

Universidad de Alcalá

Escuela Politécnica Superior

**Máster Universitario en Ingeniería de
Telecomunicación**

Trabajo Fin de Máster

**Detección y predicción de errores
en smart grids mediante modelos
de inteligencia artificial**

ESCUELA POLITECNICA
SUPERIOR

Autor: Paula Bartolomé Mora

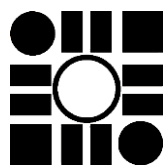
Tutor: Elisa Rojas Sánchez

Cotutor: David Carrascal Acebron

2024



Máster Universitario en Ingeniería de Telecomunicación



Universidad
de Alcalá

Madrid, 1 de julio de 2024

UNIVERSIDAD DE ALCALÁ

ESCUELA POLITÉCNICA SUPERIOR

Máster Universitario en Ingeniería de Telecomunicación

Trabajo Fin de Máster

**Detección y predicción de errores en smart grids mediante
modelos de inteligencia artificial**

Autor: Paula Bartolomé Mora

Tutor: Elisa Rojas Sánchez

Cotutor: David Carrascal Acebron

Tribunal:

Presidente:

Vocal 1º:

Vocal 2º:

AQUI DEDICATORIA

Agradecimientos

AQUI AGRADECIMIENTOS

Resumen

AQUI RESUMEN

Palabras clave:

Abstract

In this Master's Thesis (TFM) we present... AQUI ABSTRACT

Keywords:

AQUI CITA CELEBRE

AQUI AUTOR

Índice general

Resumen	v
Abstract	vii
1. Introducción	1
1.1. Presentación	1
1.2. Objetivos	2
1.3. Estructura del TFM	3
2. Estado del arte	5
2.1. Smart Grids	5
2.1.1. Flujo energético y comunicación bidireccional	6
2.1.2. Prosumidores y microgrids	8
2.1.2.1. Mercado energético	9
2.1.3. Estructura de una Smart Grid	12
2.1.3.1. Generación	12
2.1.3.2. Distribución	15
2.1.3.3. Consumo	19
2.1.4. Protocolos	19
2.1.5. Topología de red	19
2.1.6. Seguridad en Smart grids	19
2.1.7. Estándares y regulación en smart grids	21
2.1.8. Beneficios ambientales	21
2.1.9. Situación en España	21
2.1.10. Desafíos futuros de las smart grids	22
2.1.10.1. Retos técnicos	22
2.1.10.2. Retos económicos	22
2.2. IoT	22
2.2.1. IoE	23
2.2.2. Redes LLN	24
2.2.2.1. RPL	24
2.3. DEN2NE	24

2.4. Big Data	24
2.4.1. Computación en nube	25
2.5. Machine Learning/IA	27
2.6. Herramientas software	28
2.6.1. BRITE	28
2.6.1.1. Definición de topologías	28
2.6.1.2. Modelos de topologías	29
2.6.1.3. Automatización de la ejecución	30
2.6.1.4. Graficación de topologías	31
3. Diseño y análisis de datos	33
3.1. Análisis de los conjuntos de datos	33
3.1.1. <i>SustDataED</i>	37
3.1.2. Estructura del dataset <i>SustDataED</i>	37
3.1.2.1. Medidas de consumo de energía	38
3.1.2.2. Medidas de eventos de potencia	40
3.1.2.3. Medidas de producción de energía	40
3.1.2.4. Medidas demográficas	41
3.1.2.5. Medidas de eventos de usuario	41
3.1.2.6. Medidas de condiciones ambientales y climáticas	42
3.1.3. Conclusiones del análisis del dataset <i>SustDataED</i>	42
3.1.4. Global Solar Atlas	42
3.2. Preprocesado	42
4. Desarrollo y evaluación del modelo	43
4.1. Selección de modelos	43
4.2. Modelo 1	43
5. Conclusiones y trabajo futuro	45
Bibliografía	47
Lista de Acrónimos y Abreviaturas	53
A. Anexo I - Pliego de condiciones	55
B. Anexo II - Presupuesto	57

C. Anexo III - Manuales de usuario e Instalación	59
C.1. Herramienta BRITE	59
C.1.1. Virtualización del sistema operativo Linux	59
C.1.2. Compilación e instalación de BRITE	59
C.1.3. Ejecución de BRITE	60

Índice de figuras

1.1. Proceso de detección y predicción de errores	2
2.1. Flujo bidireccional de datos y energía en una <i>smart grid</i> [9]	7
2.2. Infraestructura de una interfaz AMI para el flujo bidireccional de energía y datos [7]	7
2.3. Estructura de la respuesta a la demanda en el contexto de una <i>smart grid</i> con prosumidores [?]	9
2.4. Modelo basado en las interacciones de los agregadores con los prosumidores y el mercado eléctrico [13]	10
2.5. Estrategia de <i>Demand Side Management</i> (DSM) basada en interacciones individuales de cada consumidor con la compañía [16]	11
2.6. Estrategia de DSM para smart grids basada en interacciones entre los usua- rios y la compañía [16]	12
2.7. Estructura de componentes de una SG [5]	13
2.8. Esquema de funcionamiento de una unidad CHP [17]	13
2.9. Gestión de operaciones en una unidad CHP [18]	14
2.10. Representación del proceso P2G [19]	15
2.11. Instalación de un SVC en Sylling (Noruega) [21]	16
2.12. Representación de un sistema de transmisión de energía eólica con líneas HVDC [25]	18
2.13. Red nacional de transmisión energética [30]	21
2.14. Infraestructura de componentes de un sistema IoT [33]	23
2.15. Infraestructura de <i>Big Data</i> enfocada al ámbito de las <i>Smart Grid</i> (SG)s [26]	26
2.16. Plataforma de análisis de <i>Big Data</i> en el entorno de SGs [38]	26
2.17. Modelos soportados por <i>Boston University Representative Internet Topology</i> <i>Generator</i> (BRITE) [4]	29
3.1. Representación del período temporal que comprende el proceso de recolec- ción de datos para cada hogar	38
C.1. Características del sistema operativo	60

Índice de tablas

2.1. Tecnologías FACTS de 1 ^a y 2 ^a generación	17
3.1. Comparación entre datasets públicos en el ámbito NILM [42] [53] [54] [43] .	36
3.2. Tabla de medidas de consumo energético [51]	39
3.3. Tabla de medidas de eventos de potencia [51]	40
3.4. Tabla de medidas de producción de energía [51]	40
3.5. Tabla de medidas de producción de energía [51]	41
3.6. Tabla de medidas de eventos de usuario [51]	41
3.7. Tabla de medidas de condiciones ambientales [51]	42

Índice de Códigos

C.1. Instalación de paquetes	60
C.2. Ejecución de la interfaz gráfica de BRITE	60
C.3. Ejecucion de BRITE por la línea de comandos	60

1. Introducción

En este primer capítulo, se presentarán a modo de introducción los aspectos más relevantes del Trabajo Fin de Máster (TFM) desarrollado, como son la transformación de las redes energéticas en nuevas redes inteligentes de energía (del inglés SG) y su consecuente necesidad de monitorización, además del creciente auge de la Inteligencia Artificial (IA) y de las técnicas de aprendizaje automatizado (del inglés *Machine Learning* (ML)).

De la misma manera, se hará hincapié en los objetivos que se establecen para este TFM, describiendo así la secuencia de acciones que se seguirá para poder llevarlos a cabo. Dichos objetivos serán los pilares sobre los que se sustente el diseño y el desarrollo de un modelo eficiente basado en ML para la detección y predicción de errores en SGs. De forma adicional, se incluirá un último apartado en referencia a la estructura de capítulos que seguirá este TFM, así como una explicación breve sobre los temas que se abordarán en cada uno de ellos.

1.1. Presentación

En los últimos años, el concepto de red eléctrica convencional está siendo completamente transformado como resultado del creciente desarrollo de las SGs [1] [2]. La posibilidad de brindar a tiempo real una monitorización del proceso de distribución eléctrica y a su vez, integrar una participación activa del consumidor convierte a estas redes en uno de los pilares de la transición hacia la descarbonización.

En este contexto, se ha convertido en un objetivo primordial el diseño de nuevos protocolos que provean un encaminamiento y una distribución energética eficientes. Es por ello que, desde el equipo de investigación NetIS de la *Universidad de Alcalá* (UAH) se ha desarrollado el algoritmo *Distributed ENergy ENvironments and Networks* (DEN2NE) [3], con el fin de automatizar la gestión de los recursos en entornos del Internet de las Cosas (del inglés Internet of Things (IoT)), enfocándose especialmente en SGs. Teniendo esto en cuenta, se puede expresar que DEN2NE pretende realizar una distribución óptima de los recursos energéticos de cada uno de los nodos que presenta una determinada topología de red para conseguir finalmente un balance a nivel global.

Sin embargo, el algoritmo se encuentra actualmente con un desafío a afrontar, basado en la imperiosa necesidad de poder detectar y predecir de una forma precisa los posibles errores que se puedan producir en el proceso de distribución energética entre nodos. En caso de fallos en la red, DEN2NE debería tener la capacidad de responder de una forma rápida y específica, pues es de vital importancia tomar acciones a nivel de distribución para evitar que los errores se agraven y escalen al resto de la red.

1.2. Objetivos

Tomando en consideración el apartado anterior, se define como objetivo principal de este TFM el desarrollo de técnicas de ML para identificar de forma precisa los posibles errores que se pueden producir en una SG durante el proceso de distribución energética que, como se ha introducido, se aplicará mediante el algoritmo DEN2NE.

La identificación de patrones se apoyará en el análisis de grandes volúmenes de datos provenientes de implementaciones de SGs reales. Estos serán previamente procesados de una forma exhaustiva para reducir su tamaño y seleccionar la información verdaderamente necesaria para el posterior desarrollo de técnicas de ML. En la Figura 1.1 se determina el diagrama de flujo que define la secuencia de acciones a realizar en este TFM.

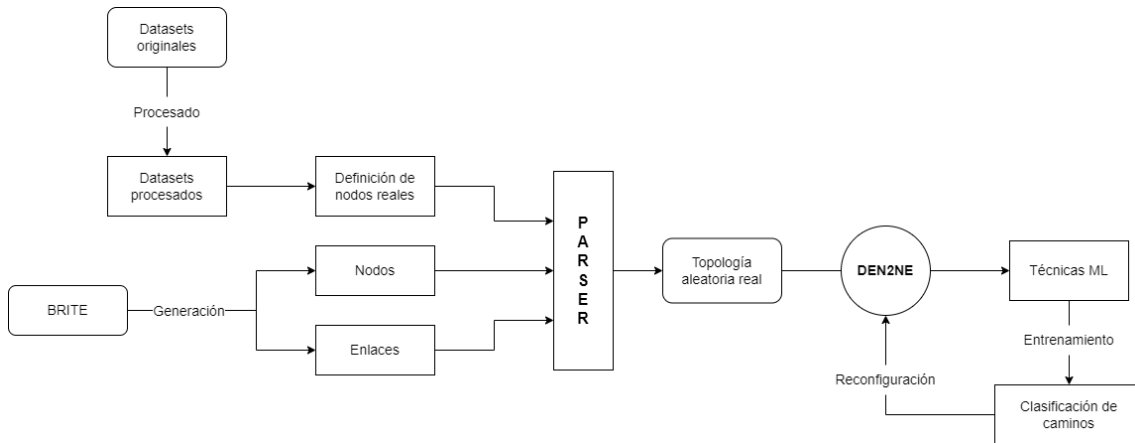


Figura 1.1: Proceso de detección y predicción de errores

Por tanto, el presente TFM permitirá determinar un modelo final de predicción de fallos para entornos de SGs. Entrando en detalle, se pueden definir los diferentes objetivos en los siguientes puntos:

- **Estudio del estado del arte y análisis de las diferentes fuentes de datos reales disponibles.** Se realizará un estudio en profundidad del contexto en el que se engloban las SGs y los protocolos de distribución de recursos, poniendo el enfoque en DEN2NE. Será necesaria una investigación sobre la disponibilidad de fuentes de datos reales, además de una evaluación de su viabilidad para este TFM.
- **Definición y procesamiento del volumen de datos para el posterior desarrollo.** El procesamiento de los datos adquiridos será fundamental para simplificar su comprensión. Además, se llevará a cabo la selección de la cantidad de información realmente útil para el desarrollo.
- **Planteamiento de diferentes escenarios.** A partir del volumen final de datos que se elaborará, se plantearán una serie de escenarios mediante la generación de topologías de nodos con la herramienta BRITE [4].
- **Generación y simulación de topologías.** Las topologías se importarán a un simulador propietario realizado en python para obtener el conjunto de datos final.
- **Desarrollo y entrenamiento de diferentes modelos de ML.** Se realizará un estudio sobre las posibles técnicas de ML a emplear y se procederá a seleccionar las más apropiadas, según el contexto en el que se engloba este TFM. Por consiguiente, se llevará a cabo el desarrollo y el entrenamiento de los mismos tomando como base los datos procesados anteriormente.
- **Análisis y evaluación de los resultados obtenidos.** Los resultados de detección y predicción de errores de los diferentes modelos de ML serán analizados. Se determinará de forma concluyente cuál aporta una mayor efectividad para poder diagnosticar con un alto porcentaje de precisión los errores que se producen en una SG.

1.3. Estructura del TFM

En esta sección se expone de forma general la estructura de capítulos que tendrá la memoria del presente TFM con una breve descripción del contenido que abarcarán cada uno de ellos.

Capítulo 1: Introducción. Se comenzará esta memoria con una presentación de los aspectos más significativos. Se expondrán las motivaciones que han originado la realización de este TFM y los objetivos que se pretenden alcanzar con el mismo.

Capítulo 2: Estado del arte. Se establecerá el marco teórico en el que se sitúa este TFM y se documentarán los conceptos principales para introducir en un contexto lo suficientemente consistente el diseño y desarrollo práctico posterior. Además, se expondrá de forma detallada el funcionamiento del algoritmo DEN2NE.

Capítulo 3: Diseño y análisis de datos. Se llevará a cabo un análisis sobre las posibilidades de diseño a través de una investigación exhaustiva sobre las fuentes de datos de implementaciones reales existentes. Por consiguiente, se definirá todo el procesamiento que deberá llevarse a cabo sobre los datos para seleccionar la información útil sobre la que entrenar los modelos de ML a desarrollar.

Capítulo 4: Desarrollo y evaluación de los modelos. Se describirá en detalle el desarrollo de las técnicas de ML que han sido seleccionadas según el contexto en el que se engloba este TFM. Los modelos obtenidos serán entrenados, analizados y por consiguiente, evaluados. Los resultados definirán finalmente el modelo más apropiado para diagnosticar con alta precisión los errores que se pueden producir en el entorno de una SG.

Capítulo 5: Conclusiones y trabajo futuro. Se finalizará la memoria con un capítulo enfocado a las conclusiones obtenidas tras la realización de este TFM. Además, se indicarán las posibles vías de trabajo a futuro.

Bibliografía. Se incluirán todas las referencias de artículos, libros, páginas web u otros materiales que han sido consultados y empleados para elaborar esta memoria. Se seguirá el estilo de citación del Institute of Electrical and Electronics Engineers (IEEE), así como las recomendaciones oficiales de la normativa sobre TFMs de la UAH.

Anexos. Se añadirán los manuales de usuario y de instalación de herramientas. También, se expondrán las especificaciones del *hardware* sobre el que se ha desarrollado este TFM y una estimación de su presupuesto.

