Modular Advanced Metering Infrastructure to Reduce Electricity Theft and a Cluster-Based Illegal Loads Detection

Roberto Morales-Caporal, *Senior Member, IEEE*

***Abstract*—This article introduces the development of a modular advanced metering infrastructure (AMI) for a single-phase power system, focusing on the reduction of non-technical losses (NTLs). In emerging economies, electromechanical meters are within the reach of consumers, and can be easily tampered with. With the proposed modular AMI, the physical security of smart meters is increased, thereby reducing this practice. The hardware design and the functionality of each main component of the modular AMI are explained. In addition, a cluster-based strategy to detect illegal electrical loads directly connected to power distribution lines is presented. The detection algorithm does not require extensive processing or complicated analysis of a large amount of data. The experimental and numerical results confirm the functionality of the developed AMI, an adequate detection of energy theft, and the feasibility to reduce the NTLs.**

***Index Terms*—AMI, clustering, LAN, NAN, mesh network, metering system, smart grid, smart meter, wireless communication.**

I. Introducción

A. Contexto

Hoy en día, en América Latina y el Caribe (ALC) sigue siendo común el robo de energía eléctrica, y la mayoría de estos ilícitos se cometen a través de: 1) la manipulación de los medidores, 2) la conexión ilegal de cargas eléctricas a las líneas de distribución [1], [2]. Y en el caso de existir una [*Advanced Metering Infrastructure*] (AMI), mediante el acceso remoto al medidor, para implantar malware en el firmware [3].

En ALC, operan AMIs en pocos lugares, por lo que el impacto negativo a la economía de las compañías de suministro de electricidad (en el caso de México, la Comisión Federal de Electricidad (CFE)), debido al robo por medio de ataques físicos aún supera con creces a las pérdidas por ciberataques.

Las mejoras continuas en las tecnologías de la información y la comunicación han permitido la implementación de AMIs para la medición y el control de la electricidad a través de medidores "inteligentes" [*Smart Meters*] (SMs), lo que ha llevado al establecimiento de *Smart Grids* (SGs) [4]. Estos avances tecnológicos han tenido lugar principalmente en países desarrollados, pero la conveniencia de utilizar e implementar estas tecnologías también es tema de investigación en países en vías de desarrollo, en donde por razones legislativas, de infraestructura o sociales, la tecnología debe de ser adaptada [5], [6]. No obstante, a la par, también han incrementado las estrategias de los usuarios maliciosos, que aprovechan la vulnerabilidad física y/o cibernética del sistema para utilizar el servicio a un costo reducido o sin pagar. Debido a esto, los investigadores y desarrolladores continuamente proponen novedosos métodos de detección de robo de electricidad [*Electricity Theft Detection*] (ETD). Los autores de [7] realizan un análisis exhaustivo de los diferentes tipos de robo, explican los métodos existentes para identificarlos, y hacen una comparación métrica entre ellos. Sin embargo, en el siguiente subtema se presenta una comparación propia de los métodos de ETD más actuales y relacionados al presente trabajo.

Roberto Morales-Caporal, is with Tecnológico Nacional de México/Instituto Tecnológico de Apizaco, Tlaxcala, México.
e-mail: roberto.mc@apizaco.tecnm.mx.

B. Trabajos Relacionados

Los estudios sobre ETD se pueden clasificar principalmente en dos tipos: los basados en hardware y los basados en datos.

1) Métodos Basados en Hardware

El autor en [8] presentó un método de ETD que consiste en instalar un equipo para medir la corriente eléctrica tanto en el transformador de distribución como en cada SM de usuario, y se evalúa el consumo total. Si existe una diferencia de corriente significativa, se confirma el robo de electricidad. En [9], se presentó un sistema basado en varios sensores y utiliza la combinación de datos cuando existen eventos anómalos y datos de consumo habituales, para modelar y detectar comportamientos relacionados con robos. Los resultados experimentales mostraron que con esta estrategia se puede identificar el robo de energía. Los autores de [10] usaron un procesador ARM-cortex M3 y otros componentes de hardware para detectar el robo de electricidad de cuatro formas diferentes entre las que están el cortocircuito del medidor y su manipulación para evitar la medición. Se construyó un prototipo para evaluar la capacidad de detección de los diferentes tipos de robo.

En [11], se propuso el uso de un medidor de potencia a la salida del transformador de distribución que alimenta un grupo de clientes. Una unidad central de supervisión recibe las mediciones de consumo de cada cliente y las del transformador y las compara. Si hay una diferencia significativa de potencia, se confirma el robo de energía. En [12], los autores presentaron el diseño e implementación de un sistema de gestión de energía basado en internet de las cosas (IoT). Se analizan los perfiles de carga para el desarrollo de numerosos enfoques de gestión de la energía. En [13], se investigó una encapsulación de claves por medio de hardware para SMs, este fue implementado en un FPGA. Los resultados experimentales muestran que

TABLA I
COMPARACIÓN CUALITATIVA DEL SISTEMA DESARROLLADO CONTRA OTROS SISTEMAS PREVIAMENTE PROPUESTOS.

