

Técnicas de tratamiento avanzado de datos y su aplicación en comunidades energéticas locales

Área de investigación: Predicción y tratamiento avanzado de datos

Línea de investigación: Tratamiento avanzado de datos para la gestión de las microrredes en tiempo real

Participantes: Instituto Tecnológico de Galicia (ITG) e Instituto Tecnológico de la Energía (ITE)

Resumen: En este documento se resumen los resultados obtenidos en la línea de investigación “Tratamiento avanzado de datos” de la red Cervera HySGrid+ en relación a las técnicas de tratamiento avanzado de datos y su aplicación en comunidades energéticas locales. A continuación, se introduce la importancia del tratamiento de datos y de la interoperabilidad en las Comunidades Energéticas Locales (CEL), se enumeran las principales técnicas de pre-procesado identificadas y se describe el diseño de la arquitectura software necesaria para garantizar la interoperabilidad con los sistemas de comunicación y poder aplicar las técnicas de tratamiento avanzado de datos identificadas.

Fuentes de datos en las CEL

Las “redes inteligentes” hacen uso de múltiples fuentes de información para su gestión y operación, cubriendo los datos del proceso de generación, transmisión, distribución y consumo de energía eléctrica.

La recolección y análisis de estos datos es la base para la programación de la generación y almacenamiento de la energía en la comunidad, así como de la gestión de la respuesta a la demanda de cara al éxito de la participación de las comunidades energéticas locales en los mercados emergentes.

La heterogeneidad de estándares de comunicación en toda la cadena de valor que rodea a las CEL para el uso de la energía desde el usuario hacia la red, y viceversa, supone un reto en sí mismo que debe ser superado para lograr la compatibilidad entre los sistemas que deban coordinarse para la prestación de servicio, bien sea interno o externo a la comunidad. Dependiendo del ámbito: residencial, terciario y/o industrial, la presencia de determinados protocolos será más extendida, siendo necesaria la compatibilización de éstos de cara a la interoperabilidad de recursos. Algunos de los protocolos que son de aplicación en el ámbito de las CEL se reflejan a continuación.

Protocolos de comunicación con contadores inteligentes
Protocolos de comunicaciones que permitan acceder a la lectura del consumo energético en diferentes contextos como Home to Smart Meter (IEC-870-5-102, ZigBee, SEP 2.0 ...), Smart Meter to Data Concentrator (OSGP, IEC-61850, DLMS-COSEM, METERS and MORE, PRIME, ...) o Data Concentrator to Data Collector (TCP-IP, IEC 60870-5 (CIM), FTP, IEC-61850 ...).
Modbus
Protocolo con aplicación especialmente extensa y estando presente en los mecanismos de transmisión de información en la red, en especial en smart meters y pasarelas que captan la información desde este estándar y la vuelcan en tiempo real hacia soluciones cloud o aplicaciones de servicios terceros.
OCPP
El protocolo OCPP (Open Charge Point Protocol) desarrollado por la OCA (Open Charge Alliance) tiene como objetivo determinar un procedimiento abierto y universal para que las estaciones de recarga se comuniquen con un sistema central, independientemente del fabricante, comercializadora o gestor del punto de recarga.
OPC-UA
OPC está considerado como el conjunto de protocolos de comunicación estándar entre dispositivos y sistemas industriales. La especificación más actual es OPC Unified Architecture (UA).
OpenADR
Estándar de comunicación abierto e interoperable que facilita el intercambio de información de la red inteligente entre operadores de sistemas, empresas eléctricas, agregadores, proveedores de servicios de energía y clientes finales
IEC 61850
Estándar ampliamente extendido entre los dispositivos de las subestaciones eléctricas y un potencial candidato para la comunicación entre dispositivos involucrados en la CEL y la interacción con la utilities para la prestación de servicios de a la red.
Protocolos IoT
MQTT, AMQP y XMPP, entre otros

Técnicas de pre-procesado

Las técnicas de pre-procesado de datos son necesarias para mejorar la calidad de los mismos, debido a fallos en las comunicaciones o en las lecturas de los sensores. Éstas incluyen las técnicas de integración, limpieza, transformación de los datos “brutos”, reducción de la dimensionalidad de los conjuntos de datos o estimación de los huecos de datos mediante algoritmos de regresión. A continuación, se enumeran algunas de las técnicas de pre-procesado más relevantes identificadas a contemplar en el tratamiento de datos en las CEL.

Reducción de la dimensionalidad

La reducción de la dimensionalidad es el proceso de reducir el número de variables o parámetros que se consideran en el análisis. Se presentan a continuación un conjunto de

técnicas de reducción de dimensionalidad aceptadas en un entorno Big Data: Missing Values Ratio, Low Variance Filter o Maximum relevance, High Correlation Filter o Minimum redundancy, Método basado en Rankings, Random Forests o Ensemble Trees y Principal Component Analysis (PCA), entre otros.

Detección de datos atípicos

Para la detección de datos atípicos pueden emplearse las siguientes técnicas de pre-procesado: redes neuronales, métodos basados en máquinas de vectores de soporte (SVM), teoría difusa y análisis clúster, métodos basados en GSA (Gap Statistic Algorithm) y métodos basados en k-NN (k-Nearest Neighbor), entre otros.

Métodos de llenado

Los métodos de llenado se basan en el uso de algoritmos para el cálculo del valor que falta, como métodos de regresión o valor medio (en períodos cortos de tiempo) o la creación de perfiles de demanda diaria a partir de datos históricos (largos períodos). Cabe destacar la interpolación lineal (IL), la estimación a partir de la media histórica (MH) y el promedio ponderado entre interpolación lineal y media histórica.