2. Estado del arte

En el presente capítulo se entrará en profundidad en los conceptos más relevantes relacionados con el TFM. Además, se describirán las diversas herramientas de software que se emplearán. El propósito principal de presentar el estado del arte es establecer un marco teórico consistente que permita conseguir una mayor comprensión del contexto previamente a la exposición del desarrollo del proyecto.

2.1. Smart Grids

En las últimas décadas, se ha observado una transformación integral del modelo asociado a la red eléctrica convencional. El aumento de la energía demandada por los usuarios finales y los requisitos cada vez mayores de la industria ha tenido como consecuencia que algunos países hayan intentado diseñar redes eléctricas agrupadas en un conjunto de grandes redes nacionales. De este modo, todas las fuentes energéticas disponibles pueden estar conectadas para ser gestionadas conjuntamente en función de la demanda recibida, además de conseguir una coordinación a alto nivel. [5]

Sin embargo, una red eléctrica no se puede definir como una entidad única e independiente, pues se trata de una agregación de redes, compañías energéticas y operadores que trabajan en distintos niveles de comunicación. Es por ello, que la idea de desarrollar una gran red nacional encuentra cierto equilibrio energético, pero no llega a los altos porcentajes de eficiencia que pueden proveer las redes inteligentes energéticas o de otra forma, las *smart grids* (SG).

La iniciativa sobre redes inteligentes del IEEE, califica a las *smart grids* [6] como "imperativas y revolucionarias que implican nuevas capacidades de comunicación y control, fuentes de energía, modelos de generación y adherencia a estructuras regulatorias transjurisdiccionales". Por otro lado, en el año 2009, el Departamento de Energía de Estados Unidos [7] determinó en su reporte anual sobre *smart grids* los siguientes objetivos principales que se perseguían con el desarrollo de este tipo de redes:

- Permitir una participación activa de los clientes en el sistema.
- Integrar todas las opciones de generación y almacenamiento de energía.
- Ofertar nuevos productos y servicios.
- Proporcionar energía de calidad y satisfacer un gran rango de necesidades.
- Optimizar el uso de los recursos y aumentar la eficiencia.
- Dar una respuesta rápida ante emergencias y problemas producidos en la red.

Tomando en consideración estos objetivos, se puede definir una *smart grid* como una red inteligente con la capacidad de distribuir los suministros de energía de forma optimizada a los usuarios, basándose en la información que recoge de los mismos. Es por ello que supone una actualización digital de las redes de distribución y transmisión a larga distancia para incorporar sistemas de monitorización y control a tiempo real. [8]

Su base se asienta en el diseño de sistemas inteligentes coordinados capaces de obtener tanto la información respectiva a la demanda o requisitos energéticos de cada zona como la de la disponibilidad de recursos a partir de las diferentes fuentes de producción existentes. Todo ello conduce al desarrollo de un plan estratégico de distribución energética para poder conectar a todas las entidades participantes en la red entre sí. [1]

2.1.1. Flujo energético y comunicación bidireccional

Cabe destacar que una de las principales características diferenciadoras de las SGs respecto a las redes energéticas convencionales es que se fundamentan en una comunicación bidireccional. Este factor determinante se puede apreciar en la Figura 2.1 y se traduce en una producción y una distribución dinámica y personalizada hacia cada usuario de la red.

Es decir, en una red tradicional al ser unidireccional, se provee energía desde el distribuidor hacia el consumidor sin llevar a cabo ningún análisis estadístico sobre el consumo que se está produciendo en las líneas finales en un determinado instante temporal. En cambio, en el contexto de las SGs, se pone el foco en las acciones de los usuarios y en la consecuente asignación de patrones de consumo eléctrico que permita predecir el comportamiento futuro de los mismos. [?]

Sin embargo, como se expondrá más adelante en la Sección 2.1.3.3, este proceso de clasificación de usuarios requerirá del análisis de grandes volúmenes de datos adquiridos a tiempo real, por lo que se aumentará la complejidad del sistema a costa de alcanzar la eficiencia.

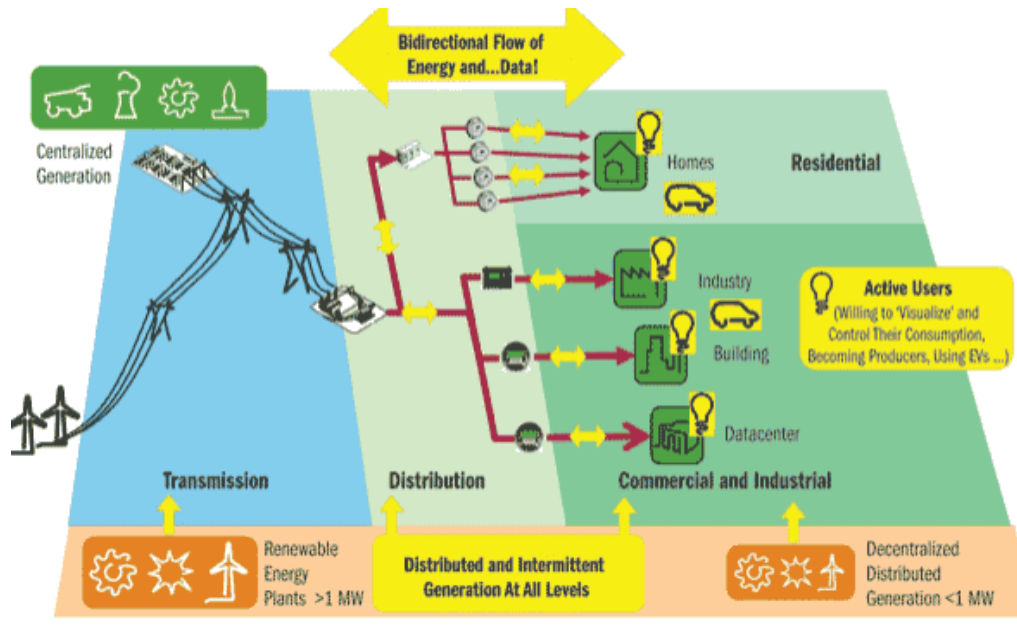


Figura 2.1: Flujo bidireccional de datos y energía en una *smart grid* [9]

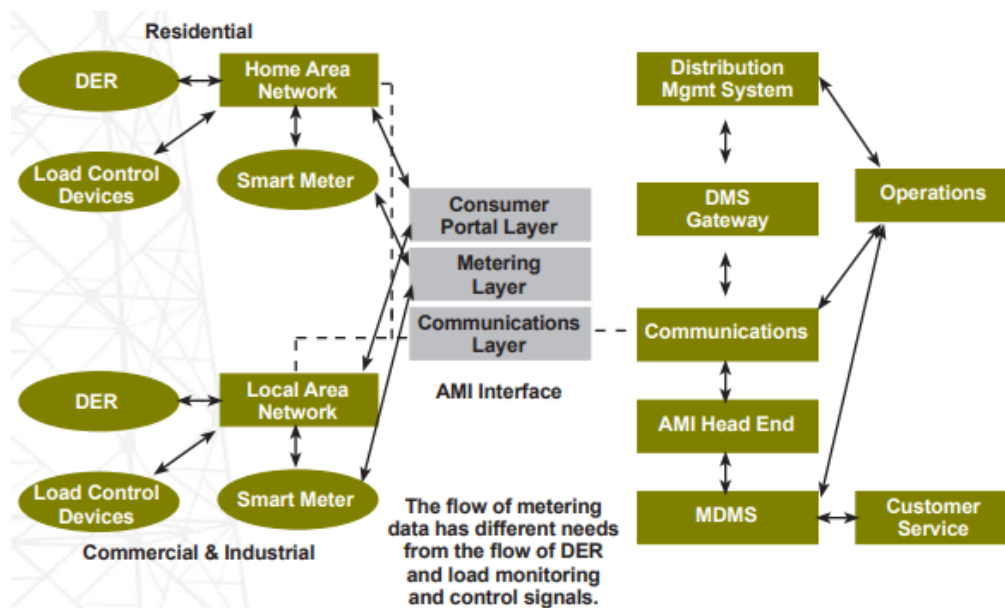


Figura 2.2: Infraestructura de una interfaz AMI para el flujo bidireccional de energía y datos [7]

Por lo tanto, bajo la premisa de la característica de red bidireccional, se puede expresar que los propios usuarios son el gran pilar sobre el que se cimienta el sistema. En comparación a la red eléctrica tradicional, toman un papel mucho más activo monitorizando continuamente su comportamiento eléctrico y recopilando información para trasladarla al

resto de la red. Esto tiene como ventaja una gran reducción de los costes derivados de la distribución y transmisión en el sistema, ya que se evita proporcionar más cantidad de energía de la requerida por las líneas finales. [8]

Es por ello que se emplea una Infraestructura Avanzada de Medición (del inglés Advanced Metering Infrastructure (AMI)) con una interfaz constituida por varias capas o niveles como se puede apreciar en la Figura 2.2. De esta forma, se dividen los flujos de comunicación entre las diferentes entidades que componen la red para permitir una mejor gestión de los mismos. [7]

2.1.2. Prosumidores y microgrids

Dentro del presente contexto, en referencia a los usuarios es preciso introducir el término de prosumidor [?]. Se puede determinar como prosumidora a toda entidad o usuario final que tiene la capacidad de producir de forma alternativa recursos energéticos, además de poder consumirlos. En la Figura 2.3 se puede visualizar cómo se estructura la SG en un conjunto de distintas comunidades o agregaciones de prosumidores con el fin de crear microrredes de distribución y almacenamiento energético (del inglés *microgrids*).

En este caso es imprescindible que los equipos y dispositivos dedicados a la generación de electricidad se encuentren dentro de la microrred o en una ubicación cercana a la misma. Así, se podrá abastecer a los usuarios del conjunto de forma eficiente, reduciendo los costes relacionados con la transmisión energética.

Este tipo de esquema puede ser aplicado en diferentes entornos, como pueden ser los residenciales, industriales o agrícolas, para conseguir descentralizar la gestión de los recursos. Sin embargo, debe existir cierta gestión externa y coordinada de todas estas microrredes o comunidades de usuarios. Esta es llevada a cabo por una entidad denominada como operador del sistema distribuido (del inglés *Distribution System Operator* (DSO)) [10] [11]. Algunas de sus funciones principales se basan en asegurar el abastecimiento de recursos en el interior de las microrredes en función de la demanda existente y en supervisar el estado operativo la infraestructura de red para garantizar su estabilidad y seguridad.

Por otro lado, se encuentra también la figura del centro de servicios auxiliares [10], que se encarga de recibir la información sobre los recursos energéticos de los prosumidores y cuantificar en función de la oferta y demanda el beneficio económico que obtendrán los prosumidores. Estos normalmente son gestionados por los operadores del mercado eléctrico, pero puede variar según la región o el país.

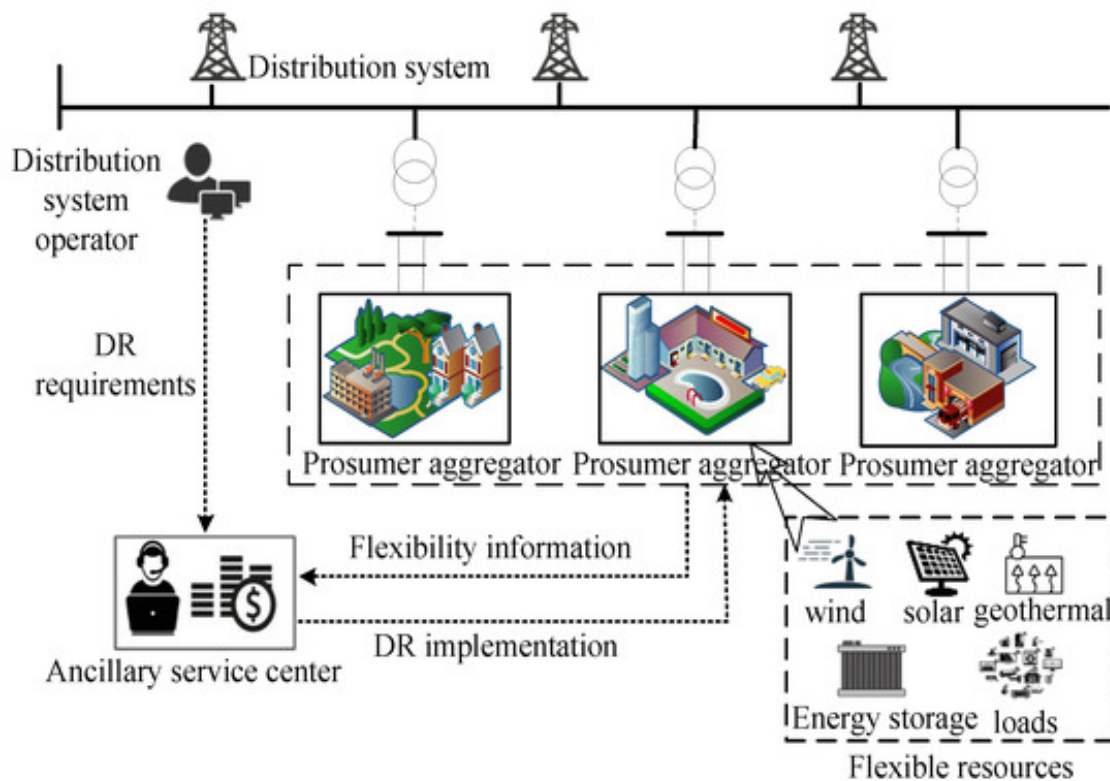


Figura 2.3: Estructura de la respuesta a la demanda en el contexto de una *smart grid* con prosumidores [?]

2.1.2.1. Mercado energético

Las agregaciones de prosumidores posibilitan el autoconsumo y la compartición de los recursos generados por sí mismos con los demás participantes de la microrred, incluso con el resto de la SG. La mayoría de los prosumidores, sobre todo en entornos residenciales son de pequeño tamaño y tienen mayores complicaciones para participar en los mercados eléctricos por sí mismos.

Por ello, los agregadores [12] [10] se encargan de facilitar esta tarea para que una cantidad de prosumidores locales puedan actuar en conjunto como fijadores de precios de compra/-venta en el mercado eléctrico global. Es decir, cada agregador se encarga de representar a los prosumidores que lo componen y actúa como intermediarios entre los mismos y el conjunto de mercados energéticos.

A cambio, cada prosumidor debe pagar un plan mensual a la compañía para cubrir los costes eléctricos y así, dejar la responsabilidad de las transacciones a los agregadores para que operen de forma flexible. En esta operativa también incide el DSO, el cual se

encarga de que las transacciones de compra/venta sigan el cumplimiento de los estándares y regulaciones establecidos en el mercado energético.