Ref.	Seguridad física	Ciber-seguridad	Algoritmos de ETD	Volumen de datos	Iden. peque. robos	Iden. robo fuera de los SMs	Costo Compu-tacional	Descone-xión de usuario	Precisión
[8]	X	X	Suma algebraica	Bajo	✓	✓	Bajo	✓	—
[9]	X	✓	Aprendizaje Bayesiano	Medio	✓	✓	Medio	✓	—
[10]	X	X	Cálculo tradicional	Bajo	✓	✓	Bajo	X	—
[11]	X	X	Suma algebraica, Identificación	Bajo	✓	✓	Medio	✓	—
[12]	X	X	Matemáticas discretas	Alto	X	X	Medio	X	—
[13]	X	✓	CRYSTALS-Kyber	Bajo	X	X	Medio	X	—
[14]	X	✓	RNA profunda	Alto	X	X	Alto	X	0.97
[15]	X	✓	RNCs, MVS	Alto	X	X	Muy alto	X	0.96
[16]	X	✓	Data driven, RNAs	Alto	X	X	Muy alto	X	—
[17]	X	✓	Machine learning	Muy alto	✓	X	Muy alto	X	0.985
[18]	X	✓	RNA profunda, Redes Bayesianas	Muy alto	✓	X	Muy alto	X	0.97
[19]	X	✓	Machine learning, RNCs, RNAs	Muy alto	✓	X	Muy alto	X	0.975
[20]	X	✓	Suma acumulativa	Alto	✓	X	Alto	X	0.98
[21]	X	✓	Clustering, MVS	Alto	✓	X	Alto	✓	0.98
[22]	X	X	Clustering, DensityClust	Alto	✓	X	Medio	X	0.985
[24]	X	X	Flujo de datos de valor binario	Bajo	X	X	Bajo	X	F1-0.93
[25]	X	X	Clustering	Medio	X	X	Bajo	X	0.926
[26]	X	X	Arboles de decisión	Bajo	X	X	Medio	X	0.92
AMI-M	✓	✓	Suma algebraica, DensityClust	Bajo	✓	✓	Bajo	✓	0.96

el codiseño de hardware del esquema CRYSTALS-Kyber implementado en el FPGA reduce su tiempo de ejecución en comparación con su implementación de software.

2) Métodos Basados en Datos

Las estrategias basadas en el análisis de datos se pueden clasificar a grandes rasgos en: 1) algoritmos de aprendizaje supervisado y 2) algoritmos de aprendizaje no supervisado.

- *Algoritmos de aprendizaje supervisado*

Los autores de [14] propusieron utilizar dos redes neuronales artificiales (RNAs) para un clasificador de robo. La primera red analiza el consumo energético diario, y la segunda, integra datos no secuenciales, como la potencia contratada o la información geográfica. Con esta estrategia se logró obtener una precisión del clasificador (ACC) de 0.97. En [15], se propuso una combinación de redes neuronales convolucionales (RNCs) y máquinas de vectores de soporte (MVS). Las RNCs se utilizan para extraer características significativas y las MVS clasifican las características extraídas en robo y no robo. Se reportó una ACC de 0.96. En [16], se presentó un método de tres módulos. El primero maneja valores faltantes y datos de consumo no estandarizados. El segundo emplea un enfoque de equilibrio de clases híbrido para manejar el conjunto de datos desequilibrados y el tercero utiliza un clasificador basado en RNAs para confirmar los casos de robo. En [17], se propuso una combinación de algoritmos de aprendizaje automático (machine learning) (ML) para el manejo de valores atípicos y su normalización. Se emplearon tres métodos diferentes de ML, con lo que se obtuvieron resultados comparativos y una ACC máxima de 0.985. En [18] se propuso utilizar un clasificador basado en una RNA profunda. Se abordaron las debilidades de los conjuntos de datos y los problemas de desequilibrio a través de su interpolación, y se utilizó un optimizador bayesiano para determinar características "anormales", con lo que logró una ACC de 0.97. Los autores de [19] proponen un ETD basado en ML entrenado solo con datos benignos para detectar ataques ocultos (manipulación de consumo) mediante la combinación secuencial de un codificador automático, RNCs y RNAs.

Los métodos de ETD basados en inteligencia artificial sufren de tener una tasa de detección relativamente baja, pero una tasa de falso positivo (T_FP) relativamente alta [20]. Esto se debe a dos inconvenientes: 1) datos desequilibrados, ya que, la cantidad de datos de usuarios maliciosos es significativamente menor a la de los datos de usuarios honestos, y 2) factores no maliciosos, que pueden ser confundidos con robo, tales como: pequeñas fallas, eventos especiales, cambio de inquilinos, etc.