Métodos de transformación

En algunos métodos de predicción las series de tiempo no se procesan secuencialmente, por lo que es necesario codificar la posición de cada token. Esta codificación posicional se puede lograr representando la marca de tiempo como tres elementos: hora, día y mes. Para representar cada tipo de datos con veracidad, se descompone cada elemento en una componente de seno y de coseno.

Arquitectura software para el tratamiento y almacenamiento de datos

Una comunidad energética es un nuevo modo de generar, usar y gestionar la energía a nivel local, a través de la cooperación entre agentes -ciudadanía, administración local y pymes-, contribuyendo a crear un sistema energético descentralizado, justo, eficiente y colaborativo, al basarse en la participación de los propios usuarios.

Las comunidades energéticas necesitan disponer de sistemas de monitorización en tiempo real que permitan anticiparse tanto a las necesidades marcadas por el mercado eléctrico, como a las necesidades de la propia comunidad energética. Los diferentes puntos de medida definen un escenario multisensor que obliga a disponer de diferentes tipos de conexión, protocolos y seguridad para volcar la información al sistema de almacenamiento (Big Data).

Como resultado de las necesidades de comunicación y técnicas de pre-procesado identificadas, se ha diseñado una herramienta software que permita agregar información de diferentes fuentes, realizar una evaluación automática de la información que se introduce, así como acceder a las diferentes fuentes de datos de la comunidad de manera remota, para realizar un análisis sobre estos y mostrar de forma clara y precisa la información que permita gestionar la comunidad.

La arquitectura propuesta posee estas funcionalidades permitiendo el almacenamiento de información en diferentes bases de datos en función de las características del dato (Series

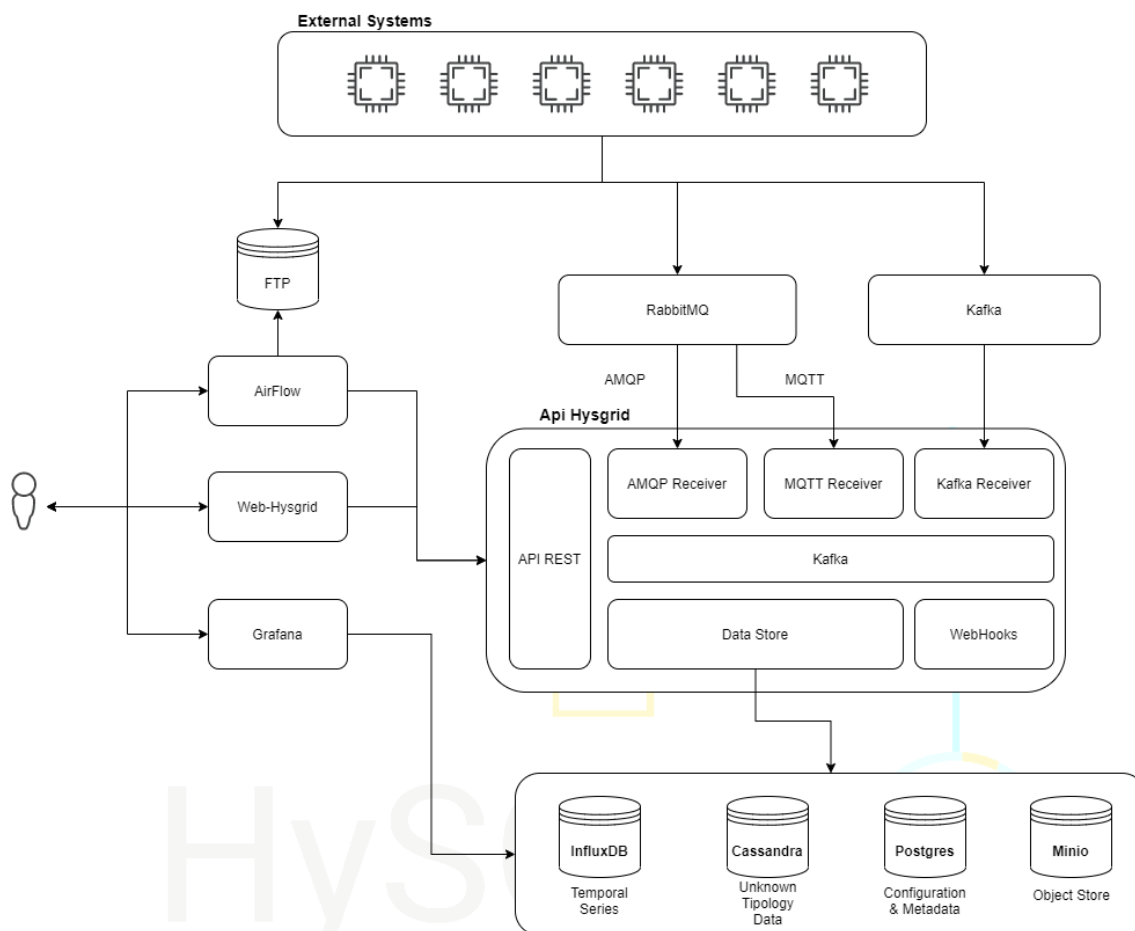


Figura 1. Arquitectura para tratamiento de datos

Por otro lado, es necesario integrar y orquestar la información de diferentes sensores y para ello se propone realizar dicha integración utilizando AirFlow como herramienta de ETL (del inglés Extract Transform Load) y RabbitMQ y Kafka para la comunicación con los sensores.