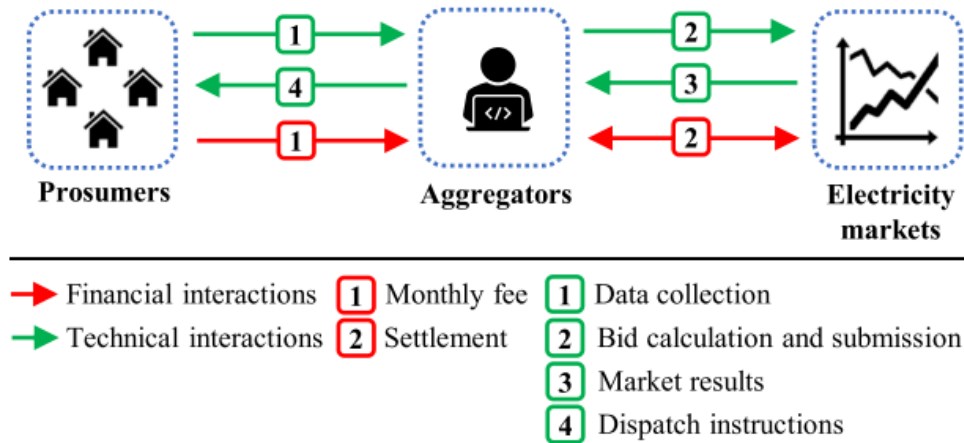


Figura 2.4: Modelo basado en las interacciones de los agregadores con los prosumidores y el mercado eléctrico [13]

A partir del funcionamiento expuesto y de la Figura 2.4, se pueden diferenciar los tipos de interacciones que se producen en este modelo [13]:

- Interacciones técnicas
 1. Recopilación y adquisición de datos sobre los recursos de los prosumidores para su posterior procesamiento en los agregadores.
 2. Cuantificación y presentación de ofertas desde los agregadores hacia los mercados mayoristas de electricidad.
 3. Comunicación de los resultados de diferentes mercados a los agregadores.
 4. Envío de instrucciones o recursos distribuidos en función de los resultados del mercado.
- Interacciones financieras
 1. Pago de tarifa mensual de cada prosumidor al agregador.
 2. Liquidación de las transacciones trasladadas desde los prosumidores en el mercado eléctrico (cobro por demanda y pago por generación).

En cuanto al proceso de fijación de precios, generalmente, se adoptan tarifas minoristas dinámicas según el instante temporal, como pueden ser la fijación de precios por tiempo de uso (del inglés *Time-of-Use Pricing* (TOU)) o en tiempo real (del inglés *Real Time Pricing*

(RTP)). Dentro de este contexto es preciso introducir los programas de gestión del lado de la demanda (del inglés DSM). Como término, un programa DSM es un programa basado en el control de las interacciones de consumo y de gestión de cargas residenciales desde el lado del cliente.

Considerando el modelo de interacciones expuesto anteriormente, se puede expresar que un DSM tiene como pilar fundamental la distribución de recursos energéticos de manera eficiente y acorde a la demanda y a la oferta. Permite reducir el coste de la adquisición energética y los costes asociados a la distribución como consecuencia de minimizar el número de interacciones necesarias entre el prosumidor y el resto de entidades del sistema. [14]

Por lo tanto, volviendo a los modos de dinamización de las tarifas, si por ejemplo se emplea RTP, se buscará lograr el equilibrio de la demanda en tiempo real, modificando las cargas de los consumidores en las horas pico [15]. Es decir, cada uno de los usuarios actúa individualmente en función de los precios dinámicos en el tiempo, comunicándose directamente con la compañía energética como se puede visualizar en la Figura 2.5.

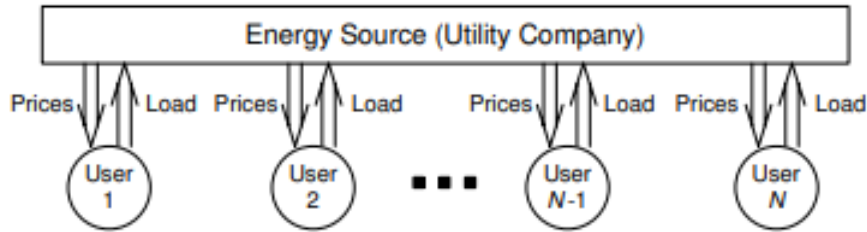


Figura 2.5: Estrategia de DSM basada en interacciones individuales de cada consumidor con la compañía [16]

Con este proceso, el consumidor como puede saber en tiempo real la tarifa a la que está consumiendo energía, puede trasladar su propia carga desde las horas donde el precio es más alto (horas pico) a las de precio menor (horas valle). Esta gestión produce por tanto, menores picos de demanda y contribuyendo a la bajada de los precios. [14] [16]

No obstante, el diseño de un programa DSM ideal en el contexto de las SGs debería permitir también las interacciones entre los mismos prosumidores dentro de una zona residencial o microgrid. Estas interacciones generalmente se automatizan a través de equipos dedicados a una comunicación digital bidireccional. Como se aprecia en la Figura 2.6, este proceso tiene como fin conducir a una coordinación de las acciones respectivas a las cargas de un área determinada.

Por tanto, en este caso la monitorización de las cargas totales para todos los nodos participantes en un instante determinado aportará la información necesaria sobre la relación entre la potencia pico y promedio (del inglés *Peak-to-Average Ratio* (PAR)) y contribuirá a la fijación del precio unitario para ese mismo instante. [16]

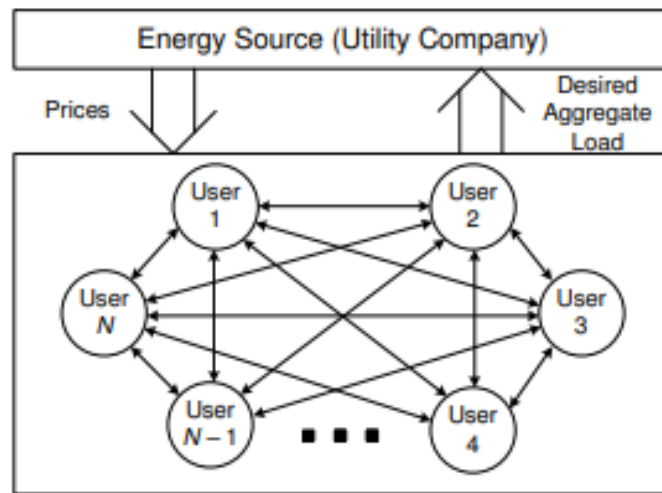


Figura 2.6: Estrategia de DSM para smart grids basada en interacciones entre los usuarios y la compañía [16]

2.1.3. Estructura de una Smart Grid

Una SG está constituida por múltiples elementos diferentes como se puede visualizar en la Figura 2.7. Cada uno de ellos está dedicado a uno de los procesos principales, que se pueden dividir en generación, distribución y consumo. [5]

2.1.3.1. Generación

Como ya se introducía en la Sección 2.1.2, dentro de una SG, el prosumidor se trata de la figura que potencia principalmente el uso de energías renovables, las cuales pueden ser obtenidas por ejemplo a partir de generadores térmicos fotovoltaicos (del inglés *Photovoltaic Thermal* (PVT)) o de turbinas eólicas. Para posibilitar posteriormente el uso de la energía generada, esta debe ser convertida y acondicionada mediante dispositivos dedicados, como son unidades combinadas de calor (del inglés *Combined Heat and Power* (CHP)) o de conversión a gas (del inglés *Power-to-Gas* (P2G)). [10]

Las unidades CHP [17] llevan a cabo un aprovechamiento del calor residual que deriva del proceso de generación de electricidad. Se emplean como respaldo eléctrico, ya que se

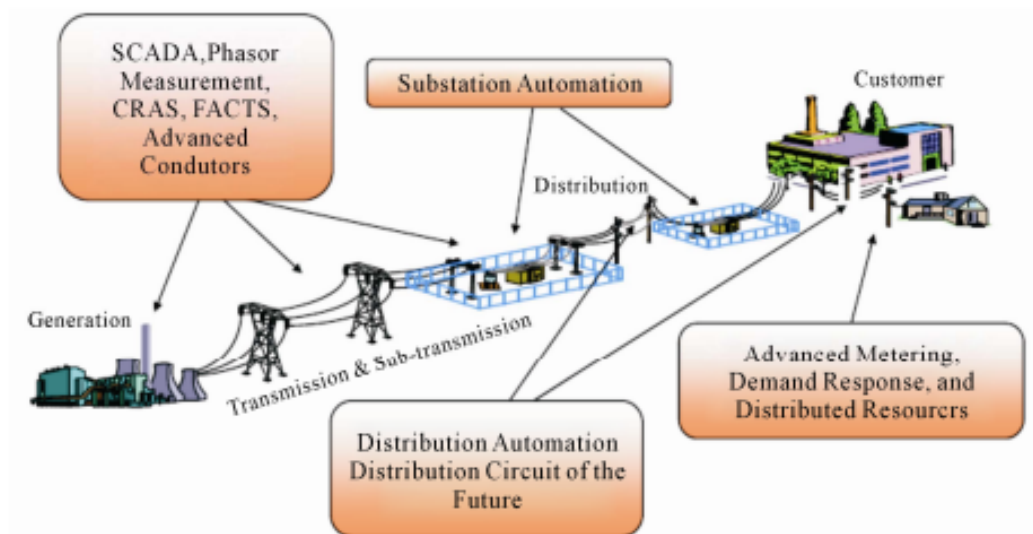


Figura 2.7: Estructura de componentes de una SG [5]

componen de sistemas de almacenamiento de energía térmica (del inglés *Thermal Energy Storages* (TES)) para poder separar la producción y el uso del calor y la energía. Luego, este calor almacenado se puede emplear en aplicaciones térmicas como la calefacción u otros procesos industriales. Es por ello que las unidades CHP mejoran la eficiencia del sistema energético, contribuyendo a un uso mucho más eficaz de los recursos y a una reducción de las pérdidas que se producen en la transmisión. En la Figura 2.8 se representa a modo esquemático el funcionamiento de una unidad CHP. [18]

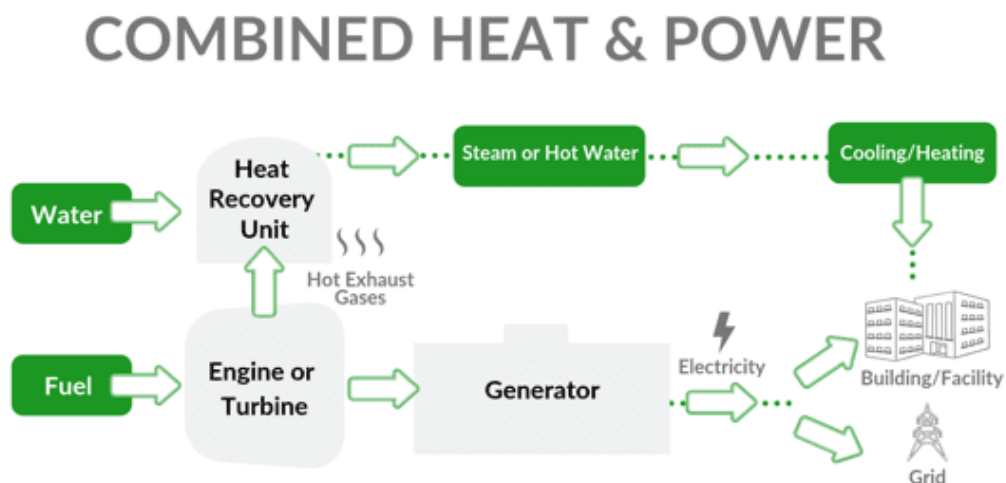


Figura 2.8: Esquema de funcionamiento de una unidad CHP [17]

Además del TES, utilizado para almacenar el calor producido, una unidad CHP también se constituye de un motor acoplado a un generador eléctrico para llevar a cabo el proceso de generación de electricidad y calor. Como se puede apreciar en la Figura 2.9, todas las operaciones de la unidad y las interacciones que se producen entre módulos se coordinan desde el gestor de operaciones para que se produzca un correcto funcionamiento.

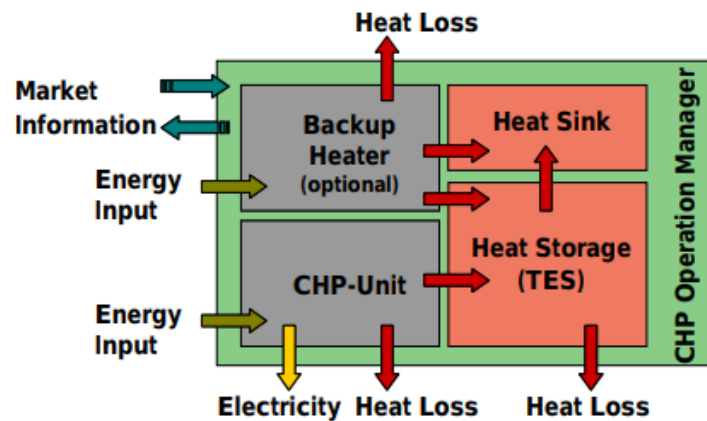


Figura 2.9: Gestión de operaciones en una unidad CHP [18]

Por otro lado, la unidad P2G, mencionada anteriormente, se encarga de la conversión de la electricidad en gases sintéticos, como son el hidrógeno o el metano. Es de gran importancia, ya que este tipo de gases son más fácil de transportar y almacenar que la electricidad, por lo que se simplifica la gestión energética. [10]

Como se puede apreciar en la Figura 2.10, el objetivo del P2G se basa en descomponer el agua mediante un proceso de electrólisis a partir de electricidad proveniente de fuentes renovables. Después, el hidrógeno que se produce se puede utilizar directamente o procesarlo de forma adicional en metano.

Dentro del contexto de la generación de energía también es importante destacar algunos equipos, como son los sistemas de control y adquisición de datos (del inglés *Supervisory Control and Data Acquisition* (SCADA)) [20]. Estos pueden ser instalados en generadores, como son paneles solares fotovoltaicos o turbinas eólicas, consiguiendo una monitorización remota de los mismos. Mediante la información recopilada a tiempo real permiten conocer los niveles de generación y establecer una predicción de la disponibilidad energética que habrá en el sistema o en un área determinada del mismo.

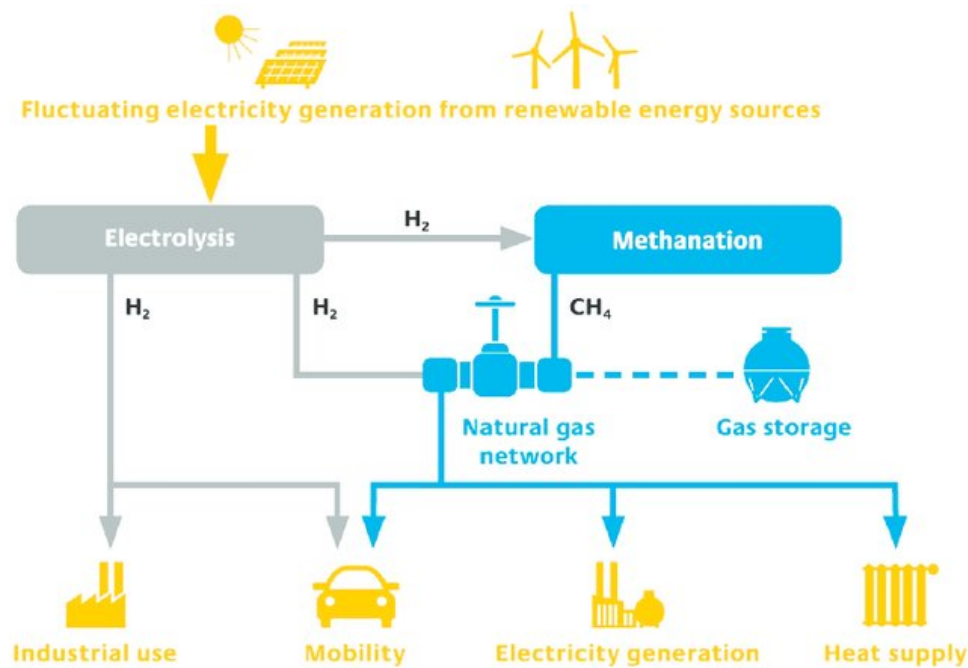


Figura 2.10: Representación del proceso P2G [19]

2.1.3.2. Distribución

Un dispositivo importante en el campo de la distribución energética es la Unidad de Medición de Fasores (del inglés *Phasor Measurement Unit* (PMU)). Se emplea para medir con una alta precisión los fasores de tensión y corriente en la red eléctrica, proporcionando información relevante sobre las magnitudes y fases de las ondas sinusoidales.