- *Algoritmos de aprendizaje no supervisado*

Los autores en [21] presentaron un ETD basado en patrones de consumo total de cada barrio, el cual se mide a la salida del transformador de distribución y se compara contra el consumo informado por los SMs. Si se detectan patrones de consumo "anormales", se etiquetan como sospechosos y se usa un clasificador MVS para confirmar el robo. En [22], se propuso una técnica de agrupamiento de datos. El grado de "anormalidad" del perfil de carga se computa de acuerdo con la distancia del centro de clúster y se utiliza la teoría denominada DensityClust [23]. Se simularon seis tipos de robo y se realizaron comparaciones con otros métodos similares. Los resultados mostraron que el algoritmo detecta diferentes tipos de robo con ACCs entre 0.927 y 0.985. En [24], se propuso un método que utiliza los datos "normales" de consumo para entrenar el modelo. Los experimentos se realizaron utilizando datos reales y simulados, incluido el preprocesamiento de vectores de consumo diario mediante normalización. Se reportó un valor-F1 de 0.93. En [25], se presentó un método basado en la agrupación y un factor de valores atípicos. Primero se analizaron los perfiles de carga con *k*-means. Luego, los perfiles que se encuentran lejos de los centros de clústeres son seleccionados como datos atípicos, y con el factor de valores atípicos se confirma el robo. Los autores en [26] proponen una metodología de detección de fraude utilizando técnicas de minería de datos y arboles

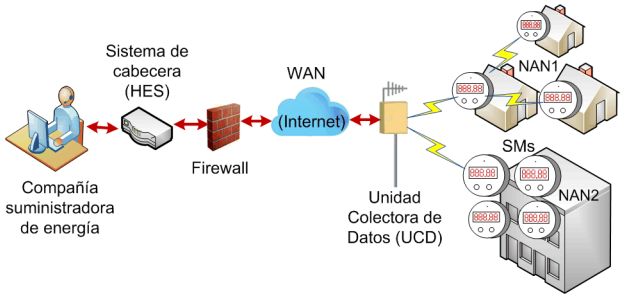


Fig. 1. Concepto simplificado de una AMI convencional.

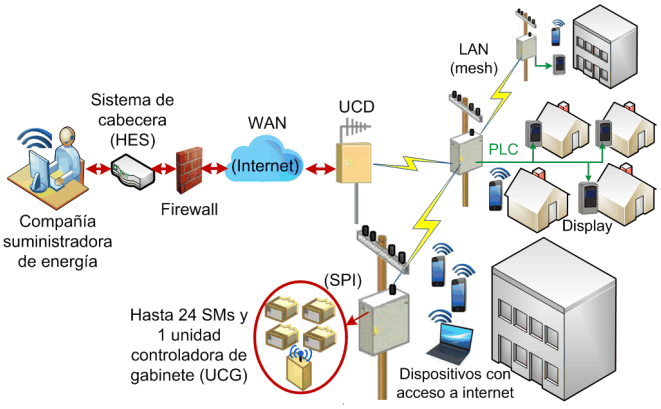


Fig. 2. concepto simplificado de la AMI-M desarrollada.

de decisión para identificar anomalías en los patrones de consumo. Los resultados mostraron una ACC de 0.92.

C. *Contribuciones*

De la tabla I, se observa que ninguna de las estrategias de ETD proponen una solución a los ataques físicos de los SMs, y únicamente las basadas en hardware tienen la capacidad de detectar el robo fuera de los SMs, que son las dos formas de robo más comunes en ALC; por lo que, la principal novedad de este sistema AMI Modular (AMI-M) es el de proteger de manipulaciones físicas a los SMs. Adicionalmente, se propone una estrategia de ETD de cargas ilegales conectadas a las líneas de distribución. Para esto, se utiliza una combinación de los métodos [11] y [22], pero modificados para poder ser utilizados en la AMI-M, ya que en [22], no se menciona que datos se computan (corriente, potencia, etc.), y se utiliza un elevado volumen de datos, debido a que no se normalizan.

Las contribuciones del sistema AMI-M en relación con una AMI convencional son: 1) evita la manipulación física de los SMs, 2) visualiza el consumo a usuarios con y sin acceso a Internet, 3) propuesta para identificar cargas eléctricas ilegales fuera de los SMs, 4) conexión/desconexión remota del servicio, y 5) incrementada ciberseguridad.

El resto del artículo se organiza como sigue: la sección II describe la operación y los principales componentes de una AMI convencional y de la AMI-M. En la sección III, se presenta el diseño, la implementación y la operación del hardware para la AMI-M, sus sistemas de comunicación y de ciberseguridad. En la sección IV, se explica la técnica propuesta de detección de cargas ilegales, la cual se basa en un algoritmo de agrupamiento. La sección V muestra los resultados del funcionamiento de la AMI-M y de la estrategia para ETD, y en la sección VI, se dan las conclusiones.

II. AMI

A. *AMI Convencional*

La Fig. 1 (no se muestran las conexiones eléctricas) muestra el concepto simplificado de una AMI convencional.

