado a la red. Donde  $h, b \in \mathbb{R}^m$  son vectores que contiene  $m$  armónicos.

Lo primero que se recomienda hacer es restar la interferencia electromagnética base que existe en el hogar, como se muestra en la Ecuación 4.2, esto con el fin de tener un solo agregado de las interferencias que producen el conjunto de dispositivos eléctricos.

$$h^* = h - b \quad (4.2)$$

Con esto el problema se puede plantear como el sistema lineal de ecuaciones mostrado en la Ecuación 4.3, donde  $x$  es el vector solución al problema que determinara si el dispositivo  $d$  se encuentra presente o no en la mezcla de interferencias electromagnéticas  $h^*$ .

$$\begin{aligned} F_{1,1}x_1 + F_{1,2}x_2 + F_{1,3}x_3 + \dots + F_{1,d}x_d &= h_1^* \\ F_{2,1}x_1 + F_{2,2}x_2 + F_{2,3}x_3 + \dots + F_{2,d}x_d &= h_2^* \\ F_{3,1}x_1 + F_{3,2}x_2 + F_{3,3}x_3 + \dots + F_{3,d}x_d &= h_3^* \\ &\dots \\ F_{m,1}x_1 + F_{m,2}x_2 + F_{m,3}x_3 + \dots + F_{m,d}x_d &= h_m^* \end{aligned} \quad (4.3)$$

Si el sistema de ecuaciones lo expresamos en forma matricial como se observa en la Ecuación 4.4, obtenemos un sistema de la forma  $Ax = b$  como se muestra en la Ecuación 4.5.

$$\begin{bmatrix} F_{1,1} & F_{1,2} & F_{1,3} & \dots & F_{1,d} \\ F_{2,1} & F_{2,2} & F_{2,3} & \dots & F_{2,d} \\ F_{3,1} & F_{3,2} & F_{3,3} & \dots & F_{3,d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{m,1} & F_{m,2} & F_{m,3} & \dots & F_{m,d} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} h_1^* \\ h_2^* \\ h_3^* \\ \vdots \\ h_m^* \end{bmatrix} \quad (4.4)$$

$$Fx = h^* \quad (4.5)$$

Una de las principales problemáticas para resolver este sistema de ecuaciones es que el sistema tiene  $m$  ecuaciones con  $d$  incógnitas, donde  $m$  es mucho mayor que  $d$ , lo que provoca generalmente que el sistema no tenga solución. Prueba de ello es que si se tomaran solo  $d$  de las  $m$  ecuaciones y se resolviera el sistema, la solución sería distinta de la que se produce si se resuelve el sistema tomando otras  $d$  ecuaciones distintas de las primeras.

Al no tener solución exacta se debe calcular una solución aproximada. La mejor aproximación a la solución del sistema será la que minimice el error como se muestra en la Ecuación 4.6, tomando como medida de distancia la euclidiana, llamada aproximación  $L_2$ .

$$\text{minimiza} \quad \sum_{i=1}^m \left( h_i^* - \sum_{j=1}^d F_{ij} x_j \right)^2 \quad (4.6)$$

La solución a la minimización como se muestran en [Chvatal83] da como resultado la Ecuación 4.7.

$$\sum_{j=1}^d \left( \sum_{i=1}^m F_{ik} F_{ij} \right) x_j = \sum_{i=1}^m h_i^* F_{ik} \quad (k = 1, 2, \dots, d) \quad (4.7)$$

El resultado de la Ecuación 4.7 nos genera un nuevo sistema de ecuaciones pero ahora solo con  $d$  ecuaciones y  $d$  incógnitas como se muestra en las Ecuaciones 4.8 y 4.9.

$$\begin{bmatrix} F_{1,1}^* & F_{1,2}^* & F_{1,3}^* & \dots & F_{1,d}^* \\ F_{2,1}^* & F_{2,2}^* & F_{2,3}^* & \dots & F_{2,d}^* \\ F_{3,1}^* & F_{3,2}^* & F_{3,3}^* & \dots & F_{3,d}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{d,1}^* & F_{d,2}^* & F_{d,3}^* & \dots & F_{d,d}^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} h_1^{**} \\ h_2^{**} \\ h_3^{**} \\ \vdots \\ h_d^{**} \end{bmatrix} \quad (4.8)$$

$$F^* x = h^{**} \quad (4.9)$$

La solución de la Ecuación 4.9 se puede calcular directamente al aplicar algún método que resuelva el sistema de ecuaciones lineales, por ejemplo: por sustitución y eliminación de Gauss-Jordan, entre otros. El resultado obtenido será la mejor aproximación  $L_2$  a la solución de la Ecuación 4.5 planteada desde un inicio.

Cabe destacar que como se trata de un problema de clasificación, solo se va a reportar si el dispositivo  $d$  se encuentra en funcionamiento o no. Por lo que las diferentes combinaciones que se pueden tener al contar con  $d$  dispositivos eléctricos distintos son  $2^d$  posibles combinaciones.

Entonces el espacio de búsqueda serán los vértices del hipercubo unitario de  $d$  dimensiones. Si  $d$  es igual a 3 ( $d_x, d_y, d_z$ ) los posibles resultados se encuentran en los vértices que forma el cubo, como se muestra en la figura 4.1.

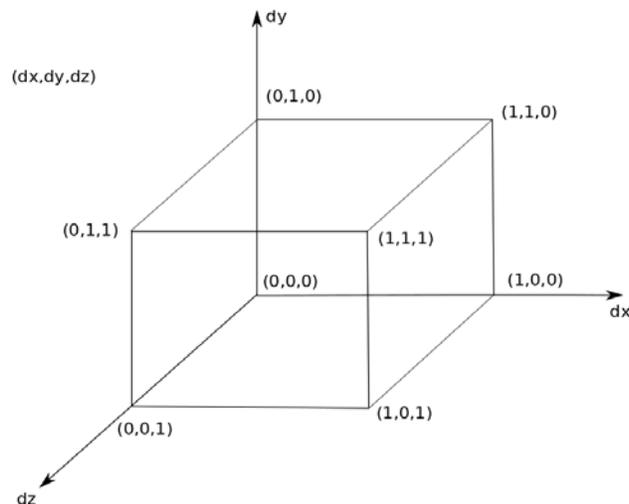


Figura 4.1: Cubo que forma el espacio de búsqueda cuando se cuenta con solo 3 dispositivos eléctricos. El espacio de búsqueda está formado por las  $2^3$  combinaciones que existen.

El último paso en la metodología propuesta consiste en que al momento de resolver la ecuación 4.9, al ser una solución aproximada esta nos proporcionará las coordenadas cercanas a la solución exacta. Aplicando un simple redondeo estaremos escogiendo el vértice más cercano como solución (la combinación de dispositivos en  $F$  que genera el agregado de interferencia electromagnética  $h$ ), en la Figura 4.2 se muestra esto a más detalle.