También, es empleado en los generadores y debe contar con una alta tasa de muestreo para capturar eventos transitorios o cambios rápidos en la red. En otros términos, debe tener la capacidad de identificar o detectar a tiempo real posibles anomalías que se puedan dar en la distribución y por ello, se trata de un componente imprescindible para comprobar y garantizar la estabilidad de la SG.

La confiabilidad y la seguridad de la red también reside en los Sistemas de Transmisión de Corriente Alterna Flexibles (del inglés *Flexible AC Transmission System* (FACTS)) [21] [22]. Estos vienen dados por la necesidad de superar las limitaciones técnicas introducidas por las redes eléctricas, como son las térmicas o las respectivas al voltaje. En otros términos, incrementan la potencia transmitida y aportan flexibilidad al permitir modificar de forma dinámica los parámetros eléctricos ante cambios en la configuración de la red.

Los FACTS incluyen todos los elementos electrónicos basados en tecnología de alta

potencia y que son empleados dentro de una SG para la transferencia de energía de CA y el control de la potencia reactiva. También, realizan tareas de reducción de impedancia en las líneas de transmisión y de optimización del factor de potencia y pueden actuar tanto a nivel individual como de forma coordinada con otros controladores.

La primera generación de FACTS emplea interruptores controlados por tiristores, mientras que la segunda tiene como base convertidores estáticos de conmutación. En la Tabla 2.1 se visualizan las tecnologías FACTS más relevantes. [23] [22]

Poniendo el enfoque en la garantía de estabilidad de una SG, uno de los sistemas más importantes de los expuestos en la Tabla 2.1 es el SVC [21]. Los compensadores estáticos son capaces de detectar grandes caídas de tensión resultantes a la producción de un cortocircuito o a la pérdida de líneas de transmisión en un área de la red. Esto es importante, ya que una detección rápida de una avería permite restaurar en un corto período de tiempo la tensión del área afectada sin que el problema escale a otras partes de la red.

En otros términos, aísla este área del resto de la red eléctrica. Además, el SVC asegura que el proceso de restauración se produzca paulatinamente para que los efectos producidos por el cortocircuito sean prácticamente imperceptibles en los puntos de carga del área afectada. En la Figura 2.11 se puede apreciar un ejemplo de instalación de un SVC en el municipio noruego Sylling y se encuentra conectada a un sistema de 420 kV.



Figura 2.11: Instalación de un SVC en Sylling (Noruega) [21]

Generación	Tipo	Descripción
1º	<i>Thyristor-Controlled Series Capacitor</i> (TCSC)	Controla el flujo de potencia tanto reactiva como activa por la línea de transmisión mediante el ajuste de impedancia en serie y amortigua las oscilaciones.
	<i>Static Var Compensator</i> (SVC)	Absorbe o suministra potencia reactiva según las necesidades de una línea de transmisión a través de la variación de la susceptancia en paralelo. Ayuda a mantener el voltaje estable en la red y proveen un aumento de la capacidad de transferencia de energía.
2º	<i>Static Synchronous Compensator</i> (STATCOM)	Compensa la potencia reactiva al igual que el SVC, pero en este caso empleando electrónica de potencia para proporcionar una respuesta más rápida.
	<i>Unified Power Flow Controller</i> (UPFC)	Combina las funciones de un TCSC y un SVC, teniendo la capacidad de controlar en una línea de transmisión tanto la impedancia en serie como la susceptancia en paralelo.
	<i>Static Synchronous Series Compensator</i> (SSSC)	Modifica dinámicamente la impedancia y es capaz de controlar la fase y la amplitud de la tensión en la línea.
	<i>Interline Power Flow Controller</i> (IPFC)	Conecta varias líneas de transmisión en paralelo. No obstante, modula la impedancia y la fase de la tensión de cada línea de transmisión de forma independiente, lo que permite operar sobre el flujo de potencia de una forma controlada. Mejora la capacidad de transmisión y reduce las pérdidas energéticas.

Tabla 2.1: Tecnologías FACTS de 1ª y 2ª generación

Siguiendo en el ámbito de la distribución energética, es importante también exponer las tecnologías enfocadas a la transmisión de electricidad a larga distancia, como son las Corriente Continua de Alto Voltaje (del inglés *High Voltage Direct Current* (HVDC)) [24]. Como su nombre indica, utilizan corriente continua para ello, lo que minimiza las pérdidas

en el proceso de transmisión respecto al caso de la Corriente Alterna de Alto Voltaje (del inglés *High Voltage Alternating Current* (HVAC)), además de permitir portar mucha más potencia.

Esto es fundamental cuando se pretende integrar al sistema global fuentes de energía ubicadas en áreas remotas, sobre todo dentro del contexto de las SGs. Generalmente, las fuentes de energías renovables, como pueden ser la solar o la eólica, se encuentran alejadas de los puntos geográficos con mayor demanda y se requiere una respuesta rápida ante cambios de carga para mantener la estabilidad del sistema.

Es decir, cuando se transmiten grandes cantidades de potencia a largas distancias con líneas de HVAC, se generan ángulos eléctricos entre los voltajes y las corrientes en la línea muy grandes que pueden producir oscilaciones descontroladas y por tanto, desencadenar inestabilidades que llegarán a escalar a lo largo del sistema.

Por ello, las tecnologías HVDC simplifican la interconexión entre microrredes y permiten una mayor integración de sistemas dedicados al almacenamiento energético para facilitar la gestión de los recursos disponibles. Como desventaja, se encuentra el alto coste que supone su implementación, ya que al final solamente se puede emplear en el caso de aplicaciones punto a punto.

En la Figura 2.12 se muestra como ejemplo una infraestructura de transmisión de energía eólica a larga distancia. Esta energía proviene de una fuente remota y se transporta a través del medio marino para llegar hasta los usuarios. Como se puede apreciar, es necesario implementar estaciones de conversión de alto voltaje de corriente alterna a continua y viceversa en los extremos de las líneas HVDC.



Figura 2.12: Representación de un sistema de transmisión de energía eólica con líneas HVDC [25]

2.1.3.3. Consumo

Entrando en detalle en las ubicaciones finales del sistema y por tanto, en el proceso de consumo, es imprescindible conocer la cantidad de energía demandada por los clientes que pertenecen a una SG. La instalación de sensores o medidores inteligentes (del inglés *smart meters*) en las viviendas posibilitan el registro constante de datos respectivos al consumo de energía, niveles de voltaje, corriente y factor de potencia. Estos son almacenados, analizados y procesados por las distribuidoras de energía para obtener a partir de los mismos la información necesaria sobre el comportamiento de los usuarios. [26]

Como se había introducido en el Apartado 2.1 el proceso de adquisición de datos por parte de los usuarios se trata de uno de los pilares más relevantes de cara a la optimización energética del sistema. Los objetivos principales que se pretenden con este proceso se engloban en la reducción del consumo, la optimización de la distribución y la maximización del beneficio. Respecto a este último, las compañías energéticas a partir de su base de clientes estudian las estrategias de categorización de los mismos para definir el sistema de fijación de precios dinámicos. [26]

No obstante, la gestión de los sensores finales en las SGs se caracteriza por su alta complejidad, debido a los grandes volúmenes de datos a manejar. Se requiere un procesamiento eficiente a tiempo real para analizar datos capturados de múltiples fuentes con el fin de evitar latencias o bloqueos en el sistema. Es por ello que se dificulta el empleo de herramientas convencionales de gestión de bases de datos y se requiere una solución más avanzada mediante la aplicación de tecnologías enfocadas al *Big Data*. Se profundizará más sobre ello en la Sección 2.4.

2.1.4. Protocolos

2.1.5. Topología de red

2.1.6. Seguridad en Smart grids

Como se ha introducido en la Sección referente a la figura del prosumidor (ver Sección 2.1.2), el DSO tiene la capacidad de operar ante problemas y fallos en la red tomando decisiones y dando una respuesta rápida. En determinados casos las incidencias se pueden dar por la configuración del propio sistema, si esta no se realiza de forma correcta y eficiente. No obstante, también es preciso contemplar la posibilidad de ataques externos que atenten contra el sistema de la SG.

Uno de los ataques más peligrosos en este ámbito es el de Inyección de Datos Falsos (del inglés *False Data Injection* (FDI)) [27] y, como su nombre determina, se basa en

la inyección de paquetes maliciosos en la red con el fin de bloquear los servicios de la misma y conseguir la autorización para realizar operaciones restringidas. El proceso se puede realizar actuando directamente a través de los sensores finales (*smart meters*) o secuestrando el canal de comunicación. Algunas técnicas empleadas por los atacantes son las siguientes [27]:

- **Fallo del dispositivo:** Se aplica un ataque de Denegación de Servicio (del inglés *Distributed Denial of Service* (DDoS)) a los sensores. Cabe destacar que generalmente los sensores que se encuentran en las ubicaciones finales tienen una capacidad limitada de conexión que derivaría en grandes latencias en la comunicación de datos con el resto de entidades de la red. En este momento el atacante se encarga de suplantar al host en cuestión para enviar los paquetes falsos en su nombre.
- **Cracking:** Se descifran las contraseñas para obtener acceso a los equipos mediante ingeniería social o fuerza bruta. De la misma forma que el anterior, la limitación de los recursos computacionales de los sensores provoca que no existan mecanismos de contraseñas lo suficientemente seguros. Además, la mayoría de ellos emplean como protocolos de comunicación Modbus/TCP o DNP 3.0/TCP, los cuales implementan una transmisión de texto en formato plano y sin cifrado. Por ello, un atacante puede tener la posibilidad de monitorear y capturar el tráfico si consigue acceso.
- **Envenenamiento de tablas *Address Resolution Protocol* (ARP) o *ARP Spoofing* [28]:** Consiste en enviar mensajes falsos ARP para monitorizar el tráfico y capturar o alterar los paquetes que se están transmitiendo por la red.

Teniendo esto en cuenta, es imprescindible definir los mecanismos de detección y mitigación a emplear para evitar que se produzcan ataques FDI [27]:

- **Mecanismo de autenticación estrictos:** Los datos que se suministran al DSO o a otras entidades de control de la SG deben de ser autenticados para comprobar su confiabilidad. Para ello, se emplean por ejemplo, marcas de tiempo para evitar ataques de repetición y protocolos de seguridad como *Transport Layer Security Protocol* (TLS) o *Secure Sockets Layer* (SSL), además de *Secure Hash Algorithm* (SHA) y *Hash Message Authentication Code* (HMAC).
- **Gestión dinámica de claves:** Se actualizan las claves con una determinada frecuencia para incrementar la seguridad de los estándares IEEE 802.11s y evitar ataques DDoS.
- **Empleo del protocolo *Secure Neighbor Discovery* (SEND) [29]:** Para evitar los ataques de *ARP Spoofing* se emplean pares de claves *Rivest, Shamir y Adleman* (RSA) para

poder garantizar en el proceso de enrutamiento que los mensajes que tienen como origen un determinado host pertenecen al mismo.

2.1.7. Estándares y regulación en smart grids

2.1.8. Beneficios ambientales

2.1.9. Situación en España

En el caso de España, nuestro país ocupa el quinto lugar en el mercado eléctrico europeo, solamente por detrás de Alemania, Francia, Reino Unido e Italia.

Además, en los últimos años ha crecido en consideración y está creciendo rápidamente. La demanda eléctrica estimada para el año 2001 fue de 210.4 mil millones de kilovatios-hora (bkwh), un aumento del 5 [30]

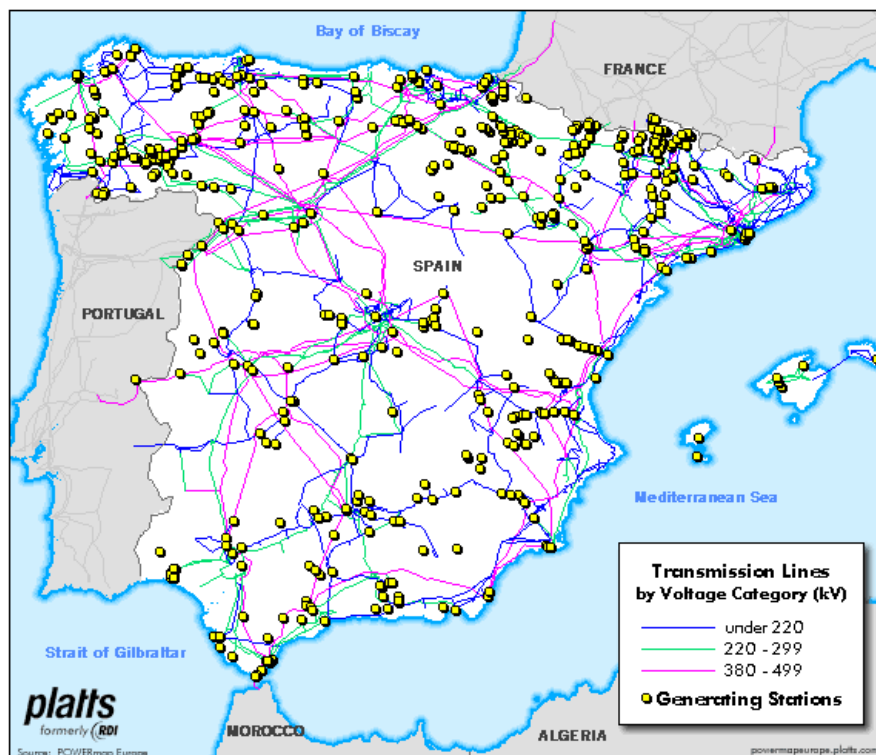


Figura 2.13: Red nacional de transmisión energética [30]

2.1.10. Desafíos futuros de las smart grids

2.1.10.1. Retos técnicos

2.1.10.2. Retos económicos

2.2. IoT

El Internet de las Cosas (del inglés IoT) [31] [32] se define como la integración de múltiples sensores o dispositivos físicos interconectados entre sí que se comunican a través de una red inalámbrica y recopilan información en tiempo real.

La evolución del IoT tal y como se conoce en la actualidad tiene sus inicios en los años 90. Sin embargo, en ese entonces estas tecnologías se enfrentaban con la problemática del gran tamaño que presentaban los chips, por lo que se produjo un avance lento de las mismas. A medida que se permitió la reducción del volumen los dispositivos electrónicos y se mejoró la eficiencia y la capacidad de computación de los chips, se facilitó el progreso del IoT.

En la actualidad, la integración de la tecnología 5G o de quinta generación móvil, ha supuesto el incremento de la velocidad de transmisión, procesamiento y análisis de los datos en los sistemas IoT. La capacidad de administrar a gran escala multitud de dispositivos físicos y la próxima implementación del 6G, depara un futuro en el que las tecnologías IoT estarán presentes en numerosos ámbitos y sectores.