1) SM El medidor "inteligente" (SM) de electricidad es un dispositivo electrónico que registra el consumo del usuario cada determinado tiempo (15 min. para Norteamérica) y lo transmite a la empresa suministradora para generar facturas de consumo. Para esto, es común que varios SMs conformen una red de comunicación de área cercana (NAN) usando protocolos IEEE 802.15.4 (ZigBee) o IEEE 802.11 (Wi-Fi). Los SMs pueden realizar enrutamiento y determinar el camino de menor costo a través de la NAN hacia una Unidad Colectora de Datos (UCD), y el SM que se encuentra más cerca de una UCD hace la función de *router* para transferir la información [27].

2) UCD

La UCD sirve como un concentrador de datos y como una puerta de enlace que coordina la comunicación entre los dispositivos de la NAN con el sistema de cabecera, por medio de una red de área amplia (WAN) [28]. Dependiendo de la topología de la AMI, la UCD puede formar una red local (LAN) con otras UCDs para retransmitir la información.

3) Comunicaciones

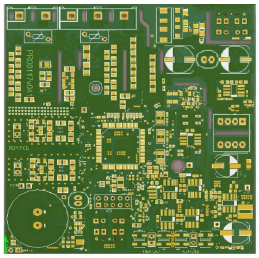
La comunicación de la AMI es bidireccional y puede ser implementada en diferentes topologías y con diferentes tecnologías [4], [28]. Hoy en día, las AMIs crean redes WAN, LAN y NAN entre los diferentes dispositivos de medición y control para lograr su objetivo. Para la red WAN, es preferible utilizar banda ancha como IEEE 802.11 (Internet) o IEEE 802.16 (WiMax), como sistema de comunicación. Sin embargo, las redes de comunicación NAN y LAN se implementan con sistema de comunicación inalámbricos de bajo costo y bajo consumo de energía, tales como: ZigBee, 6LoWPAN, Sigfox, LoraWAN, y Radio Frecuencia (RF). También es posible encontrar comunicación alámbrica como Ethernet o comunicación por línea de potencia (PLC).

4) Ciberseguridad

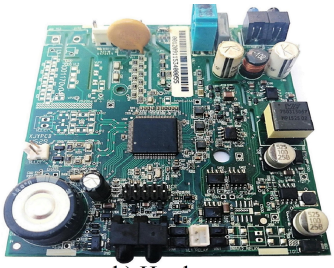
La ciberseguridad en la AMI implica principalmente que los datos desde y hacia los SMs no puedan ser manipulados, ni puedan enviarse comandos falsos de control a los relés de los SMs. Por lo tanto, la AMI requiere no solo de mecanismos de seguridad (como la autenticación), sino que, además, puede utilizar un sistema de detección de intrusos [3].

5) HES

El sistema de cabecera [*Headend System*] (HES) consiste en una instalación principal (hardware y software) que recibe y envía el flujo de datos de los SMs, para que el sistema de gestión de medición (SGM) los procese y analice [4], [28]. Si se usa un método de ETD, es en el SGM en donde corre el algoritmo. En el caso de detectarse el robo de energía, se



a) Footprint



b) Hardware

Fig. 3. SM.

Fig. 4. Versión comercial del SM [39].

factura de forma proporcional el costo del consumo "anormal" entre todos los usuarios conectados al gabinete [29].

B. AMI-M Desarrollada

La Fig. 2 muestra (no se muestran las conexiones eléctricas) un diagrama conceptual de la AMI-M desarrollada. Esta se diseñó teniendo en cuenta la norma oficial mexicana [30], la infraestructura existente y las especificaciones de la CFE [31].

1) SM desarrollado

El SM se diseñó para ser colocado de manera modular dentro de un gabinete hermético. El SM realiza la adquisición de 12 ciclos de señal subsecuentes de voltaje y corriente a una frecuencia de muestreo de 3.8 kHz, cada 15 minutos [30].

En este caso, el usuario tiene dos opciones para poder visualizar los datos de consumo: 1) desde un dispositivo con acceso a Internet. Ya que, con un usuario y contraseña, se accede a la plataforma de la empresa, donde también pueden realizar pagos, consultar historial de consumo, etc. 2) En el caso de no tener acceso a Internet (muy común en países en vías de desarrollo), el consumo se puede visualizar en un *display* que se desarrolló para este propósito. La conectividad entre el SM y el *display* se realiza mediante un protocolo PLC.

2) UCG

Los SMs dentro del gabinete se administran por un dispositivo electrónico desarrollado para este propósito, denominado Unidad Controladora del Gabinete (UCG), que realiza diferentes funciones, incluida la recopilación de datos de los SMs y su reenvío a la UCD. Además, la UCG tiene la capacidad de establecer una red malla (LAN) con otras UCGs para enviar la información a través de saltos hasta la UCDs más cercana.