Siguiendo este procedimiento podremos informar qué dispositivos se encuentran en determinado instante de tiempo, recordando que estos pasos se deben realizar por cada muestra de la interferencia electromagnética medida en el hogar a supervisar.

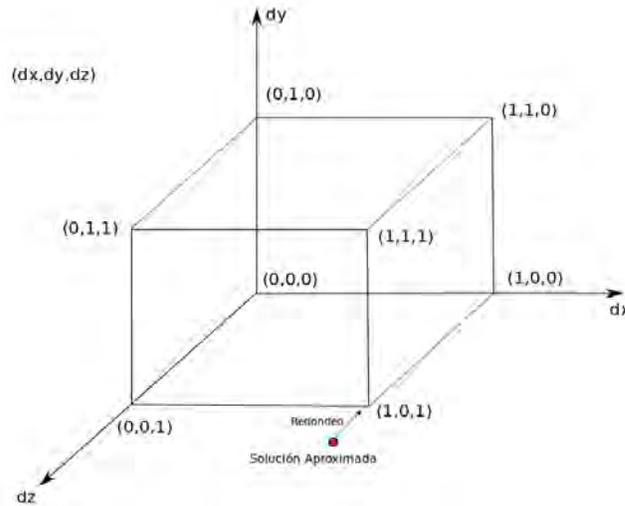


Figura 4.2: Gráfica que ejemplifica donde se ubica la solución aproximada que arroja el método y con un redondeo se acerca al vértice mas cercano en el espacio de búsqueda.

#### 4.2.1. Ejemplo Numérico

Con el fin mostrar de forma mas clara lo que hace la metodología propuesta en este capítulo para identificar los dispositivos eléctricos que se encuentran conectados a la red eléctrica, a partir de una muestra de la medición de interferencia electromagnética que generan en conjunto, a continuación se presenta un ejemplo con datos numéricos.

Supongamos que el hogar a supervisar solo cuenta con tres aparatos eléctricos, un refrigerador, un televisor y una lampara, de los cuales tenemos su firma electromagnética que consta de los primeros 12 armónicos de su respectivo espectrograma, como se muestra en la Tabla 4.1.

Además se tienen tres muestras diferentes del ruido electromagnético que se produce en el hogar combinando los distintos dispositivos como se muestra en la Tabla 4.2. Cabe destacar que estas muestras están contaminadas por ruido, ya que no es la combinación exacta de los ruidos electromagnéticos que producen.

Si tomamos la muestra marcada como Refrigerador/Televisor, como se menciona en la Ecuación 4.2 es restar a la muestra el ruido base que existe en el hogar.

$$h^* = h - b$$

$$h^* = [153, 74, 178, 83, 76, 70, 55, 93, 97, 95, 120, 92] - [2, 14, 12, 9, 13, 2, 20, 15, 7, 3, 17, 12]$$

Tabla 4.1: Firmas electromagnéticas de el refrigerador, televisor y lampara, así como el ruido electromagnético base que existe en el hogar cuando no se encuentra ningún dispositivo conectado.

<b>Armónico</b>	<b>Refrigerador</b>	<b>Televisor</b>	<b>Lampara</b>	<b>Base</b>
1	82	75	80	2
2	54	7	9	14
3	84	73	7	12
4	48	22	31	9
5	59	6	98	13
6	58	13	7	2
7	37	2	22	20
8	61	9	94	15
9	81	11	24	7
10	47	41	69	3
11	35	70	85	17
12	50	27	36	12

Tabla 4.2: Muestra que se producen al combinar las interferencias que producen los distintos dispositivos eléctricos.

<b>Armónico</b>	<b>Refrigerador Televisor</b>	<b>Televisor Lampara</b>	<b>Refrigerador Lampara</b>
1	153	151	158
2	74	29	76
3	178	101	112
4	83	66	92
5	76	115	168
6	70	19	64
7	55	40	75
8	93	126	178
9	97	40	110
10	95	117	123
11	120	170	135
12	92	78	101

$$h^* = [151, 60, 166, 74, 63, 68, 35, 78, 90, 92, 103, 80]$$

Una vez extraído el ruido electromagnético base de la muestra, obtenemos un agregado de la interferencia electromagnética que producen los posibles dispositivos que se encuentran conectados en ese instante de tiempo, a partir de esto y con el conjunto de firmas electromagnéticas podemos formular el problema como un sistema de ecuaciones lineales como se muestra en la Ecuación 4.3.

$$82x_1 + 75x_2 + 80x_3 = 151$$

$$54x_1 + 7x_2 + 9x_3 = 60$$

$$84x_1 + 73x_2 + 7x_3 = 166$$

$$48x_1 + 22x_2 + 31x_3 = 74$$

$$59x_1 + 6x_2 + 98x_3 = 63$$

$$58x_1 + 13x_2 + 7x_3 = 68$$

$$37x_1 + 2x_2 + 22x_3 = 35$$

$$61x_1 + 9x_2 + 94x_3 = 78$$

$$81x_1 + 11x_2 + 24x_3 = 90$$

$$47x_1 + 41x_2 + 69x_3 = 92$$

$$35x_1 + 70x_2 + 85x_3 = 103$$

$$50x_1 + 27x_2 + 36x_3 = 80$$

Con esto obtenemos un sistema de 12 ecuaciones y 3 incógnitas, por lo que el sistema no tiene solución exacta. Al no tener solución, se tiene que encontrar la que minimice el error para obtener una solución aproximada, para esto se utiliza la Ecuación 4.7 que nos arroja un sistema de ecuaciones como el siguiente.

$$43430x_1 + 22065x_2 + 31820x_3 = 66051$$

$$22065x_1 + 19208x_2 + 18840x_3 = 41657$$

$$31820x_1 + 18840x_2 + 40322x_3 = 50971$$

El sistema de ecuaciones obtenido ya se puede resolver directamente aplicando

algún método como la eliminación de Gauss-Jordan, método gráfico, algoritmos numéricos, etc. El resultado obtenido es el siguiente.

$$x_1 = 1.01090422$$

$$x_2 = 1.01539669$$

$$x_3 = -0.00808606$$

Donde  $x_1, x_2, x_3$  corresponde al refrigerador, televisor y lámpara respectivamente, si el resultado lo vemos de manera gráfica como se muestra en la Figura 4.3, observamos que el resultado se encuentra cerca de una de las posibles soluciones (vértices), por lo que aplicando el último paso que es redondear las cifras obtenemos el resultado [110] que nos indica que en ese instante de tiempo el refrigerador y el televisor se encontraban en funcionamiento.