Una de las grandes ventajas que proporcionan los sistemas IoT es que no es necesaria la intervención humana para posibilitar la transmisión y recepción de datos entre los elementos de la red. También, cabe destacar la reducción del coste de computación, la automatización de las tareas y la monitorización en tiempo real. No obstante, en cuanto a posibles desventajas, pueden presentar ciertas vulnerabilidades en cuestiones de privacidad y seguridad. Por ello, en la construcción de un sistema IoT se deben centrar esfuerzos en blindar los distintos elementos que lo componen frente a posibles ataques que puedan extraer información delicada o sabotear el funcionamiento.

En cuanto a la estructura de un sistema IoT, se pueden diferenciar cuatro componentes que contribuyen al funcionamiento del mismo [31] [32]:

- Dispositivos inteligentes: Se trata de cualquier elemento físico al que se le pueda asignar una dirección IP y que tenga como mínimo capacidad de computación para recopilar datos del entorno o del usuario y transmitirlos a través de la red.
 - Conectividad: En este campo entran varias posibilidades de conexión: Wi-Fi, Blue-
-

tooth, redes de área amplia y baja potencia (del inglés *Low Power Wide Area Networks* (LPWAN)), etc.

- Aplicación: Se compone de un conjunto de servicios de nube que se encargan de integrar y procesar todos los datos recopilados por los dispositivos del sistema IoT. Además, en algunos ámbitos, emplea técnicas de ML para analizar la información que recibe y optimizar la toma de decisiones en el sistema.
- Interfaz gráfica: Desde la misma, el usuario puede gestionar y controlar el conjunto de elementos de la red. La interfaz está conectada directamente a la nube y puede tratarse de una aplicación móvil o web.

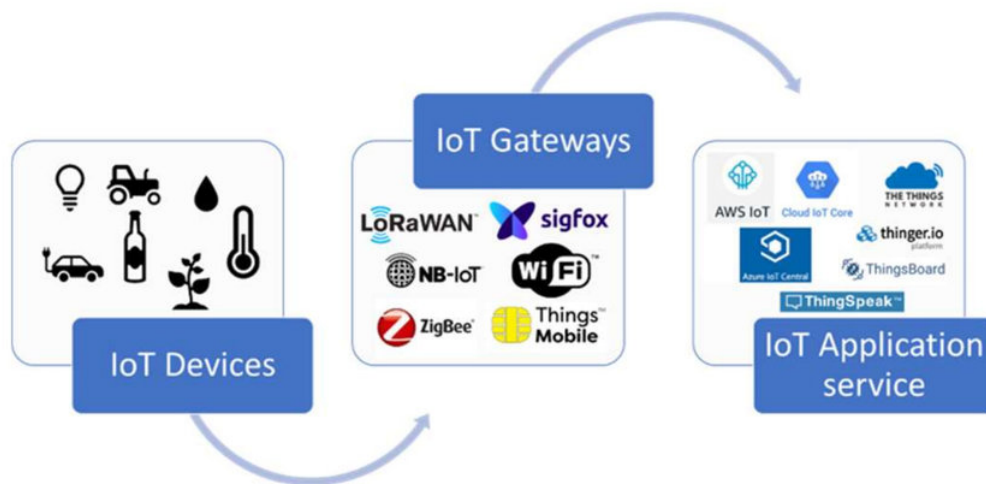


Figura 2.14: Infraestructura de componentes de un sistema IoT [33]

2.2.1. IoE

Considerando el contexto energético en el que se engloba este TFM y entrando con más detalle en el marco de las tecnologías IoT, es preciso destacar dentro de las mismas la subcategoría denominada como el Internet de la Energía (del inglés Internet of Energy (IoE)) [34]. Como concepto, el IoE comprende todo el paradigma de operación de los elementos que constituyen una red energética. Es decir, lleva a cabo la integración de todos los dispositivos, sensores o equipos informáticos en una misma estructura, la cual está basada en internet para poder monitorizar y controlar remotamente cada punto de la red.

Se puede expresar que, con la implementación de tecnologías basadas en IoE, se buscan

objetivos muy similares a las SGs. No obstante, el IoE va más allá y supone un control y una gestión de la red a mayor escala. En otros términos, no solo se constituye por los elementos de la propia infraestructura eléctrica, sino que también permite una gestión energética a nivel interna de los hogares, mediante la inclusión de electrodomésticos y otros dispositivos electrónicos.

2.2.2. Redes LLN

2.2.2.1. RPL

2.3. DEN2NE

aquí sería explicar el protocolo, esto en análisis o en estado del arte?

[3] [35]

2.4. Big Data

El *Big Data* [26] se puede definir como el manejo y el análisis de un conjunto extremadamente grande de datos, los cuales provienen de fuentes diferentes y no correlacionadas. Debido a su complejidad y volumen, no pueden ser procesados con software o herramientas tradicionales de gestión de bases de datos y se requieren soluciones tecnológicas más avanzadas. El concepto de *Big Data* se basa en tres componentes principales, denominadas como las 5 Vs [36] [37]:

- Variedad: Hace referencia a la heterogeneidad de la información. Existen distintos tipos de datos a procesar de diferentes fuentes, formas y resoluciones. Se puede simplificar su gestión si estos son estructurados (en forma de tablas de bases de datos) o dificultar, si estos no tienen una estructura definida (en formato de texto o imagen). También, pueden ser semiestructurados (en formato JSON o XML).
 - Velocidad: Es un parámetro crítico, ya que algunos entornos requieren de una toma de decisiones a tiempo real y los datos deben de ser generados, procesados y analizados rápidamente.
 - Volumen: Es imprescindible tener la capacidad de gestionar una gran cantidad de datos (en el orden de los TB) que además, es creciente en el tiempo. El almacenamiento y la minería de datos deben de ser eficientes para manejarlos de forma correcta
 - Veracidad: Mide la calidad y confiabilidad de la información que se adquiere. Se comprueba su autenticidad e integridad para asegurar que proviene de una fuente
-

legítima, por lo que se deben aplicar mecanismos de seguridad como la encriptación o herramientas de detección y mitigación de ataques a la red (ver Sección 2.1.6). También, se verifica que la información no contenga errores y que sea lo más precisa posible.

- Valor: La masividad de datos introduce mucho ruido y confusión en la información y debe de seleccionarse solamente aquella de utilidad aplicando minería de datos y procesamiento adicional.

Durante los últimos años el *Big Data* se ha adoptado en multitud de sectores y campos como en las telecomunicaciones, las finanzas, el comercio, la medicina, el transporte o la investigación científica. Como se ha expuesto en los apartados anteriores (ver Sección 2.1 y en particular, 2.1.3.3) y poniendo enfoque en los objetivos de este TFM (ver Sección 1.2), el *Big Data* tiene un gran potencial en el ámbito de las SGs.

En este contexto, la efectividad del empleo de la información reside en la búsqueda de la correlación entre los datos adquiridos de la red eléctrica y el resto de características que pueden influir en los mismos, como son el comportamiento de los usuarios finales, las condiciones de estabilidad, la disponibilidad de recursos, el estado de las cargas o las condiciones climáticas en un determinado instante temporal. La coordinación de las mediciones que se realizan en la red junto con la información que se dispone del entorno es crucial para la monitorización, el control y la operativa llevada a cabo dentro de la SG. [26]

En otros términos, un análisis efectivo del *Big Data* conducirá a una buena toma de decisiones en el sistema y en consecuencia, a su optimización. Este análisis se constituye del empleo de herramientas dedicadas al procesamiento distribuido, el almacenamiento en la nube, la minería de datos y a técnicas y algoritmos de aprendizaje automático (ML).

2.4.1. Computación en nube

En la Figura 2.15 se representa la infraestructura de procesos dedicados al *Big Data* en un ámbito de SGs. Como se puede visualizar, se compone de tres capas operativas: la superior está enfocada al almacenamiento y computación de datos, la intermedia, a la gestión, compartición e integración de datos de diferentes aplicaciones o fuentes y la inferior, a todos los procesos y técnicas que se encargan del procesamiento, minería, clusterización y clasificación final. [26]

Para implementar esta infraestructura es preciso hacer uso de un entorno de computación en nube. Dentro de este contexto, las grandes tecnológicas como Google, Amazon

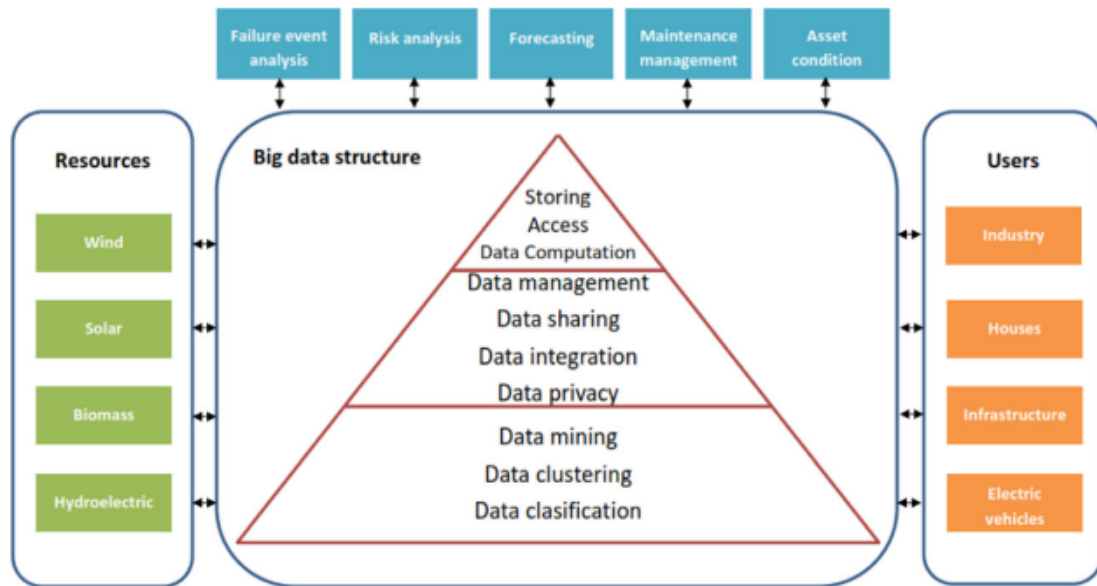


Figura 2.15: Infraestructura de *Big Data* enfocada al ámbito de las SGs [26]

y Microsoft han puesto sus esfuerzos en los últimos años en desarrollar y optimizar sus propios entornos de nube con el fin de integrar toda la gestión y análisis del *Big Data* en una plataforma. A modo de aportar una mayor comprensión de su funcionamiento, se representa en la Figura 2.16 un esquema de los servicios de nube disponibles.

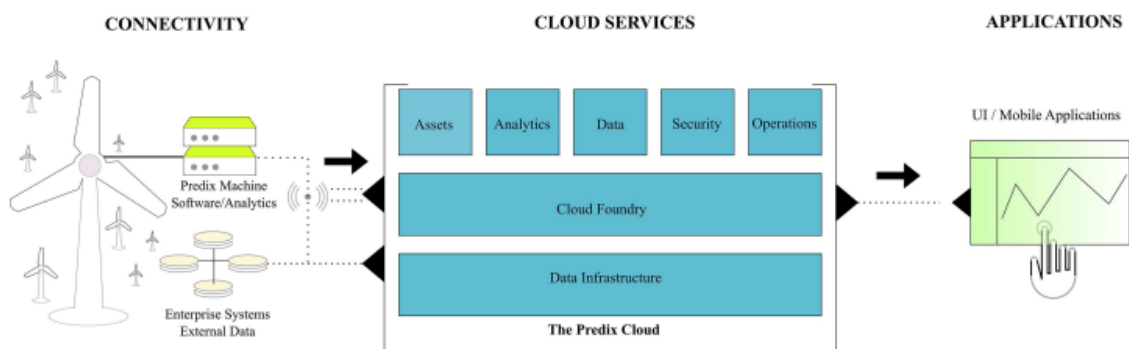


Figura 2.16: Plataforma de análisis de *Big Data* en el entorno de SGs [38]

La computación en la nube se basa en el concepto de virtualización, el cual se puede definir como la tecnología que permite la creación de diferentes entornos virtuales aislados a partir de los recursos hardware disponibles y con independencia de este. Los proveedores de servicios emplean la virtualización para crear múltiples Máquinas Virtuales (VMs) y

ofrecer recursos, como pueden ser el almacenamiento, la potencia de procesamiento o las aplicaciones de la nube, como diferentes servicios virtualizados. Esto aporta las siguientes ventajas a la computación en nube [38] [39]:

- Escalabilidad: Se pueden aumentar o reducir recursos de la nube según las necesidades en cada instante temporal. Los usuarios consiguen un autoprovisionamiento de los recursos, ya que pueden desplegar instancias virtuales cuando lo deseen de forma manual o automática.
- Flexibilidad: Se ofrecen muchas posibilidades de configuración de los servicios de la nube.
- Interoperabilidad: Se emplean Interfaces de Programación de Aplicaciones (del inglés *Application Programming Interface* (API)) para integrar plataformas y aplicaciones heterogéneas. El objetivo es soportar la comunicación entre las mismas y que se produzca una interpretación de los datos de forma correcta. Por ello, también se estandarizan los protocolos para que sean comunes.
- Seguridad: Las plataformas de servicios de nube ofrecen un gran nivel de seguridad e implementan medidas como la encriptación de los datos y restricción de acceso o de permisos. Además, el aislamiento de los recursos y de las aplicaciones en diferentes VMs permite coconfigurar e implementar diferentes políticas de seguridad según las necesidades.
- Ahorro de costes: Se implementa un modelo de pago por uso en función de los recursos que se han consumido y se evita la inversión inicial necesaria para la instalación de una infraestructura física, además de su mantenimiento en el tiempo (CAPEX y OPEX).
- Disponibilidad: Se garantiza el acceso a los servicios y la confiabilidad de los mismos. Para ello, implementa técnicas de redundancia y aporta una buena tolerancia a fallos.

2.5. Machine Learning/IA

[2]

2.6. Herramientas software

2.6.1. BRITE

BRITE [4] se presenta como una plataforma dedicada a la generación de topologías de red. Fue desarrollada en la Universidad de Boston y se caracteriza por su gran flexibilidad, ya que enfoca su arquitectura en el concepto de modelo topológico. Es decir, soporta varios modelos diferentes y cada uno de ellos viene determinado por los parámetros de entrada que se definen. Es por ello que BRITE sigue la siguiente secuencia de acciones para diseñar las topologías:

1. Definición del posicionamiento de los nodos:
2. Nivel de Interconexión de los nodos y configuración de enlaces:
3. Asignación de atributos y características de la red y de los dispositivos: Se determina el delay y el ancho de banda de los enlaces.
4. Especificación del formato de salida: Se indica un formato `.brite` para todas las topologías generadas.

Como motivo de facilitar el empleo de la herramienta, se incluirá una Sección dedicada al proceso de instalación, configuración y ejecución de la herramienta en el Anexo correspondiente a los manuales de usuario (ver Sección C.1). Además, de la infse podrá acceder a más información desde el repositorio¹ del equipo de investigación NetIS de la UAH.

2.6.1.1. Definición de topologías

La herramienta BRITE basa el proceso de creación de topologías en la definición de los siguientes parámetros de entrada en un fichero con formato `.conf`:

- *Name*: Modelo de la topología.
- *N*: Número de nodos de la topología.
- *HS* y *LS*: Dimensiones del plano. Respectivamente, hacen referencia a la longitud total del plano cuadrado y al tamaño de los cuadros interiores.
- *Node Placement*: Posicionamiento de los nodos. Se puede producir de forma totalmente aleatoria o creando zonas a lo largo del plano con mayor concentración de nodos.