3) UCD y HES

En este caso, la UCD y la HES realizan las mismas funciones que en la AMI convencional. Ya que, dentro de los requerimientos de la CFE, se consideró que existiera compatibilidad del sistema desarrollado con las UCDs y la HES ya existentes y en operación.

III. Diseño del Hardware

La metodología utilizada para desarrollar el hardware de la AMI-M fue el modelo "V", que consta de 7 etapas [32]. Además, fue diseñado bajo las Normas Internacionales de diseño electrónico IPC-2221, IPC-2231, IPC-2152 e IPC-D-325A [33], [34], y apegándose a los estándares [30] y [31].

A. Diseño e Implementación del SM

Entre las principales especificaciones y requisitos solicitados por la CFE para 1F son: 120 V, 30(100) A, 60 Hz, precisión 0.5 %, medición de kW, kWh y KVARh, puerto óptico infrarrojo, conexión/desconexión remota de servicio, y fuente propia de voltaje con respaldo de baterías [35].

El dispositivo principal que constituye el SM es la unidad de microcontrolador (MCU). En este caso, se seleccionó el ATSAM4C de Microchip [36]. Este dispositivo ofrece un nivel de integración y flexibilidad con dos procesadores de 32 bits y de 120 MHz cada uno. Cuentan con siete canales: ADC sigma-delta, y comunicación SPI [*Serial Peripheral Interface*].

La nota de aplicación [37] proporciona recomendaciones de diseño de hardware para SMs utilizando este MCU. La nota de aplicación [38] brinda consideraciones de diseño para el acondicionamiento de señal para la interfaz analógica ATSENSE-301 sigma-delta empotrada en la tarjeta SAM4CM.

La corriente de fase se mide a través de un transformador de corriente y el voltaje de fase se adquiere una vez que cruza un circuito resistivo. En [12] y [34] se pueden consultar las ecuaciones utilizadas para calcular los valores RMS de voltaje y corriente de fase, a partir de valores muestreados. Estos valores también se utilizan para determinar las potencias activa (P_{ACT_F}) y reactiva (P_{REA_F}) de fase como sigue:

$$P_{ACT_F} = K_{ACT}\frac{\sum_{k=1}^{SC} v[k] * i[k]}{SC} \tag{1}$$

$$P_{REA_F} = K_{REA}\frac{\sum_{k=1}^{SC} v_{90}[k] * i[k]}{SC} \tag{2}$$

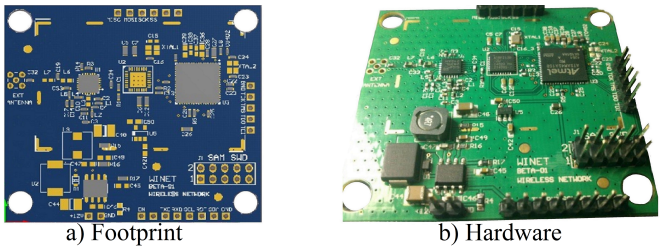
donde: $v[k]$ e $i[k]$ son las muestras del voltaje y corriente de fase en el instante de muestreo k, respectivamente. K_{ACT} y K_{REA} representan los factores de escala para la potencia activa y reactiva respectivamente. $v_{90}[k]$ representa la muestra de voltaje en el instante k desplazada 90° y SC representa el número de muestras en un lapso específico. Después, la potencia aparente (P_{APA_F}) de fase se calcula como:

$$P_{APA_F} = \sqrt{P_{ACT_F}^2 * P_{REA_F}^2} \tag{3}$$

el factor de potencia (FP) se puede evaluar después de calcular la potencia activa y aparente como $FP = \frac{P_{ACT_F}}{P_{APA_F}}$.

Una vez seleccionados todos los componentes y habiéndose diseñado los diagramas de circuito, se procedió al diseño de la PCB con ayuda de una herramienta CAD. La Fig. 3a muestra el *footprint* de la PCB diseñada, la Fig. 3b muestra el hardware durante su ensamble y la Fig. 4 muestra el SM desarrollado en su versión comercial [39].

Fig. 5. Protocolos de comunicación de la AMI-M.



a) Footprint b) Hardware

Fig. 6. UCG.

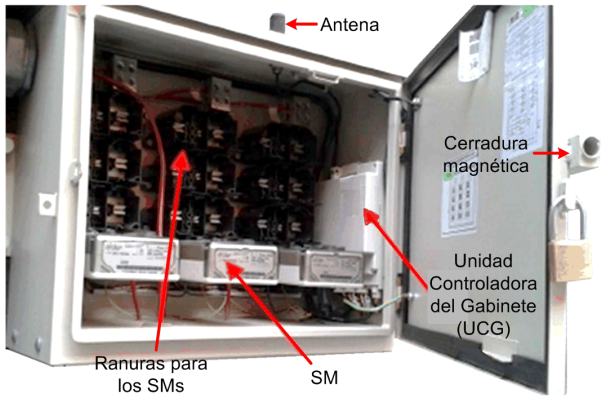


Fig. 7. Gabinete modular.