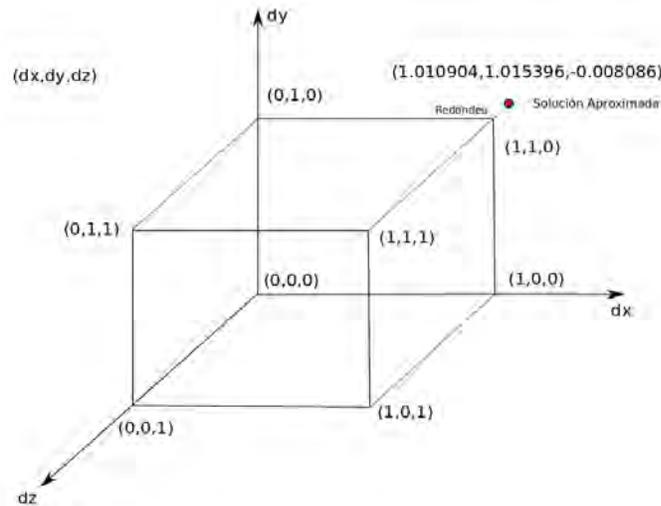


Figura 4.3: Ubicación de la solución en el espacio de búsqueda.

En la Tabla 4.3 se muestran los resultados obtenidos al aplicar la metodología propuesta en este capítulo a las tres muestras de este ejemplo. Como se puede apreciar, la metodología en este ejemplo es capaz de informar que dispositivos eléctricos se encuentran conectados en cada instante de tiempo. Por lo que en el Capítulo 5 se realizan pruebas mas exhaustivas para verificar la precisión y robustez del método.

Tabla 4.3: Resultado de aplicar la metodología a las tres muestras de ejemplo.

	<b>Refrigerador</b>	<b>Televisor</b>	<b>Refrigerador</b>
	<b>Televisor</b>	<b>Lampara</b>	<b>Lampara</b>
$x_1$	1	0	1
$x_2$	1	1	0
$x_3$	0	1	1

### 4.3. Resumen de Capítulo

A pesar de que en el capítulo anterior se realiza la alineación de eventos y extracción de características de los dispositivos eléctricos etiquetados en los conjuntos de datos, estos no se pudieron utilizar para probar la metodología propuesta en este capítulo, esto debido en gran parte a la cantidad de dispositivos que no fueron etiquetados y por consecuencia no se pudo extraer su firma electromagnética, ya que la metodología propuesta en este capítulo requiere del conjunto de firmas electromagnéticas de todos los dispositivos para su correcto funcionamiento.

Es por esto que existe la necesidad de crear un conjunto de datos de manera simulada, como se presento en la sección 4.1. Para que los conjuntos de datos generados sean lo mas cercano a la realidad se respetó lo siguiente al momento de generarlos:

- El ruido electromagnético base utilizado  $b$  es el extraído de los conjuntos de datos utilizados en el capítulo 3.
- El conjunto de firmas electromagnéticas de los dispositivos eléctricos utilizados, son los obtenidos como resultado de la alineación y extracción de características presentado en el capítulo 3.
- Dependiendo el dispositivo eléctrico, cada uno cuenta con un perfil de consumo para determinar la probabilidad de que un dispositivo se encuentre en funcionamiento dependiendo el horario del día en que se este generando la muestra.

Una vez generados los conjuntos de datos y contando con todas las firmas electromagnéticas de los dispositivos eléctricos, en la sección 4.2 se presenta una metodología para poder determinar los dispositivos que se encuentran en funcionamiento en cada instante de tiempo. A grandes rasgos el procedimiento se divide en los siguiente pasos, destacando que la metodología se ejecuta por cada una de las muestras en los conjuntos de datos:

- 
- Restar el ruido electromagnético base  $b$  a la muestra a desagregar. Para obtener un solo agregado de la interferencia electromagnética que producen los dispositivos conectados en ese instante de tiempo.
  - En base al conjunto de firmas electromagnéticas de los dispositivos y la muestra a procesar, formular el problema como un conjunto de ecuaciones lineales.
  - El sistema al no tener solución exacta, transformar el sistema de ecuaciones a otro que si tenga solución directa y nos proporcione una aproximación a los valores que resuelven el problema.
  - Redondear la solución para trasladarse al vértice mas cercano del espacio de búsqueda del problema.

En siguiente capítulo se presentan algunas pruebas y sus resultados para poder medir la precisión y robustez de la metodología aquí propuesta.



## Capítulo 5

# Pruebas y Resultados

Hasta el momento se ha descrito la metodología propuesta para realizar la identificación de los dispositivos que se encuentran conectados por cada instante de tiempo, así como la explicación del procedimiento utilizado para generar los conjuntos de datos de manera artificial para probar dicha metodología.

A continuación se muestran los resultados obtenidos, así como a las pruebas que se sometió para probar la eficacia de la metodología.

Este capítulo se divide de la siguiente manera, primero se muestran los parámetros utilizados para generar los distintos conjuntos de datos utilizados, después se exponen las pruebas realizadas contaminando los datos gradualmente con ruido blanco gaussiano para verificar lo robusto de la metodología propuesta, enseguida se exponen y discuten los resultados obtenidos, por ultimo se concluye los presentado en el capítulo.

### 5.1. Parámetros para generar los conjuntos de datos

Al acotar este trabajo a solo simulación por la limitante de no contar con un conjunto de datos obtenido de forma real al monitorear la interferencia electromagnética en un hogar, se crea uno de forma artificial.

Para que estos conjuntos de datos no sean generados totalmente de forma aleatoria se utilizan como conjunto de firmas electromagnéticas los obtenidos como resultado del Capítulo 3. De los cuales de 4 hogares distintos (H1, H2, H3 y H4) se obtienen 36 dispositivos eléctricos por cada hogar.

En las Tablas 5.1,5.2,5.3 y 5.4 se muestran los nombres de los dispositivos eléctricos

que se utilizaron para generar los conjuntos de datos artificiales de acuerdo al hogar que corresponden, donde se puede apreciar que en su mayoría se trata de lamparas ubicadas en distintas partes del hogar.

Tabla 5.1: Lista de dispositivos utilizados en el hogar 1 (H1).

1.BR1 Closet Light	13.Downstairs Bathroom Fan Light	25.Master BA Lights
2.BR1 DVD	14.Downstairs Bathroom Lights	26.Master BR Entry Lights
3.BR1 Lights	15.Downstairs Hallway Lights	27.Master BR Lights
4.BR2 Closet Light	16.Dryer	28.Master BR Walk-in Closet L
5.BR2 Lights	17.GR LCD TV	29.Outside Front Door Lights
6.Backyard Light	18.GR Lights	30.Outside Over Garage Lights
7.Balcony Lights	19.GR PS4	31.Over Sink Light
8.Central Vacuum	20.Garage Door Opener	32.Portable Vacuum
9.Coffee Maker	21.Garage Lights	33.Powder Rm Lights
10.Dining Room Lights	22.Kitchen Lights	34.Stairway Lights
11.Dishwasher	23.Kitchen Under Cabinet Ligh	35.Toaster
12.Downstairs Bathroom Fan	24.Master BA Heat Lamp	36.Trash Compactor

Tabla 5.2: Lista de dispositivos utilizados en el hogar 2 (H2).