¹<https://github.com/NETSERV-UAH/BRITE>

- *Growth Type*: Método de introducción de los nodos en la topología. Se puede realizar este proceso de forma incremental (uno a uno) o de forma aleatoria (todos a la vez).
- *m*: Número de enlaces por nodo o número de nodos vecinos a los que se conectará un nuevo nodo al unirse a la red. BRITE puede crear enlaces unidireccionales o bidireccionales y, respectivamente, la topología generada tendría un grado m o $2m$.
- *Alpha*, *Beta*: Parámetros específicos para topologías basadas en el modelo Waxman.
- *BWDist*, *BWMin*, *BWMax*: Ancho de banda de los enlaces.

2.6.1.2. Modelos de topologías

Como se ha introducido, con BRITE se posibilita el uso de múltiples modelos para crear topologías. En la Figura 2.17 se representan todos los que soporta la herramienta. No obstante, este TFM se va a enfocar en la generación de topologías a nivel de router, como son los modelos Router Waxman y Router Barabasi-Albert.

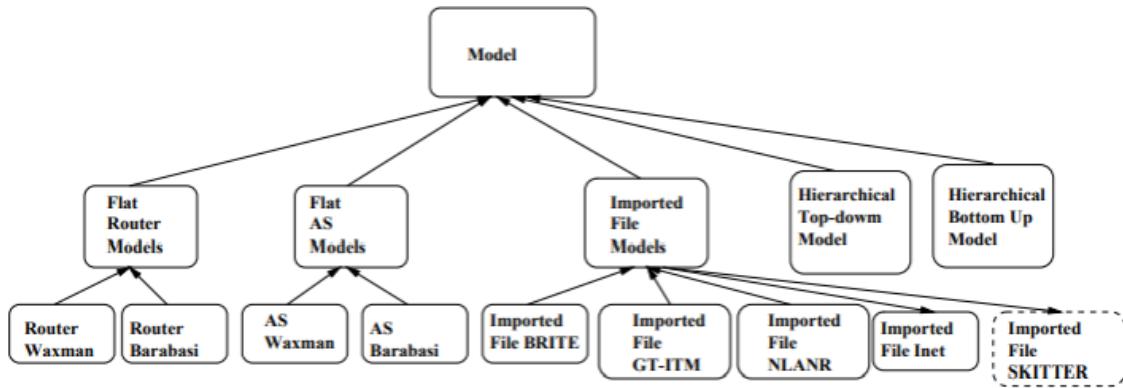


Figura 2.17: Modelos soportados por BRITE [4]

En el caso del Router Waxman, como su nombre indica, emplea un modelo de probabilidad Waxman para establecer la Interconexión de los nodos en la topología [40]:

$$P_{\text{Waxman}}(u, v) = \alpha \cdot e^{-\frac{d}{\beta \cdot L}} \quad (2.1)$$

Donde:

$P(u, v)$ es la probabilidad en función de la distancia euclidiana entre un nodo u y un nodo v de la red.

α es un parámetro específico del modelo que hace referencia a la densidad de enlaces y toma generalmente un valor igual a 0.2.

β es un parámetro específico del modelo que hace referencia al ratio enlaces largos/enlaces cortos en la topología y toma generalmente un valor igual a 0.15.

L es la máxima distancia entre dos nodos cualesquiera.

Por otro lado, el modelo Router Barabasi-Albert está basado en la generación de topologías con un incremento exponencial del número de nodos a lo largo del tiempo. Además, se permite una conexión preferencial, suponiendo que cuanto mayor grado de conectividad abarque un nodo, mayor será la probabilidad de que este añada nuevos enlaces.

Por lo tanto, el plano de la topología comienza con un número de nodos inicial N_0 y se van añadiendo los demás uno a uno. Cada uno de estos nuevos nodos se conectará a N nodos ya añadidos a la topología con una probabilidad [40]:

$$P_{\text{Barabási-Albert}}(k_i) = \frac{k_i}{\sum_j k_j} \quad (2.2)$$

Donde:

$P(k)$ es la probabilidad de conexión del nodo i con grado k .

$\sum_j k_j$ es el sumatorio de los grados de todos los nodos de la topología.

2.6.1.3. Automatización de la ejecución

A modo de facilitar la ejecución de la herramienta BRITE se aporta en el repositorio el fichero de python *generador_brite.py*, dedicado a la automatización del proceso de creación de un fichero de configuración. Este recibirá los parámetros de entrada que caracterizarán a la topología a generar (ver Sección 2.6.1.1) y escribirá el nuevo fichero.

También, se añade el fichero de python *parser.py*, que se encargará de definir la función de transformación del archivo de salida proporcionado por BRITE (en formato .brite) en dos nuevos ficheros: *Nodos.txt* y *Enlaces.txt*. Respectivamente, estos almacenarán las posiciones x e y de los nodos en el plano y la información sobre las distancias y los identificadores de los nodos que se interconectan con cada enlace.

Teniendo en cuenta estos ficheros, se proporciona un script *autogenerador.sh* donde se incluyen todas sus funcionalidades y se automatiza todo el proceso de ejecución de la herramienta. Este script sigue la siguiente secuencia de pasos:

1. Define los valores de cada uno de los parámetros de entrada
2. Ejecuta el fichero de python *generador_brite.py* aplicando los parámetros definidos.
3. Genera 10 escenarios diferentes para cada archivo de configuración a partir de semillas distintas (*seed_files*).
4. Ejecuta el fichero *parser.py* para obtener los ficheros *Nodos.txt* y *Enlaces.txt* a partir de cada topología generada.

En el caso del tercer paso, es preciso detallar que en el repositorio se aporta un directorio con 10 ficheros de semilla para posibilitar la generación de 10 topologías distintas a partir de unos mismos parámetros de entrada.

Para ejecutar múltiples pruebas sobre el algoritmo DEN2NE como motivo de este TFM, se generarán con BRITE topologías basadas en los modelos Router Waxman y Router Barabasi-Albert. En el script *autogenerador.sh* se define una introducción incremental de los nodos al plano (*GrowthType=1*), varios grados diferentes de conectividad ($m=1,2,3$) y unos valores de α y β de 0.2 y 0.15 respectivamente para el modelo Waxman.

Los nodos se incrementarán de 10 en 10, suponiendo topologías que pueden abarcar desde 10 hasta 200 nodos en total. Teniendo en cuenta el número de semillas (10), los modelos empleados (2), los grados de conectividad configurados (3) y los números de nodos posibles a fijar (20), se puede cuantificar que el número resultante de topologías distintas generadas con BRITE será igual a 1200.

2.6.1.4. Graficación de topologías

3. Diseño y análisis de datos

Este Capítulo abordará la fase del TFM que se centra en el análisis, diseño y procesado de las fuentes de datos. Se abrirá el Capítulo con una Sección enfocada al estudio de la disponibilidad de fuentes de datos de implementaciones reales de SGs. Por ello, será imprescindible llevar a cabo una investigación exhaustiva y un posterior análisis de su utilidad en referencia a las necesidades que presenta este TFM (ver Sección 3.1).

Por otro lado, una vez expuestas y justificadas las fuentes de datos seleccionadas, se añadirá una Sección dedicada a la definición detallada de las acciones que serán necesarias para llevar a cabo el procesamiento de los datos (ver Sección 3.2). Finalmente, se obtendrá un conjunto de datos reducido con la información útil sobre la que entrenar los modelos de ML a desarrollar.

3.1. Análisis de los conjuntos de datos

En ámbitos como el Procesado de Lenguaje Natural (del inglés *Natural Language Processing* (NLP)) o el reconocimiento facial, la disponibilidad de numerosas fuentes de datos ha sido imprescindible para acelerar el desarrollo de técnicas de minería de datos y de aprendizaje automático (ML).

A diferencia de estos campos, el contexto energético experimenta un menor grado de disponibilidad de datasets. La protección de la privacidad de los usuarios en el proceso de medición y análisis de su comportamiento energético supone que exista un menor número de datasets públicos enfocados a implementaciones reales. No obstante, la búsqueda de la optimización de la distribución energética y de la reducción del consumo de los usuarios, aparte de otras motivaciones ambientales, ha instado en los últimos años a múltiples ingenieros y científicos a lo largo del mundo a crear datasets públicos enfocados a la investigación (ver Tabla 3.1).

En este contexto han cobrado gran importancia las técnicas de Monitorización de Cargas no Intrusiva (del inglés *Non-Intrusive Load Monitoring* (NILM)). Como se ha introducido en la Sección 2.2.1, cada vez se produce una mayor integración de las tecnologías de IoE en el ámbito de las SGs para obtener una mayor monitorización del comportamiento

energético en los hogares. Es por ello que el fin principal de las técnicas de NILM reside en establecer una desagregación energética para poder estimar de una forma precisa el consumo individual de cada uno de los dispositivos y electrodomésticos que hay en una vivienda. Para ello, se instalan medidores en cada circuito y después, se analizan, tanto a nivel interno como a nivel externo de la vivienda, los cambios de los parámetros eléctricos en cada instante temporal. [41] [42]

Teniendo esto en cuenta, se puede expresar que las técnicas de NILM se caracterizan por su bajo coste y por su facilidad y flexibilidad de despliegue. No obstante, el proceso de medición y recolección de datos puede llegar a requerir mucho tiempo y esfuerzo. Por ello, en el contexto de la investigación es importante que los datos tengan una disponibilidad pública para progresar en el desarrollo de técnicas de minería de datos y de aprendizaje automático en el ámbito energético.

En la Tabla 3.1 se exponen los múltiples datasets residenciales que han sido estudiados en esta fase de diseño, junto con información relativa a la implementación real sobre la que se ha basado cada uno de ellos, como es la ubicación de la misma, el número de viviendas total, el lapso temporal que comprende el despliegue, la frecuencia de las muestras tomadas y los parámetros eléctricos medidos. En virtud de las motivaciones y objetivos del presente TFM, es importante tener en cuenta ciertos requisitos para evaluar los conjuntos de datos disponibles y seleccionar el más apropiado:

- Cantidad: En primera instancia, es imprescindible revisar la cantidad de datos que aporta cada conjunto. En otros términos, para llevar a cabo un análisis estadístico del comportamiento energético de una ubicación se requiere partir de un conjunto de datos que agrupe las mediciones de un gran número de edificios residenciales. Adicionalmente, este análisis debe comprender extensos períodos temporales de medición para visualizar de forma correcta el comportamiento de los usuarios y predecir patrones de consumo a lo largo del tiempo.
- Calidad: Se debe evaluar la calidad de los datos recogidos, la cual responde en parte con la resolución de las medidas que se toman. Como se puede observar en la Tabla 3.1, algunos datasets como *REDD* y *BLUED* se centran en monitorizar los edificios residenciales a una frecuencia de muestreo alta. En el contexto de las tecnologías de NILM, esto aporta una vista más representativa del comportamiento energético a nivel interno en una vivienda y permite una desagregación energética más precisa. Por lo tanto, cuanto mayor sea la frecuencia de adquisición de medidas, mayor será la resolución de los datos.

- **Ubicación:** En términos de la ubicación de la implementación real sobre la que se han adquirido las mediciones, es de importancia conocer a la hora de analizar los datos las diferencias de voltaje que se manejan en cada país. Por ejemplo, en el caso de datasets como *BLUED* o *Smart**, los cuales provienen de ciudades de Estados Unidos, trabajarán a tensiones menores de 120V, mientras que otros como *ECO* o *GREEND*, cuyos datos se han adquirido en países europeos, lo harán con tensiones de hasta 230V. [42] [43]
- **Parametrización:** Por lo general, un dataset dedicado a las tecnologías de NILM estará constituido por una colección de muestras de voltaje (V), corriente (I) y potencia activa (P), reactiva (Q) y aparente (S). Cada una de estas muestras vendrá asociada a una marca de tiempo y a un identificador que haga referencia a la vivienda o medidores a los que corresponde. Adicionalmente, algunos de los datasets que se exponen en la Tabla 3.1, como son *HUE* o *SustDataED*, también incluyen información relativa a parámetros ambientales. Esto, sobre todo en el contexto de las SGs, proporciona información fundamental sobre el impacto de las condiciones climáticas en el comportamiento de los usuarios.
- **Generación fotovoltaica:** Considerando el contexto de SGs en el que se engloba este TFM, es de vital importancia seleccionar un conjunto de datos que aporte información relativa a la generación de energía a través de fuentes renovables, como ocurre en el caso de *Smart** o *SustDataED*. En relación con esto, se expondrá en la Sección 3.1.4 el estudio de una plataforma específica dedicada a la adquisición de datos de generación energética.

Como se ha introducido en el Capítulo 1, el objetivo principal que se pretende con este TFM es desarrollar técnicas de ML para identificar y predecir de forma precisa los errores que se pueden producir en una SG durante el proceso de distribución energética. A partir de ello, se expresa la necesidad de seleccionar un conjunto de datos en el que se pueda analizar el comportamiento de múltiples usuarios, tanto de consumo como de producción, durante un extenso período de tiempo.

De igual manera, es importante que las muestras tengan una buena resolución y que proporcione información en cada instante sobre las condiciones climáticas de la ubicación donde se encuentran las viviendas. Bajo las premisas anteriores, se ha finalizado el proceso de estudio y análisis de los diferentes conjuntos de datos llegando a la conclusión de que el dataset más adecuado a emplear en este TFM es *SustDataED*.

Nombre	Localización	Nº de residencias	Período de medición (días)	Resolución de medición	Parámetros
<i>AMPds2</i> [44]	Vancouver (Canadá)	1	730	1min	I, V, P, S, F, pf
<i>BLUED</i> [45]	Pittsburg (Estados Unidos)	1	8	12KHz	I, V, eventos de switch
<i>ECO</i> [46]	Thun (Suiza)	6	244	1Hz	P
<i>GREEND</i> [42]	Italia y Austria	9	310	1Hz	P
<i>HUE</i> [47]	British Columbia, Canada	28	60	1Hz	P
<i>iAWE</i> [48]	Nueva Delhi (La India)	1	73	1Hz	V, I, P, S
<i>REDD</i> [49]	Boston (Estados Unidos)	6	119	15KHz	I, V, P
<i>Smart*</i> [50]	Massachusetts (Estados Unidos)	3	90	1min	P, S, V, I
<i>SustDataED</i> [51]	Madeira (Portugal)	50	1144	1min	I, V, P, Q, S
<i>UK-DALE</i> [52]	Reino Unido	5	499	16KHz	P, estado de switch

Tabla 3.1: Comparación entre datasets públicos en el ámbito NILM [42] [53] [54] [43]

3.1.1. *SustDataED*

El dataset *SustDataED* [51] es creado a partir del proyecto de investigación *Sustainable Interactions with Social Networks, context Awareness and Innovative Services* (SINAIS), dedicado al diseño, implementación y despliegue de sistemas NILM de bajo coste. Surge con el objetivo de proporcionar una retroalimentación ecológica a los usuarios para fomentar un comportamiento energético sostenible y un mayor uso de las fuentes de energía renovables.

El dataset *SustDataED* engloba cinco años de datos de consumo y producción energética de 50 hogares de la ciudad de Funchal (Madeira, Portugal) con una resolución o frecuencia de adquisición de nuevas muestras por minuto. Debido a esto, se caracteriza por comprender una gran cantidad de información eléctrica que tendrá que ser posteriormente procesada de forma exhaustiva (ver Sección 3.2).