El SM opera en dos modos: configuración y medición. El modo configuración debe de: 1) generar un ID y establecer comunicación SPI con la UCG (ver Fig. 5). Se seleccionó este protocolo porque es más fácil de implementar en comparación con I^2C y UART. También configura la comunicación PLC con el *display* [39], 2) guarda la configuración en la EEPROM.

El modo de medición es responsable de: 1) calcular kW, kWh, KVARh, frecuencia, FP y el perfil de consumo por día, 2) guarda los datos en la memoria flash y los envía a la UCG, y 3) controla la bobina del relé. Cuando el SM no realiza acción alguna, este pasa al modo de "dormir" (bajo consumo).

B. Diseño e Implementación de la UCG.

Este dispositivo gestiona la información y el estado de cada uno de los SMs. También es responsable de monitorear las alarmas de seguridad del gabinete. La UCG tiene la capacidad de establecer una red *mesh* (malla) de comunicación inalámbrica (LAN) adaptiva con otras UCGs [40].

Una vez habiéndose establecido los requerimientos y especificaciones de la UCG, se seleccionaron los dispositivos electrónicos. Para la comunicación inalámbrica, se seleccionó el AT86RF212, ya que es un transceptor de bajo consumo diseñado para protocolo IEEE 802.15.4, (ZigBee/6LoWPAN) [41]. Se seleccionó el MCU ATSAM4S como cerebro de la UCG [42]. Este MCU funcionan a 120 MHz e integra acelerador de lectura flash, y memoria distribuida para comunicación de alta velocidad a través de la interfaz SPI. Una vez probado el prototipo, se diseñaron los circuitos y la PCB de propósito específico (ver Fig. 6a) y se fabricó (ver Fig. 6b).

La UCG también opera en dos modos: configuración y comunicación. En el modo configuración: 1) se genera un ID, 2) se configura la interfase SPI en modo maestro, ya que es este MCU proporcionará el reloj de los dispositivos esclavos, 3) establece una conexión RF (red mesh) con otras UCGs o con una UCD (ver Fig. 5). En este caso, el RF transmite en el espectro de 900 MHz, ya que es de uso libre. El modo comunicación: 1) transmite los datos de cada SM a la red mesh o a una UCD. 2) Retransmite los datos de las otras UCGs.

C. Gabinete

El gabinete se puede montar en un poste o pedestal, y no puede ser abierto por personas ajenas a la CFE. Es de acero y puede albergar 12 o 24 SM, según el modelo (ver Fig. 7) [43]. Cuenta con un sofisticado sistema de cerraduras mecánicas y magnéticas que actúa en combinación con un sistema de alarma de apertura de puerta, el cual comunica a la CFE el ID y ubicación del gabinete al momento de abrirse.

D. Red Mesh.

En este desarrollo se propuso utilizar el protocolo de comunicación inalámbrico 6LoWPAN (IEEE 802.15.4) para poder conformar una red mesh de varias UCGs y de este modo enviar los datos a una UCD. Se seleccionó este protocolo debido a su mejor despeño de transmisión respecto a Zigbee.

La tecnología de red mesh es especialmente adecuada para ser usada en AMIs debido a su capacidad para formar enlaces de comunicación ad-hoc entre nodos vecinos de la misma red. En este caso, se ha usado un esquema de enrutamiento AODV, puesto que este protocolo de enrutamiento supera a otras técnicas en cuanto a tasa de entrega de paquetes [44].

El nodo padre (UCD) de la red genera una IP para establecer una red inalámbrica en los canales disponibles. Los módulos RF de los nodos hijos (UCGs), buscan en los canales disponibles la IP de red, y envían una solicitud de acceso hasta que se responde la solicitud y se establece el vínculo.

La UCD utiliza un modem de capa 3 para conectar hasta 100 UCGs (2,400 SMs). Y utiliza Internet de fibra óptica (WAN) para establecer el flujo de información con la HES.

E. Ciberseguridad

La ciberseguridad de la comunicación SMs-UCG a través de la interfase SPI es de confianza inherente porque no se puede modificar, ya que se encuentra en la EEPROM. Por otro lado, en este desarrollo, la ciberseguridad de la comunicación inalámbrica se basa en la autenticación simultánea de iguales (SAE) [*Simultaneous Authentication of Equals*] [45]. SAE es un estándar de ciberseguridad muy apropiado para utilizarse en redes mesh. En este esquema, dos UCGs arbitrarias pueden iniciar el proceso de autenticación y no es necesario que sean vecinas directas. Por lo tanto, no es necesario tener jerarquías de claves y un mecanismo de distribución de claves. Cuando dos UCGs se descubren entre sí, intentan un intercambio SAE. Si SAE se completa con éxito, es porque ambas UCGs conocen la contraseña de la red.