1.Bathroom 1 Lights	13.Garbage Disposal	25.Microwave
2.Bathroom 2 Lights	14.Hair Dryer	26.Office Lights
3.Bedroom 1 Lights	15.Hallway Lights	27.Outdoor Lights
4.Bedroom 2 Lights	16.Kitchen Lights	28.Phone Charger
5.Coffee Maker	17.Kitchen Table Lights	29.Printer
6.Computer	18.Laptop Charger	30.Refrigerator
7.Crockpot	19.Laundry Room Lights	31.Stairway Lights
8.Dining Room Lights	20.Living Room Lights	32.Straightening Iron
9.Dishwasher	21.Master Bath Fan	33.TV-DVR
10.Dryer	22.Master Bath Lights	34.Tea Kettle
11.Front Hall Lights	23.Master Bedroom Lights	35.Toaster
12.Front Room Lights	24.Master Closet Lights	36.Vacuum

Estos dispositivos son extraídos de los conjuntos de datos que se utilizaron en el Capítulo 3, los cuales se encontraban etiquetados, por lo cual por cada uno de los dispositivos se pudo extraer su firma electromagnética, las cuales son utilizadas para realizar una combinación lineal y poder generar los conjuntos de datos artificiales.

En la Figura 5.1(a) se muestra de manera gráfica la firma o huella electromagnética de uno de los 36 dispositivos de la casa 1 (H1), en este caso es el dispositivo número 30 que son las luces que se encuentran fuera de la cochera, se puede apreciar que la interferencia que produce afecta a muchos armónicos a lo largo del espectro de frecuencias.

Tabla 5.3: Lista de dispositivos utilizados en el hogar 3 (H3).

1.Back Porch Lights	13.Foyer Lights	25.Master Bath Fan
2.Bedroom 1 LCD TV	14.Front Porch Lights	26.Master Bath Lights
3.Bedroom 1 Lights	15.Garage Door Opener	27.Master Bluera-DVD
4.Bedroom 2 Lights	16.Garage Lights	28.Master Closet Lights
5.Bonus Room Bluera-DVD	17.Garbage Disposal	29.Master LCD TV-DVR
6.Bonus Room LED TV	18.Guest Bath Fan	30.Master Lights
7.Bonus Room Lights	19.Guest Bath Lights	31.Microwave
8.Bonus Room Wii	20.Hair Dryer	32.Office Lights
9.Computer	21.Kitchen Lights	33.Oven
10.Dining Room Lights	22.Laundry Room Lights	34.Powder Room Lights
11.Dishwasher	23.Living Room Audio-DVR-TV	35.Toaster
12.Dryer	24.Living Room Lights	36.Upstairs Hallway Lights

Tabla 5.4: Lista de dispositivos utilizados en el hogar 4 (H4).

1.Apple Macbook Pro 13	13.Electric Furnace	25.Microwave
2.Apple Macbook Pro 15	14.Entry light	26.Mixer
3.Bedroom Lamp 1	15.Forced-air Heater	27.Oven
4.Bedroom Lamps 2	16.HP Elitebook Laptop	28.PC
5.Bose iPhone Dock	17.Hairdryer	29.PS3
6.Bread Maker	18.Hallway overhead Light	30.Room Baseboard Heater
7.Closet Light	19.Kettle	31.Sandwich Maker
8.Computer Desk Lamp	20.Kitchen Counter Lights	32.Stove
9.Deck Light	21.Kitchen Lights with Dimmer	33.Toaster Oven
10.Den Baseboard Heater	22.LivingRoom overhead halogen	34.Toilet Exhaust
11.Den Overhead Light	23.Livingroom Lamp 1	35.Toilet Halogen
12.Dishwasher	24.Livingroom Lamp 2	36.Vaccum

De manera similar con el fin de ilustrar las huellas electromagnéticas extraídas y utilizadas para generar los conjuntos de datos, en la Figura 5.1(b) se muestra la interferencia electromagnética producida por las luces que se encuentran en la cocina en la casa 2 (H2).

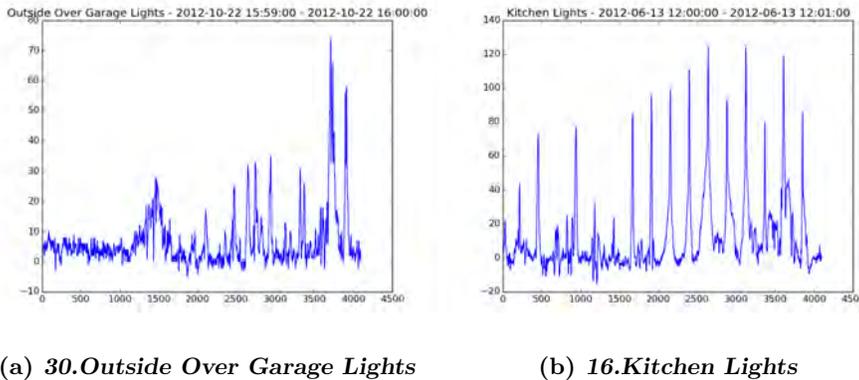


Figura 5.1: (a) Firma o huella electromagnética que produce las lámparas que se encuentran fuera de la cochera de la casa 1 (H1). (b) Firma o huella electromagnética que produce las luces de la cocina de la casa 2 (H2).

La interferencia electromagnética de uno de los dispositivos de las casas 3 y 4 se muestra en la Figura 5.2(a) y 5.2(b) respectivamente, donde la firma que producen es fácilmente distinguible una de la otra.

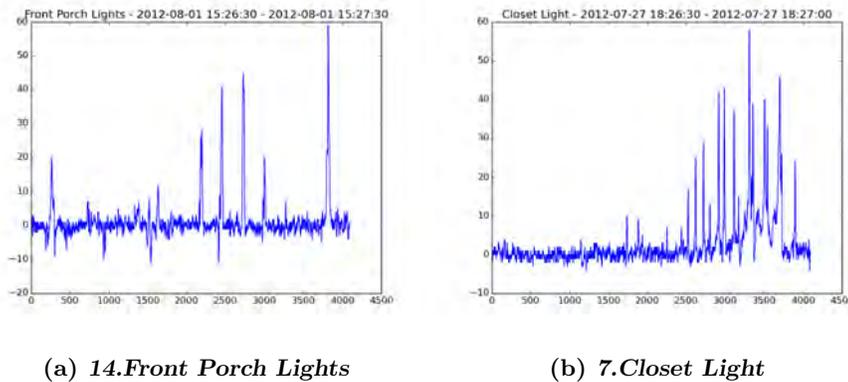


Figura 5.2: (a) Firma o huella electromagnética que produce las luces de la entrada de la casa 3 (H3). (b) Firma o huella electromagnética que produce las luces del armario de la casa 4 (H4).