Pese a la necesidad de procesamiento que conlleva, el uso de *SustDataED* en el presente TFM brinda ciertas ventajas. Entre otras, aporta un gran perspectiva del comportamiento eléctrico a largo plazo de múltiples hogares con diferentes características y ofrece un enfoque medioambiental, ya que mediante la adquisición de los datos relativos a las condiciones climáticas posibilita el análisis de la correlación de las mismas con el consumo de los usuarios y la producción de energía.

3.1.2. Estructura del dataset *SustDataED*

El proceso de adquisición y recolección de los datos de los 50 hogares que constituyen el dataset se divide en cuatro despliegues diferentes (ver Figura 3.1):

- Los despliegues 1 y 2 se pueden tratar en conjunto, ya que abarcan las mismas viviendas, suponiendo un total de 23. El despliegue 1 se inicia en julio de 2010 y finaliza en el mes de noviembre de ese mismo año cuando se reinstala una versión revisada del sistema de realimentación. En este momento, comienza el despliegue 2, que durará hasta abril del 2012.
 - El despliegue 3 se implementa en un total de 17 hogares desde el mes de agosto de 2012 hasta enero de 2013. Se introducen novedades en el sistema de adquisición de datos para reducir el número de sensores a instalar, optimizando la arquitectura de monitorización.
 - El despliegue 4 es el último que se realiza y consta de la agrupación de medidas de 10 hogares, comenzando la recolección de datos de los mismos en el mes de julio de 2013 y finalizando en marzo de 2014.
-

En la Figura 3.1 se permite visualizar de una forma gráfica los períodos temporales que abarcan cada uno de los despliegues expuestos anteriormente. Como se detallará a continuación, a modo organizativo, el dataset *SustDataED* clasifica en varias tablas los datos en función del tipo de información al que se hace referencia.

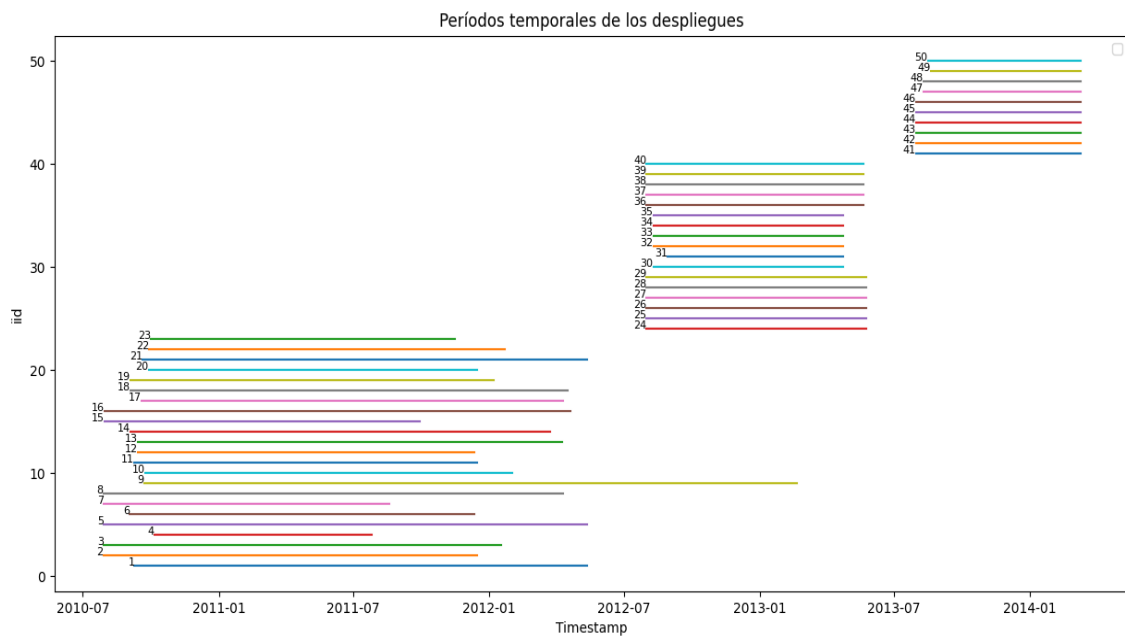


Figura 3.1: Representación del período temporal que comprende el proceso de recolección de datos para cada hogar

3.1.2.1. Medidas de consumo de energía

En la Tabla 3.2 se listan los campos que han sido definidos para la recolección de medidas de consumo de cada uno de los hogares. El dataset asocia los parámetros eléctricos a dos identificadores: uno para describir la vivienda de la que se están recogiendo los datos, y otro, para el despliegue al que pertenece. También, se almacena para cada medida una marca de tiempo con la fecha, hora, minuto y segundo en la que se ha recogido la información y se incluye el parámetro *miss_flag* para indicar si para un instante temporal no se han adquirido datos de forma correcta, suponiendo filas vacías en el dataset.

Los parámetros de tensión (V) y corriente (I) se muestrean continuamente a una frecuencia de 8KHz y las medidas de potencia activa (P), reactiva (Q) y aparente (S) se cuantifican a un ratio de 50 muestras por segundo a partir de las siguientes expresiones:

$$S = I_{\text{RMS}} \cdot V_{\text{RMS}}$$

$$P = S \cdot \cos(\phi)$$

$$Q = S \cdot \sin(\phi)$$

Donde:

I_{RMS} es la corriente eficaz (RMS).

V_{RMS} es el voltaje eficaz (RMS).

ϕ es el ángulo de fase entre la corriente y el voltaje instantáneos.

No obstante, como se había introducido en el Apartado 3.1.1 y se había definido en la Tabla 3.1, todas estas medidas, una vez son recogidas, se almacenan en el dataset a una resolución de una muestra por minuto.

Campo	Descripción	Unidades
<i>home_id</i>	Identificador de vivienda monitorizada	-
<i>timestamp</i>	Instante temporal de medida	datetime
<i>deploy</i>	Identificador del despliegue	-
<i>Imin</i>	Corriente mínima	A
<i>Imax</i>	Corriente máxima	A
<i>Iavg</i>	Corriente media	A
<i>Vmin</i>	Tensión mínima	V
<i>Vmax</i>	Tensión máxima	V
<i>Vavg</i>	Tensión media	V
<i>Pmin</i>	Potencia activa mínima	W
<i>Pmax</i>	Potencia activa máxima	W
<i>Pavg</i>	Potencia activa media	W
<i>Qmin</i>	Potencia reactiva mínima	VAR
<i>Qmax</i>	Potencia reactiva máxima	VAR
<i>Qavg</i>	Potencia reactiva media	VAR
<i>Smin</i>	Potencia aparente mínima	VA
<i>Smax</i>	Potencia aparente máxima	VA
<i>Savg</i>	Potencia aparente media	VA
<i>PFmin</i>	Factor de potencia mínimo	-
<i>PFmax</i>	Factor de potencia máximo	-
<i>PFavg</i>	Factor de potencia medio	-
<i>miss_flag</i>	Flag de valores vacíos	-

Tabla 3.2: Tabla de medidas de consumo energético [51]

3.1.2.2. Medidas de eventos de potencia

La información sobre los eventos de potencia se extrae de las medidas en crudo de potencia activa que, como se ha expresado en el apartado anterior, se producen a una frecuencia de muestreo de 50Hz. La Tabla 3.4 se dedica al proceso de detección de dichos eventos o, en otros términos, de identificación de cambios en la carga total de una vivienda.

En este caso, se emplea un método estadístico que calcula la probabilidad de que se produzca un cambio del valor de potencia media entre dos instantes. Para ello, el algoritmo trabaja con dos ventanas temporales: una de detección, que cuantifica la probabilidad del cambio en una muestra determinada y otra, de búsqueda de los valores mínimos y máximos de potencia para definir las posiciones de los eventos en la señal de potencia agregada. En referencia a esto, viene definido un cambio mínimo de potencia real de 30 W.

Campo	Descripción	Unidades
<i>home_id</i>	Identificador de vivienda monitorizada	-
<i>timestamp</i>	Instante temporal de medida	datetime
<i>deploy</i>	Identificador del despliegue	-
<i>delta_P</i>	Cambio real de potencia activa	W
<i>delta_Q</i>	Cambio real de potencia reactiva	VA
<i>trace_P</i>	Traza real de potencia activa del evento	W
<i>trace_Q</i>	Traza real de potencia reactiva del evento	VAR

Tabla 3.3: Tabla de medidas de eventos de potencia [51]

3.1.2.3. Medidas de producción de energía

El dataset *SustDataED* recoge a nivel global los datos respectivos a la producción de electricidad en intervalos de 15 minutos. Las fuentes energéticas se desagregan según su fuente de procedencia, dando protagonismo a las fuentes renovables. En la Tabla ?? se indican los campos que han sido definidos para cuantificar las medidas.

Campo	Descripción	Unidades
<i>timestamp</i>	Instante temporal de medida	datetime
<i>total</i>	Producción total	MWh
<i>thermal_fuel</i>	Electricidad producida por fuentes térmicas	MWh
<i>hydro</i>	Electricidad producida por fuentes hidroeléctricas	MWh
<i>eolic</i>	Electricidad producida por fuentes eólicas	MWh
<i>photovoltaic</i>	Electricidad producida por fuentes fotovoltaicas	MWh

Tabla 3.4: Tabla de medidas de producción de energía [51]

3.1.2.4. Medidas demográficas

Para el análisis del comportamiento eléctrico de cada uno de los hogares monitorizados se puede visualizar en la Tabla 3.5 cómo se almacena la información respectiva a sus inquilinos y a los períodos de adquisición de datos de cada vivienda.

Campo	Descripción	Unidades
<i>home_id</i>	Identificador de vivienda monitorizada	-
<i>building_id</i>	Identificador de edificio	-
<i>begin_monitoring</i>	Fecha de inicio de medición	datetime
<i>end_monitoring</i>	Fecha de fin de medición	datetime
<i>begin_feedback</i>	Fecha de inicio de realimentación en el despliegue	datetime
<i>end_feedback</i>	Fecha de fin de realimentación en el despliegue	datetime
<i>type</i>	Tipo de vivienda	-
<i>bedrooms</i>	Número de habitaciones	-
<i>adults</i>	Número de adultos inquilinos	-
<i>children</i>	Número de niños inquilinos	-
<i>contracted_power</i>	Potencia contratada	kWh

Tabla 3.5: Tabla de medidas de producción de energía [51]

De forma adicional, el dataset *SustDataED* incluye otra tabla con datos complementarios sobre las personas participantes en los despliegues, como son el género, edad, educación y ocupación.

3.1.2.5. Medidas de eventos de usuario

Una de las ventajas que proporciona la monitorización del comportamiento eléctrico de un hogar es la realimentación de la información que se produce hacia los mismos usuarios. El dataset recoge en forma de eventos las interacciones que se llevan a cabo con este sistema de realimentación a través de los campos determinados en la Tabla 3.6.

Campo	Descripción	Unidades
<i>home_id</i>	Identificador de vivienda monitorizada	-
<i>timestamp</i>	Instante temporal de medida	datetime
<i>deploy</i>	Identificador del despliegue	-
<i>type</i>	Tipo de interacción	-
<i>view_id</i>	Identificador de pantalla de visualización	-
<i>view_name</i>	Nombre de pantalla de visualización	-

Tabla 3.6: Tabla de medidas de eventos de usuario [51]

3.1.2.6. Medidas de condiciones ambientales y climáticas

Como se ha introducido al comienzo de la Sección 3.1.1, el dataset aporta un enfoque medioambiental a los procesos de consumo y producción energética al incluir la recolección de los datos relativos a las condiciones climáticas que se dan en cada instante. En la Tabla 3.7 se muestran los parámetros que se tienen en cuenta para ello.

Campo	Descripción	Unidades
<i>timestamp</i>	Instante temporal de medida	datetime
<i>temperature</i>	Temperatura exterior	°C
<i>humidity</i>	Humedad relativa	%
<i>pressure</i>	Presión relativa	hPa
<i>wind_dir</i>	Dirección del viento	-
<i>wind_speed</i>	Velocidad del viento	kmh
<i>precipitation</i>	Niveles de precipitación	mm
<i>events</i>	Eventos relevantes	-
<i>conditions</i>	Condiciones del cielo	-

Tabla 3.7: Tabla de medidas de condiciones ambientales [51]

3.1.3. Conclusiones del análisis del dataset *SustDataED*

Una vez realizado un estudio y un análisis en profundidad de los datos proporcionados por el dataset *SustDataED* se llega a una serie de conclusiones a tener en cuenta antes de iniciar la fase de preprocesamiento.

3.1.4. Global Solar Atlas

3.2. Preprocesado

4. Desarrollo y evaluación del modelo

4.1. Selección de modelos

esto aquí o en análisis?

4.2. Modelo 1

5. Conclusiones y trabajo futuro

Bibliografía

- [1] Repsol, “¿Qué son las Smart Grids o redes inteligentes? | Repsol,” <https://www.repsol.com/es/energia-futuro/tecnologia-innovacion/smart-grids/index.cshtml>, [Accessed 27-10-2023].
- [2] D. Ingeniería, “El impacto de las Smart Grids y la evolución de la Inteligencia Artificial en las redes eléctricas,” <https://es.linkedin.com/pulse/el-impacto-de-las-smart-grids-y-la-evoluci%C3%B3n-inteligencia-1f>, 2023, [Accessed 27-10-2023].
- [3] D. Carrascal, E. Rojas, J. A. Carral, I. Martinez-Yelmo, and J. Alvarez-Horcajo, “Topology-aware scalable resource management in multi-hop dense networks [Unpublished],” 2023.
- [4] A. Medina, A. Lakhina, I. Matta, and J. Byers, “BRITE: an approach to universal topology generation,” in *MASCOTS 2001, Proceedings Ninth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, 2001, pp. 346–353.
- [5] T. Vijayapriya, “Smart grid: An overview,” *Smart Grid and Renewable Energy*, vol. 02, pp. 305–311, 01 2011.
- [6] I. S. Grid, “IEEE Smart Grid 2015 Annual Report,” <https://smartgrid.ieee.org/about-ieee-smart-grid>, 2015, [Accessed 10-01-2024].
- [7] U. D. of Energy, “2009 Smart Grid System Report (July 2009),” <https://www.energy.gov/oe/articles/2009-smart-grid-system-report-july-2009>, 2009, [Accessed 14-01-2024].
- [8] I. Futura, “Smartgrids, transacciones de energía basadas en blockchain,” <https://iotfutura.com/2018/12/smargrid-basada-en-iot-y-blockchain/>, 2018, [Accessed 10-01-2024].
- [9] H.-P. c. Roy Pratt, “Resisting Temptation: the Seven Sins of Smart Grid,” <https://www.power-grid.com/news/resisting-temptation-the-seven-sins-of-smart-grid/#gref>, 2010, [Accessed 10-01-2024].