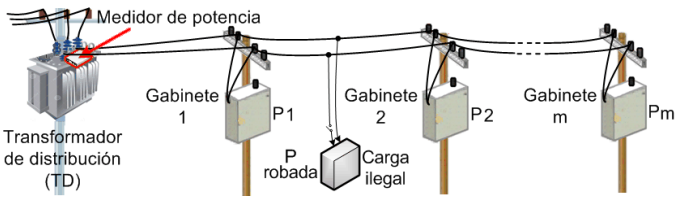


Fig. 8. Carga ilegal conectada a las líneas de distribución.

IV. Detección de Cargas Eléctricas Ilegales

Con la AMI-M, en caso de que se conecte una carga eléctrica antes de la acometida, pero después del gabinete, esta carga adicional es observada por el SM; por lo tanto, en estos casos ya no habrá pérdida. Pero, en el caso de que se conecte una carga directamente a la línea de distribución, como se ilustra en la Fig. 8, esta carga no puede ser observada por los SMs. Esta forma de robo es la más difícil de identificar y en la mayoría de los métodos de ETD no se resuelve.

En este caso, se propone utilizar una estrategia basada en hardware, similar a la propuesta en [11]. En [11], los SMs carecen de seguridad física que eviten su manipulación.

La estrategia consiste en adaptar un medidor de potencia (o corriente) a la salida del transformador de distribución (TD), como se muestra en la Fig. 8, luego se compara la suma de la demanda de potencia de los *m* gabinetes contra la suministrada por el TD. La diferencia de potencia, $\triangle P(t)$, entre la potencia suministrada por el TD $P_{TD}(t)$ en cualquier instante de tiempo de muestreo *t* y la suma de las potencias suministradas a los *m* gabinetes servidos por el TD, se evalúa como:

$$\triangle P(t) = P_{TD}(t) - \sum_{g=1}^{m} P_g(t) \tag{4}$$

donde $P_g(t)$ es la demanda de potencia en cada gabinete conectado al TD $(1 \leq g \leq m)$ en el instante t. Los registros de $\triangle P(t)$ se pueden recopilar en intervalos fijos (p. ej. 15 min.). Estas medidas se pueden expresar como un vector $\mathbf{x}= (x_1, ..., x_n)$, donde n es el número de medidas por día. Para minimizar las variaciones abruptas que pueden ocurrir por diferentes estilos de vida y pequeñas fallas, se puede procesar una normalización mín.-máx para **x**, como sigue [24]:

$$x_i \leftarrow \frac{x_i - \min\{x_i : 1 \leq i \leq n\}}{\max\{x_i : 1 \leq i \leq n\} - \min\{x_i : 1 \leq i \leq n\}} \tag{5}$$

En este estudio, se propone que el vector **x** sea sometido a un algoritmo de agrupamiento similar al propuesto en [23], para confirmar si se lleva a cabo un robo. Se seleccionó esta estrategia porque, en comparación con otros métodos de agrupamiento, DensityClust supera las desventajas de elegir el radio de vecindad y el umbral de densidad. También, debido a su elevada precisión y por su bajo costo computacional [46].

El algoritmo evalúa dos cantidades para cada punto de datos *i*: su densidad local ρ_i y su distancia δ_i desde puntos de mayor densidad. Ambas cantidades dependen únicamente de la distancia d_{ij} entre los puntos de datos.

La densidad local ρ_i del punto de datos *i* se expresa como:

$$\rho_i = \sum_j \chi(z); \quad z = (d_{ij} - d_c) \tag{6}$$

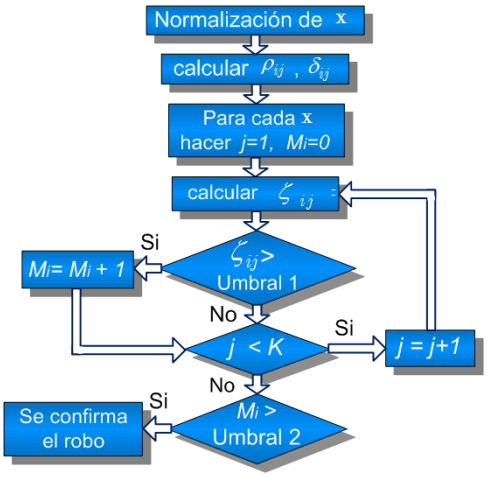


Fig. 9. Diagrama de flujo del algoritmo DensityClust.

donde $\chi(z) = 1$ si $z < 0$, y d_c es la distancia de corte.

Debido a que ρ_i es discreto en (6), se usa el kernel Gaussiano para estimarlo, como sigue:

$$\rho_i = \sum_{j \neq i} e^{-\left(\frac{d_{ij}}{d_c}\right)^2} \tag{7}$$

A medida que cambia la distancia de corte d_c, ρ_i definido en la ecuación (7) cambia más suavemente que en la ecuación (6). δ_i se mide calculando la distancia mínima entre el punto *i* y cualquier otro punto con mayor densidad, es decir:

$$\delta_i = \min_{j:\rho_j > \rho_i} d_{ij} \tag{8}$$

Y el punto con mayor densidad, δ_i se puede evaluar como:

$$\delta_i = \max_j d_{ij} \tag{9}$$

d_c se elige de tal manera que el promedio de ρ_i es aproximadamente del 1 al 2 % del número total de puntos. Los puntos de datos con ρ_i máximo global o local tienen un δ_i mucho mayor. Estos puntos se pueden elegir como centros de clústeres.