Además de contar con las firmas electromagnéticas de los distintos dispositivos,

para que los conjuntos de datos generados se acerquen a la realidad se elaboro un perfil de consumo por cada uno de los dispositivos. Por ejemplo, el televisor tiene mas probabilidad que se encuentre encendido entre las 20:00 y las 22:00 horas que entre las 01:00 y las 05:00, por lo que en los horarios con mayor probabilidad de acuerdo al dispositivo es posible que se tome como un dispositivo en funcionamiento al momento de generar el conjunto de datos.

Con los parámetros y consideraciones expuestas se procede a generar los conjuntos de datos con el procedimiento mostrado en la Sección 4.1.

## 5.2. Pruebas

Una vez obtenidos los conjuntos de datos se procede a probar la eficacia y robustez de la metodología propuesta en este trabajo para realizar la identificación de los dispositivos conectados en un instante de tiempo. Para esto contaminaremos cada una de las muestras de los conjuntos de datos con ruido gaussiano (distribución normal) con media cero ( $\mu = 0$ ) y variando la desviación estándar ( $\sigma$ ), donde  $\sigma$  determinara la amplitud del ruido.

Se proponen dos formas distintas de contaminar las muestra mediante ruido blanco gaussiano para probar como se comporta la metodología ante el ruido. La primera es contaminar de manera individual cada una de las componentes que conforman la muestra (cada muestra consta de 4096 componentes), tomando como  $\sigma$  el valor de la muestra que se quiere contaminar, es decir, el numero generado por la distribución normal sera directamente proporcional a la amplitud de la componente a afectar. Este tipo de ruido lo denominaremos como ruido *Proporcional*.

En la Figura 5.3 se muestra de manera gráfica con el fin de ejemplificar, como por cada una de las componentes varía el ruido que se agrega a la componente, a las de mayor amplitud el ruido es mayor en comparación con las de menor amplitud.

Este tipo de ruido generado tiene como objetivo que sea proporcional a cada una de las componentes y observar como se comporta la metodología al ir aumentando el ruido gradualmente.

Por otra parte, la segunda forma que se propone para contaminar las muestras es que en lugar de seleccionar la desviación estándar por cada unas de las componentes, es que la desviación estándar sea igual a la media de toda la muestra, con esto las componentes son contaminadas de igual forma independientemente si es grande o chica la amplitud de la componente. Este tipo de ruido lo denominaremos como ruido *Promedio*. Esto se ejemplifica

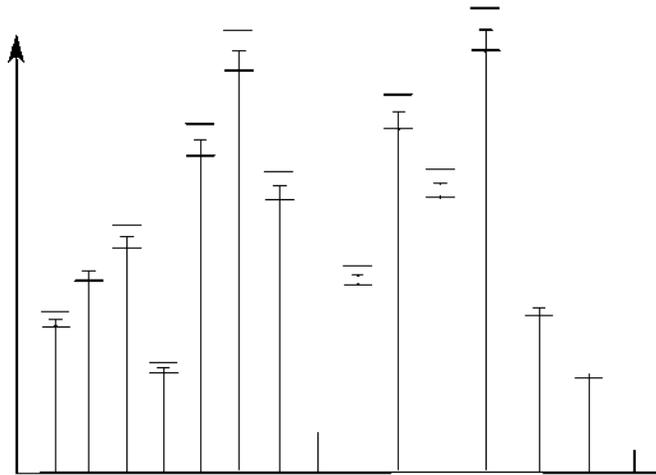


Figura 5.3: Gráfica que ejemplifica el ruido generado para cada una de las componentes, entre mayor sea la componente mayor será el ruido generado.

de manera gráfica en la Figura 5.4.

Cabe mencionar que en ambos casos la desviación estándar comienza en cero y va aumentando gradualmente hasta convertirse en los valores de los casos ya propuestos, esto con el fin de observar como se comporta la metodología conforme se va aumentando el ruido.

### 5.3. Resultados

Una vez generados los conjuntos de datos, se procede a realizar la identificación de los dispositivos eléctricos que se encuentran en funcionamiento en cada instante de tiempo, para lo cual se hace uso de la metodología propuesta en este trabajo, obteniendo los resultados presentados en este capítulo.

En la Tabla 5.5 se muestran los resultados de aplicar la metodología al conjunto de datos generado con los 36 dispositivos de la casa numero 1 (H1), donde las columnas de la tabla se interpretan de la siguiente forma:

- Días Simulados: El número de días que abarca los conjuntos de datos generados.
- % de  $\sigma$ : Porcentaje utilizado de la desviación estándar para generar el ruido con el que se contaminan los conjuntos de datos. 0% significa que los conjuntos de datos no

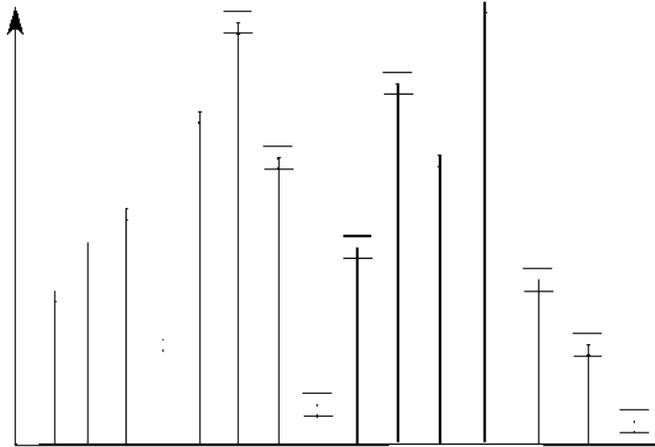


Figura 5.4: Gráfica que ejemplifica el ruido generado tomando como desviación estándar la media de toda la muestra a contaminar.

contienen ruido.

- Casa: Es la casa donde se probó la metodología.
- Numero de Dispositivos: La cantidad de dispositivos con los que se generaron los conjuntos de datos.
- Num. errores de tantos posibles: El numero de veces que se predijo mal la presencia o ausencia de un dispositivo en la señal.
- Precisión: Porcentaje del número de veces que se clasificó de forma correcta. (Total de Eventos - Número de Errores) / Total de eventos.