-
- [10] W. Lou, S. Zhu, J. Ding, T. Zhu, M. Wang, L. Sun, F. Zhong, and X. Yang, “Transactive demand–response framework for high renewable penetrated multi-energy prosumer aggregators in the context of a smart grid,” *Applied Sciences*, vol. 13, p. 10083, 09 2023.
- [11] Iberdrola, “DSO, ¿cómo transformar la gestión de redes hacia un modelo más inteligente?” <https://www.iberdrola.com/innovacion/operadores-de-sistemas-de-distribucion>, [Accessed 14-01-2024].
- [12] G. Sun, S. Shen, S. Chen, Y. Zhou, and Z. Wei, “Bidding strategy for a prosumer aggregator with stochastic renewable energy production in energy and reserve markets,” *Renewable Energy*, vol. 191, pp. 278–290, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960148122005274>
- [13] J. Iria and F. Soares, “An energy-as-a-service business model for aggregators of prosumers,” *Applied Energy*, vol. 347, p. 121487, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0306261923008516>
- [14] Z. M. Bakare Mutiu Shola, Abdulkarim Abubakar, “A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction,” *Energy Informatics*, vol. 6, 03 2023.
- [15] L. Tao and Y. Gao, “Real-time pricing for smart grid with distributed energy and storage: A noncooperative game method considering spatially and temporally coupled constraints,” *International Journal of Electrical Power & Energy Systems*, vol. 115, p. 105487, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0142061519317867>
- [16] A.-H. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, “Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid,” *Smart Grid, IEEE Transactions on*, vol. 1, pp. 320 – 331, 01 2011.
- [17] “Overview of CHP (Combined Heat & Power),” <https://www.a1energy.net/blogs/post/Overview-of-CHP-Combined-Heat-Power>, 2021, [Accessed 12-01-2024].
- [18] P. Wolfrum, M. Kautz, and J. Schäfer, “Smart operation of chp units,” *IFAC Proceedings Volumes*, vol. 45, no. 21, pp. 61–66, 2012, 8th Power Plant and Power System Control Symposium. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1474667016319450>
-

-
- [19] J. Moore and B. Shabani, “A critical study of stationary energy storage polices in australia in an international context: The role of hydrogen and battery technologies,” *Energies*, vol. 9, p. 674, 08 2016.
- [20] K. Sayed and H. Gabbar, “Chapter 18 - scada and smart energy grid control automation,” in *Smart Energy Grid Engineering*, H. A. Gabbar, Ed. Academic Press, 2017, pp. 481–514. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128053430000188>
- [21] B. T. Rolf Grünbaum, Mojtaba Noroozian, “FACTS, poderosos sistemas para una transmisión flexible de la energía,” https://library.e.abb.com/public/89318066821a283ec1256fda003b4d40/5_1999s.pdf, 1999, [Accessed 04-01-2024].
- [22] M. Donsion, J. Guemes, and J. Rodriguez, “Power quality. benefits of utilizing facts devices in electrical power systems,” in *2007 7th International Symposium on Electromagnetic Compatibility and Electromagnetic Ecology*, 2007, pp. 26–29.
- [23] M. A. Masoum and E. F. Fuchs, “Chapter 9 - the roles of filters in power systems and unified power quality conditioners,” in *Power Quality in Power Systems and Electrical Machines (Second Edition)*, second edition ed., M. A. Masoum and E. F. Fuchs, Eds. Boston: Academic Press, 2015, pp. 779–886. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128007822000099>
- [24] E. Csanyi, “Using HVDC Technology For Transmitting Electricity,” <https://electrical-engineering-portal.com/using-hvdc-technology-for-transmitting-electricity/>, 2012, [Accessed 18-01-2024].
- [25] S. Cuthrell, “Offshore Wind Needs Major HVDC Transmission Expansion,” <https://eepower.com/news/offshore-wind-needs-major-hvdc-transmission-expansion/>, 2023, [Accessed 18-01-2024].
- [26] N. Mostafa, H. S. M. Ramadan, and O. Elfarouk, “Renewable energy management in smart grids by using big data analytics and machine learning,” *Machine Learning with Applications*, vol. 9, p. 100363, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666827022000597>
- [27] D. Wang, X. Guan, T. Liu, Y. Gu, Y. Sun, and Y. Liu, “A survey on bad data injection attack in smart grid,” 12 2013, pp. 1–6.
- [28] J. Jiménez, “Suplantación de ARP: qué es y cómo afecta a nuestra red,” <https://www.redeszone.net/tutoriales/redes-cable/ataques-arp-spoofing-evitar/>, 2023, [Accessed 15-01-2024].
-

-
- [29] “Descubrimiento seguro de vecino de IPv6,” <https://www.juniper.net/documentation/mx/es/software/junos/neighbor-discovery/topics/topic-map/ipv6-secure-neighbor>, 2021, [Accessed 15-01-2024].
- [30] G. E. N. Institute, “National Energy Grid Spain,” http://www.geni.org/globalenergy/library/national_energy_grid/spain/index.shtml, 2024, [Accessed 11-01-2024].
- [31] “¿Qué es la Internet de las cosas? Definición y explicación,” <https://www.kaspersky.es/resource-center/definitions/what-is-iot>, [Accessed 24-01-2024].
- [32] “¿Qué es IoT (Internet de las cosas)?” <https://aws.amazon.com/es/what-is/iot/>, [Accessed 24-01-2024].
- [33] S. Violino, S. Figorilli, C. Costa, and F. Pallottino, “Internet of beer: A review on smart technologies from mash to pint,” *Foods*, vol. 9, p. 950, 07 2020.
- [34] H. Shahinzadeh, J. Moradi, G. B. Gharehpetian, H. Nafisi, and M. Abedi, “Internet of energy (ioe) in smart power systems,” 02 2019, pp. 627–636.
- [35] E. Rojas, “DEN2NE - Distributed ENergy ENvironments and NEtworks,” <https://github.com/NETSERV-UAH/den2ne-Alg>, [Accessed 10-11-2023].
- [36] “Las 7 V del Big data: Características más importantes,” <https://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/>, 2019, [Accessed 12-01-2024].
- [37] Telefónica, “Las 5 V del Big Data. ¿Cómo pueden beneficiar a las empresas?” <https://www.telefonica.com/es/sala-comunicacion/blog/5-v-big-data-beneficios/>, 2020, [Accessed 13-01-2024].
- [38] B. Bhattarai, S. Paudyal, Y. Luo, M. Mohanpurkar, K. Cheung, R. Tonkoski, R. Hovsopian, K. Myers, R. Zhang, P. Zhao, M. Manic, S. Zhang, and X. Zhang, “Big data analytics in smart grids: State-of-the-art, challenges, opportunities, and future directions,” *IET Smart Grid*, vol. 2, 03 2019.
- [39] R. Hat, “Diferencias entre la nube y la virtualización,” <https://www.redhat.com/es/topics/cloud-computing/cloud-vs-virtualization>, 2023, [Accessed 13-01-2024].
- [40] E. Zegura, K. Calvert, and S. Bhattacharjee, “How to model an internetwork,” *Proceedings - IEEE INFOCOM*, vol. 2, 02 2002.
-

-
- [41] J. Figueras, “¿Conoces NILM? Aplica Inteligencia Artificial en la gestión energética,” <https://www.linkedin.com/pulse/conoces-nilm-aplica-inteligencia-artificial-en-la-gesti%C3%B3n-figueras/>, 2018, [Accessed 23-01-2024].
- [42] A. Monacchi, D. Egarter, W. Elmenreich, S. D’Alessandro, and A. M. Tonello, “Greend: An energy consumption dataset of households in italy and austria,” in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2014, pp. 511–516.
- [43] Y. Himeur, A. Alsalemi, F. Bensaali, and A. Amira, “Building power consumption datasets: Survey, taxonomy and future directions,” *Energy and Buildings*, vol. 227, p. 110404, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S037877882030815X>
- [44] S. Makonin, B. Ellert, I. V. Bajić, and F. Popowich, “Electricity, water, and natural gas consumption of a residential house in canada from 2012 to 2014,” *Scientific data*, vol. 3, no. 1, pp. 1–12, 2016.
- [45] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges, “Blued: A fully labeled public dataset for event-based non-intrusive load monitoring research,” *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, pp. 1–5, 01 2012.
- [46] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini, “The eco data set and the performance of non-intrusive load monitoring algorithms,” in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, ser. BuildSys ’14. New York, NY, USA: Association for Computing Machinery, 2014, p. 80–89. [Online]. Available: <https://doi.org/10.1145/2674061.2674064>
- [47] S. Makonin, “Hue: The hourly usage of energy dataset for buildings in british columbia,” *Data in Brief*, vol. 23, p. 103744, 03 2019.
- [48] N. Batra, M. Gulati, A. Singh, and M. Srivastava, “It’s different: Insights into home energy consumption in india,” 11 2013.
- [49] J. Kolter and M. Johnson, “Redd: A public data set for energy disaggregation research,” *Artif. Intell.*, vol. 25, 01 2011.
- [50] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, “Smart*: An open data set and tools for enabling research in sustainable homes,” *Proc. SustKDD.*, 01 2012.
-

- [51] L. Pereira, F. Quintal, R. Gonçalves, and N. J. Nunes, “Sustdata: A public dataset for ict4s electric energy research,” *Proceedings of the 2014 conference ICT for Sustainability*, 2014.
 - [52] J. Kelly and W. Knottenbelt, “The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes,” *Scientific Data*, vol. 2, no. 150007, 2015.
 - [53] D. P. B. Renaux, F. Pottker, H. C. Ancelmo, A. E. Lazzaretti, C. R. E. Lima, R. R. Linhares, E. Oroski, L. d. S. Nolasco, L. T. Lima, B. M. Mulinari, J. R. L. d. Silva, J. S. Omori, and R. B. d. Santos, “A dataset for non-intrusive load monitoring: Design and implementation,” *Energies*, vol. 13, no. 20, 2020. [Online]. Available: <https://www.mdpi.com/1996-1073/13/20/5371>
 - [54] K. Hühne, “Nilm datasets,” <https://github.com/k-nut/nilm-datasets/blob/master/nilm-datasets.csv>, [Accessed 23-01-2024].
-

Lista de Acrónimos y Abreviaturas

AMI	Advanced Metering Infrastructure.
API	<i>Application Programming Interface.</i>
ARP	<i>Address Resolution Protocol.</i>
BRITE	<i>Boston University Representative Internet Topology Generator.</i>
CHP	<i>Combined Heat and Power.</i>
DDoS	<i>Distributed Denial of Service.</i>
DEN2NE	<i>Distributed ENergy ENvironments and Networks.</i>
DSM	<i>Demand Side Management.</i>
DSO	<i>Distribution System Operator.</i>
FACTS	<i>Flexible AC Transmission System.</i>
FDI	<i>False Data Injection.</i>
HMAC	<i>Hash Message Authentication Code.</i>
HVAC	<i>High Voltage Alternating Current.</i>
HVDC	<i>High Voltage Direct Current.</i>
IA	Inteligencia Artificial.
IEEE	Institute of Electrical and Electronics Engineers.
IoE	Internet of Energy.
IoT	Internet of Things.
IPFC	<i>Interline Power Flow Controller.</i>
LPWAN	<i>Low Power Wide Area Networks.</i>
ML	<i>Machine Learning.</i>
NILM	<i>Non-Intrusive Load Monitoring.</i>
NLP	<i>Natural Language Processing.</i>
P2G	<i>Power-to-Gas.</i>
PAR	<i>Peak-to-Average Ratio.</i>
PMU	<i>Phasor Measurement Unit.</i>
PVT	<i>Photovoltaic Thermal.</i>
RSA	<i>Rivest, Shamir y Adleman.</i>

RTP	<i>Real Time Pricing.</i>
SCADA	<i>Supervisory Control and Data Acquisition.</i>
SEND	<i>Secure Neighbor Discovery.</i>
SG	<i>Smart Grid.</i>
SHA	<i>Secure Hash Algorithm.</i>
SINAIS	<i>Sustainable Interactions with Social Networks, context Awareness and Innovative Services.</i>
SSL	<i>Secure Sockets Layer.</i>
SSSC	<i>Static Synchronous Series Compensator.</i>
STATCOM	<i>Static Synchronous Compensator.</i>
SVC	<i>Static Var Compensator.</i>
TCSC	<i>Thyristor-Controlled Series Capacitor.</i>
TES	<i>Thermal Energy Storages.</i>
TFM	<i>Trabajo Fin de Máster.</i>
TLS	<i>Transport Layer Security Protocol.</i>
TOU	<i>Time-of-Use Pricing.</i>
UAH	<i>Universidad de Alcalá.</i>
UPFC	<i>Unified Power Flow Controller.</i>
VMs	<i>Máquinas Virtuales.</i>

A. Anexo I - Pliego de condiciones

B. Anexo II - Presupuesto

C. Anexo III - Manuales de usuario e Instalación

A modo de facilitar al lector el acceso y el uso de las herramientas empleadas en este TFM, se dedicará este anexo a la creación de un manual de usuario. Se desarrollarán de forma detallada los procedimientos de instalación y configuración del software empleado, además de exponer su funcionamiento completo. De la misma forma, se incluirá la información necesaria para comprender conceptualmente los scripts que se han empleado con el fin de automatizar y optimizar el desarrollo de los casos prácticos.

C.1. Herramienta BRITE

El objetivo de incluir esta sección es exponer al lector la información de instalación y configuración de la plataforma de generación de topologías de red BRITE, descrita previamente en la Sección 2.6.1.

C.1.1. Virtualización del sistema operativo Linux

Previamente a exponer los pasos de instalación del BRITE, es preciso indicar que se ha empleado una máquina virtual con un sistema operativo Linux para soportar las funcionalidades de la herramienta. En este caso, se ha utilizado como plataforma VMWare¹ y una versión del sistema operativo Ubuntu 20.04.1 LTS como se puede visualizar en la Figura C.1.

C.1.2. Compilación e instalación de BRITE

Para proceder a instalar la herramienta, se accede al repositorio del equipo de investigación NetIS de la UAH² y se descarga y se descomprime el fichero .zip con todos los contenidos necesarios. Después, se realiza un *make* desde el directorio raíz de la herramienta. Se pueden producir errores por la falta de paquetes *g++* o *openjdk++*, que se soluciona con los comandos dados en el Código C.1.

¹<https://www.vmware.com/es/products.html>

²<https://github.com/NETSERV-UAH/BRITE>

Device Name	ubuntu >
Memory	1.9 GiB
Processor	AMD® Ryzen 5 3400g with radeon vega graphics × 2
Graphics	SVGA3D; build: RELEASE; LLVM;
Disk Capacity	21.5 GB
OS Name	Ubuntu 20.04.1 LTS
OS Type	64-bit
GNOME Version	3.36.3
Windowing System	X11
Virtualization	VMware

Figura C.1: Características del sistema operativo

Código C.1: Instalación de paquetes

```
1 sudo apt install g++
2 sudo apt install openjdk-11-jdk
```

C.1.3. Ejecución de BRITE

Se puede ejecutar BRITE mediante la interfaz gráfica (GUI) (ver Código C.2) o por comandos. En el caso de los comandos se da la posibilidad de utilizar los ejecutables programados en C++ o en Java (ver Código C.3), obteniéndose los mismos resultados a partir de una misma entrada.

Código C.2: Ejecución de la interfaz gráfica de BRITE

```
1 $./britegui &
```

Código C.3: Ejecucion de BRITE por la línea de comandos

```
1 $bin/brite <archivo de entrada .conf> <archivo de salida .brite> <seed_file>
2 $C++/cppgen <archivo de entrada .conf> <archivo de salida .brite> <seed_file>
```

Como se especifica en el Código C.3, se requiere un archivo de configuración de entrada (.conf), en el cual se determinen los parámetros de la topología, y otro de salida, especificado en formato .brite. Como se ha expuesto en el Capítulo referente al Estado del Arte (ver Sección 2.6.1.3), para facilitar el uso de BRITE se aportan en el repositorio una serie de ficheros de Python y scripts dedicados a la automatización del proceso de creación de los ficheros de configuración y del tratamiento de la salida.

Universidad de Alcalá
Escuela Politécnica Superior



ESCUELA POLITECNICA
SUPERIOR



Universidad
de Alcalá