En los resultados de la Tabla 5.5 se puede observar que cuando los conjuntos de datos no contienen ruido ( $\% \text{ de } \sigma \text{ igual a } 0$ ), la metodología identifica de forma correcta los dispositivos que se encuentran en funcionamiento en todas las muestras, obteniendo una precisión del 100%. Esto debido a que la metodología plantea el problema como un sistema de ecuaciones y al no contener ruido, el sistema de ecuaciones tiene una solución exacta y única.

Al aumentar el ruido se puede observar que la efectividad de la metodología comienza a descender, esto debido a que cuando aumenta el ruido, la metodología puede

Tabla 5.5: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 1 (H1) contaminados con ruido gaussiano del tipo que denominamos *Proporcional*.

Días Simulados	% de $\sigma$	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H1	36	0 de 3110400	100.0
1	10	H1	36	0 de 3110400	100.0
1	20	H1	36	41 de 3110400	99.9986818416
1	30	H1	36	893 de 3110400	99.9712898663
1	40	H1	36	5038 de 3110400	99.8380272634
1	50	H1	36	14507 de 3110400	99.533596965
1	60	H1	36	28900 de 3110400	99.0708590535
1	70	H1	36	46996 de 3110400	98.48906893
1	80	H1	36	67539 de 3110400	97.8286072531
1	90	H1	36	89383 de 3110400	97.1263181584
1	100	H1	36	111324 de 3110400	96.4209104938

interpretar que otro dispositivo se encuentra en funcionamiento aunque no sea así. Sin embargo, en el peor de los casos probado cuando la desviación estándar del ruido generado es igual a la amplitud de la componente a contaminar, sigue teniendo una muy buena precisión con el 96.42 % clasificado correctamente, lo cual demuestra que la metodología es bastante robusta ante el ruido que pueda existir.

En las Tablas 5.6, 5.7 y 5.8 se muestra la efectividad de la metodología con otros conjuntos de datos generados con distintos dispositivo de las casas 2, 3 y 4 (H2, H3 y H4) respectivamente. Al igual que los resultados de la Tabla 5.5 se observa la metodología tiene una precisión del 100 % ante señales sin ruido o muy poco ruido, y va perdiendo precisión conforme aumenta el ruido, aunque en el peor de los casos probado (% de  $\sigma$  igual a 100) conserva una precisión aproximada del 96 %, lo cual demuestra que la metodología propuesta es bastante robusta al ruido en las señales.

Los resultados de la segunda prueba propuesta se muestran en las Tablas 5.9, 5.10, 5.11 y 5.12, donde  $\sigma$  es igual a la media de cada muestra a contaminar con ruido. En estos resultados se puede observar que la metodología contiene una precisión mayor cuando el ruido va aumentando en comparación con los resultados de la prueba con ruido que denominamos *Proporcional*.

Los resultados de las Tablas 5.9, 5.10, 5.11 y 5.12 se deben a que las firmas electromagnéticas de los dispositivos que están compuestas de 4096 componentes, en su mayoría son componentes de baja amplitud, ya que la interferencia electromagnética que producen

Tabla 5.6: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 2 (H2) contaminados con ruido gaussiano del tipo que denominamos *Proporcional*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H2	36	0 de 3110400	100.0
1	10	H2	36	0 de 3110400	100.0
1	20	H2	36	131 de 3110400	99.995788323
1	30	H2	36	1541 de 3110400	99.9504565329
1	40	H2	36	5824 de 3110400	99.8127572016
1	50	H2	36	13421 de 3110400	99.5685120885
1	60	H2	36	24009 de 3110400	99.2281057099
1	70	H2	36	37264 de 3110400	98.8019547325
1	80	H2	36	52597 de 3110400	98.3089956276
1	90	H2	36	69523 de 3110400	97.7648212449
1	100	H2	36	87747 de 3110400	97.1789158951

los dispositivos eléctricos solo excitan un pequeño numero de armónicos en comparación con toda la señal. Entonces la media de cada señal será pequeña y por consiguiente el ruido generado también será menor, lo que provoca una mejor precisión en la identificación de dispositivos que se encuentran en operación.

## 5.4. Resumen de Capítulo

Lo primero que se planteó en este capítulo para probar la metodología propuesta, son los parámetros utilizados para generar los conjuntos de datos de prueba. Donde se crean 4 conjuntos de datos, simulando que se obtuvieron de monitorear 4 hogares distintos a lo largo de 1 día cada uno de ellos.

Cada uno de los 4 hogares cuenta con 36 dispositivos eléctricos distintos, donde por cada dispositivos se cuenta con su firma electromagnética que producen. Para formar los conjuntos de datos por cada instante de tiempo (1 segundo) se realiza una combinación lineal de las firmas electromagnéticas de los dispositivos siempre y cuando estos se encuentre en funcionamiento.

Para saber si un dispositivo se encuentra en funcionamiento en un instante de tiempo o no, cada dispositivo cuenta con un perfil de consumo (creado de la experiencia de observar los horarios que generalmente se encuentran en funcionamiento los dispositivos), donde si por ejemplo, se trata de un televisor, es mucho mas probable que este se encuentre

Tabla 5.7: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 3 (H3) contaminados con ruido gaussiano del tipo que denominamos *Proporcional*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H3	36	0 de 3110400	100.0
1	10	H3	36	0 de 3110400	100.0
1	20	H3	36	1 de 3110400	99.9999678498
1	30	H3	36	43 de 3110400	99.9986175412
1	40	H3	36	445 de 3110400	99.9856931584
1	50	H3	36	1928 de 3110400	99.9380144033
1	60	H3	36	5048 de 3110400	99.8377057613
1	70	H3	36	10474 de 3110400	99.6632587449
1	80	H3	36	18229 de 3110400	99.4139338992
1	90	H3	36	28394 de 3110400	99.0871270576
1	100	H3	36	40865 de 3110400	98.6861818416

Tabla 5.8: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 4 (H4) contaminados con ruido gaussiano del tipo que denominamos *Proporcional*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H4	36	0 de 3110400	100.0
1	10	H4	36	3 de 3110400	99.9999035494
1	20	H4	36	1005 de 3110400	99.9676890432
1	30	H4	36	7583 de 3110400	99.7562049897
1	40	H4	36	22068 de 3110400	99.2905092593
1	50	H4	36	43272 de 3110400	98.6087962963
1	60	H4	36	68455 de 3110400	97.7991576646
1	70	H4	36	96618 de 3110400	96.8937114198
1	80	H4	36	125854 de 3110400	95.9537680041
1	90	H4	36	155812 de 3110400	94.9906121399
1	100	H4	36	186162 de 3110400	94.0148533951

Tabla 5.9: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 1 (H1) contaminados con ruido gaussiano del tipo que denominamos *Promedio*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H1	36	0 de 3110400	100.0
1	10	H1	36	0 de 3110400	100.0
1	20	H1	36	1 de 3110400	99.9999678498
1	30	H1	36	47 de 3110400	99.9984889403
1	40	H1	36	390 de 3110400	99.9874614198
1	50	H1	36	1727 de 3110400	99.9444765947
1	60	H1	36	4755 de 3110400	99.8471257716
1	70	H1	36	9764 de 3110400	99.6860853909
1	80	H1	36	16803 de 3110400	99.4597800926
1	90	H1	36	25532 de 3110400	99.1791409465
1	100	H1	36	35587 de 3110400	98.8558706276

Tabla 5.10: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 2 (H2) contaminados con ruido gaussiano del tipo que denominamos *Promedio*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H2	36	0 de 3110400	100.0
1	10	H2	36	0 de 3110400	100.0
1	20	H2	36	6 de 3110400	99.9998070988
1	30	H2	36	316 de 3110400	99.989840535
1	40	H2	36	1731 de 3110400	99.9443479938
1	50	H2	36	4898 de 3110400	99.8425282922
1	60	H2	36	9837 de 3110400	99.6837384259
1	70	H2	36	16702 de 3110400	99.4630272634
1	80	H2	36	24992 de 3110400	99.1965020576
1	90	H2	36	34424 de 3110400	98.8932613169
1	100	H2	36	45041 de 3110400	98.5519225823

Tabla 5.11: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 3 (H3) contaminados con ruido gaussiano del tipo que denominamos *Promedio*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H3	36	0 de 3110400	100.0
1	10	H3	36	0 de 3110400	100.0
1	20	H3	36	1 de 3110400	99.9999678498
1	30	H3	36	22 de 3110400	99.9992926955
1	40	H3	36	251 de 3110400	99.9919302984
1	50	H3	36	938 de 3110400	99.969843107
1	60	H3	36	2282 de 3110400	99.9266332305
1	70	H3	36	4477 de 3110400	99.8560635288
1	80	H3	36	7408 de 3110400	99.7618312757
1	90	H3	36	11080 de 3110400	99.6437757202
1	100	H3	36	15578 de 3110400	99.4991640947

Tabla 5.12: Resultados obtenidos de aplicar la metodología al conjunto de datos generado con los dispositivos de la casa 4 (H4) contaminados con ruido gaussiano del tipo que denominamos *Promedio*.

Días Simulados	% de sigma	Casa	Numero de Dispositivos	Num. errores de tantos posibles	Precisión %
1	0	H4	36	0 de 3110400	100.0
1	10	H4	36	0 de 3110400	100.0
1	20	H4	36	103 de 3110400	99.9966885288
1	30	H4	36	1591 de 3110400	99.9488490226
1	40	H4	36	6579 de 3110400	99.7884837963
1	50	H4	36	14883 de 3110400	99.5215084877
1	60	H4	36	25981 de 3110400	99.1647055041
1	70	H4	36	38546 de 3110400	98.7607381687
1	80	H4	36	51971 de 3110400	98.3291216564
1	90	H4	36	65963 de 3110400	97.8792759774
1	100	H4	36	79886 de 3110400	97.4316486626

en funcionamiento entre las 20:00 y 22:00 horas que entre las 04:00 y 06:00 horas.

Una vez creados los cuatro conjuntos de datos estos se contaminan de dos formas distintas utilizando ruido gaussiano con media igual a zero. La primer forma es contaminar cada una de las componente de cada una de las muestras de forma individual, haciendo la desviación estándar igual al valor de la componente a contaminar, esto con el objetivo de que el ruido generado sea directamente proporcional a la amplitud de la componente a contaminar.

La segunda prueba que se propone es hacer la desviación estándar ( $\sigma$ ) igual a la media de cada muestra (una muestra consta de 4096 componentes), y con esto contaminar de igual forma con la misma función de probabilidad cada una de las componentes de la muestra a contaminar con ruido. Para observar como se comporta la metodología conforme se aumenta el ruido se propone en los dos casos de prueba aumentar gradualmente  $\sigma$ , donde primero se toma el 0% de  $\sigma$ , es decir, la señal pura sin ruido, y se va aumentando de 10% en 10% hasta que sea 100%.

Los resultados obtenidos en ambas pruebas resultaron ser muy buenos, ya que con los conjuntos de datos donde el ruido es pequeño, la metodología tiene una precisión de casi el 100%, mientras que en el peor de los casos es del 97.57% en promedio. Por lo que se concluye que la metodología es bastante robusta ante el ruido que pueda existir en la señal.

A pesar de que los resultados son muy bueno, tienen la desventaja de que son con conjuntos de datos generados de forma artificial mediante una computadora. Por lo que en el siguiente capítulo se concluye lo trabajado en esta investigación y se proponen algunos trabajos futuro que puedan mejorar y confirmar la metodología aquí propuesta para la identificación de dispositivos eléctricos conectados en un instante de tiempo.



## Capítulo 6

# Conclusiones y trabajo futuro

La desagregación de cargas en los sistemas *NILM* es un tema que tiene varias décadas bajo investigación y que hasta la fecha no se ha podido resolver de manera precisa para su comercialización, esto debido en gran parte a que la mayoría de investigaciones que se han echo en este tema, se trata de realizar la clasificación haciendo uso de métodos estocásticos, donde para funcionar se tiene que someter a una etapa de entrenamiento mediante un conjunto de datos etiquetado previamente.

Utilizar métodos estocásticos para realizar la clasificación para la desagregación de cargas en sistemas *NILM*, tiene la gran desventaja que el modelo obtenido al finalizar la etapa de entrenamiento se ajusta únicamente para clasificar el conjunto de dispositivos eléctricos con el que fue entrenado, por lo que si se agregue un nuevo dispositivo el modelo ya no funciona y se tiene que someter nuevamente a una etapa de entrenamiento.

No es factible que el usuario final cada que compre un nuevo dispositivo tenga que realizar la etapa de entrenamiento, evitar esto es una de las principales motivaciones de este trabajo, ya que la metodología propuesta solo requiere de las firmas electromagnéticas de los dispositivos eléctricos para poder funcionar, independientemente del numero de aparatos eléctricos con que se cuente, si se agrega o no dispositivos nuevos o el lugar donde se aplique, por lo cual la metodología propuesta sería ideal para aplicar a este tipo de problemas.

A continuación se muestran las conclusiones generales obtenidas de esta investigación y se proponen algunos trabajos futuros que se derivan.

## 6.1. Conclusiones Generales

Los resultados obtenidos son bastante prometedores al clasificar de forma correcta el 97.7% de los casos aproximadamente. Al tratarse de un problema entero la metodología resulta ser robusta ante el ruido que pueda existir en las mediciones.

Aunque se trate solo de una simulación los resultados reflejan que el enfoque propuesto tiene una gran precisión y que podría dar muy buenos resultados si se implementara de forma real.

Cabe destacar que la metodología propuesta se basa en las siguientes premisas para su correcto funcionamiento, ya que posiblemente sin estas condiciones el enfoque propuesto no serviría.

- La interferencia electromagnética que producen los dispositivos eléctricos es única y repetible, por lo que pueden ser utilizadas como huellas o firmas electromagnéticas para diferenciarlos.
- El ruido electromagnético producido en un hogar en un instante de tiempo es igual a la combinación lineal de las firmas electromagnéticas de los dispositivos que se encuentran en funcionamiento en ese instante de tiempo.

El enfoque propuesto resulta ser bastante interesante y eficiente siempre y cuando se cumpla las premisas expuestas anteriormente, ya que la metodología puede funcionar en cualquier lugar y con cualquier número de dispositivos, ya que solo ocupa el conjunto de firmas electromagnéticas para su correcto funcionamiento.

Este trabajo aporta al estado del arte de los sistemas *NILM* un enfoque distinto y con muy buenos resultados para la desagregación de cargas, ya que no se encontró registro de alguna investigación que utilizara un enfoque similar para identificar los dispositivos que se encuentran conectados en un instante de tiempo.

A continuación se exponen algunos trabajos futuros que se derivan de esta investigación, donde uno a resaltar es buscar la forma de que los fabricantes proporcionen la firma electromagnética de los dispositivos que producen, esto debido a que la metodología solo funciona si se cuenta con todas las firmas de los dispositivos que se encuentran en el hogar a monitorear.

## 6.2. Trabajos Futuros

En esta investigación al tratarse solo de análisis queda mucho por delante que hacer, para confirmar que la metodología propuesta pueda ser utilizada para la desagregación de cargas en los sistemas *NILM*. A continuación se listan una serie de trabajos futuros que se derivan de esta investigación y que podrían confirmar la eficacia del enfoque propuesto.

- Buscar la forma, mediante alguna norma, que los fabricantes proporcionen de manera publica las firmas electromagnéticas de los dispositivos eléctricos que producen.
- Elaborar una base de datos de las firmas electromagnéticas que sea accesible desde cualquier sitio, para que el sistema *NILM* pueda acceder a dichas firmas de los dispositivos con los que se vaya a probar, independientemente de el lugar donde se encuentre.
- Probar la metodología propuesta con un conjunto de datos real, obtenido de monitorear la interferencia electromagnética producida en un hogar en un periodo de tiempo, con el fin de verificar que la metodología si funciona en el mundo real.
- Mejorar la metodología al momento de resolver el sistema de ecuaciones, planteando la solución tomando en cuenta un conjunto de restricciones, por ejemplo que no existan valores negativos, es decir, que no pueda la solución informar que un dispositivo se encuentra menos un veces en funcionamiento.
- Para que la metodología sea mas exacta hay que buscar la forma de caracterizar y filtrar el ruido electromagnético que existe en los hogares, ya que este es el causante de que la precisión de la metodología disminuya.



# Referencias

- [Baranski03] Baranski, M. y Voss, J. Nonintrusive appliance load monitoring based on an optical sensor. *En Power Tech Conference Proceedings, 2003 IEEE Bologna*, tomo 4, págs. 8–pp. IEEE, 2003.
- [Chang08] Chang, H.-H., Yang, H.-T., y Lin, C.-L. Load identification in neural networks for a non-intrusive monitoring of industrial electrical loads. *En Computer Supported Cooperative Work in Design IV*, págs. 664–674. Springer, 2008.
- [Chvatal83] Chvatal, V. *Linear programming*. Macmillan, 1983.
- [Cortes95] Cortes, C. y Vapnik, V. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [Farinaccio99] Farinaccio, L. y Zmeureanu, R. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, 30(3):245–259, 1999.
- [Figueiredo11] Figueiredo, M. B., De Almeida, A., y Ribeiro, B. An experimental study on electrical signature identification of non-intrusive load monitoring (nilm) systems. *En Adaptive and Natural Computing Algorithms*, págs. 31–40. Springer, 2011.
- [Gupta10] Gupta, S., Reynolds, M. S., y Patel, S. N. Electrisense: single-point sensing using emi for electrical event detection and classification in the home. *En Proceedings of the 12th ACM international conference on Ubiquitous computing*, págs. 139–148. ACM, 2010.
- [Hart85] Hart, G. Prototype nonintrusive appliance load monitor, 1985.

- [Hart92] Hart, G. W. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.
- [Kaggle] Kaggle. Belkin energy disaggregation competition. <https://www.kaggle.com/c/belkin-energy-disaggregation-competition/data>. Accessed: 2014-09-30.
- [Lam07] Lam, H., Fung, G., y Lee, W. A novel method to construct taxonomy electrical appliances based on load signaturesof. *Consumer Electronics, IEEE Transactions on*, 53(2):653–660, 2007.
- [Li12] Li, J., West, S., y Platt, G. Power decomposition based on svm regression. *En Modelling, Identification & Control (ICMIC), 2012 Proceedings of International Conference on*, págs. 1195–1199. IEEE, 2012.
- [Marceau00] Marceau, M. L. y Zmeureanu, R. Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. *Energy Conversion and Management*, 41(13):1389–1403, 2000.
- [Norford96] Norford, L. K. y Leeb, S. B. Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings*, 24(1):51–64, 1996.
- [Patel07] Patel, S. N., Robertson, T., Kientz, J. A., Reynolds, M. S., y Abowd, G. D. *At the flick of a switch: Detecting and classifying unique electrical events on the residential power line (nominated for the best paper award)*. Springer, 2007.
- [Perona90] Perona, P. y Malik, J. Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):629–639, 1990.
- [Rojas96] Rojas, R. *Neural networks: a systematic introduction*. Springer Science & Business Media, 1996.
- [Ruzzelli10] Ruzzelli, A. G., Nicolas, C., Schoofs, A., y O’Hare, G. M. Real-time recognition and profiling of appliances through a single electricity sensor. *En*

- 
- Sensor Mesh and Ad Hoc Communications and Networks (SECON), 2010 7th Annual IEEE Communications Society Conference on*, págs. 1–9. IEEE, 2010.
- [sener10] sener. Prospectiva del sector eléctrico 2010-2025. *Secretaría de Energía*, 2010.
- [Sillverman51] Sillverman, B., Jones, M., Fix, E., y Hodges, J. An important contribution to nonparametric discriminant analysis and density estimation. *International Statistical Review*, 57(3):233–247, 1951.
- [Zeifman11] Zeifman, M. y Roth, K. Nonintrusive appliance load monitoring: Review and outlook. *Consumer Electronics, IEEE Transactions on*, 57(1):76–84, 2011.
- [Zoha12] Zoha, A., Gluhak, A., Imran, M. A., y Rajasegarar, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12):16838–16866, 2012.