Cuando el algoritmo de clúster corre con el kernel Gaussiano y cada punto (ρ_i, δ_i) se gráfica en un sistema de coordenadas, se obtiene el gráfico de decisión. ζ_i representa el grado de anormalidad, y se evalúa como $\zeta_i = \frac{\delta_i}{\rho_i + 1}$.

El diagrama de flujo del algoritmo se muestra en la Fig. 9. Primero, se normaliza el vector **x** de cada día *j* (hasta cierto número *K* de días). Con los perfiles normalizados se calculan ρ_i, δ_i y ζ_{ij}, para identificar la magnitud de los datos "anormales". Después, se evalúa un umbral 2, que concierne con el período de tiempo a evaluar M_i, p. ej. identificar un número de datos "anormales" en el lapso de una semana o un mes; entonces, en caso de cumplirse la comparación de umbrales, los datos "anormales" serán confirmados como robo.

La tasa de verdadero positivo (T_VP) se define como la proporción de positivos correctamente identificados como positivos. La tasa de falso positivo (T_FP) es la proporción de resultados positivos que son FP y se evalúan como:

$$T\_VP = \frac{VP}{FP + FN}, \quad T\_FP = \frac{FP}{FP + VN} \tag{10}$$

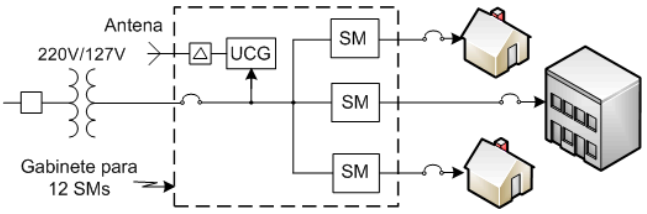
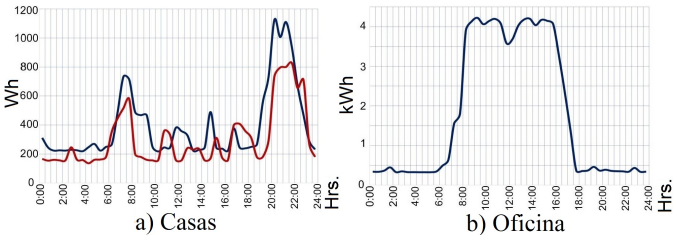


Fig. 10. Diagrama unifilar de la conexión experimental.



a) Casas b) Oficina

Fig. 11. Demanda de electricidad por día.

donde, FN, representa falso negativo, y VN representa verdadero negativo. La precisión del clasificador (ACC) se define como la proporción de resultados correctos y se evalúa como.

$$ACC = \frac{VP + VN}{VP + FN + FP + VN} \tag{11}$$

V. Resultados

A. SM

Para verificar el desempeño de los SM desarrollados, se instalaron tres de ellos dentro de un gabinete con capacidad para 12 SMs (ver Fig. 7). El diagrama de conexión de la instalación experimental se muestra en la Fig. 10. Previo a las pruebas experimentales, cada uno de los SMs fueron calibrados según la especificación de CFE [47].

La Fig. 11a muestra los datos medidos de la demanda de energía en dos casas diferentes durante un día de primavera, y la Fig. 11b muestra los datos obtenidos de la medición del consumo en un edificio de oficinas, en el mismo día. Estos resultados muestran el correcto funcionamiento de los SMs.

B. UCG

La Fig. 12 muestra el mapeo de las pruebas de conectividad de la red mesh. Esta prueba se realizó en Apizaco, México; se seleccionó el centro de la ciudad por la densidad de edificación (cada manzana mide aproximadamente 120 m x 120 m).

Los íconos conectados con líneas continuas representan UCGs con buena conectividad, la línea discontinua representa problemas de conectividad. El icono azul representa la UCD. Las líneas muestran cómo se dirige el tráfico de datos por la red. La red mesh calcula la trayectoria óptima a través de un mínimo número de saltos entre nodos con la mejor señal.

La tasa de pérdida de datos se probó a diferentes distancias en la ciudad. Los resultados se presentan en la Tabla II. Se observa que no hay pérdida de paquetes cuando la distancia entre dispositivos es menor a 500 m. Existen pequeñas pérdidas de paquetes entre los 600 m y los 800 m, y se presenta mayor pérdida después de los 800 m, esto se debe principalmente a que se utiliza un módulo RF de muy bajo consumo (1W).

Fig. 12. Prueba de conectividad con obstáculos.

TABLA II
TASA DE PÉRDIDA DE PAQUETES EN DIFERENTES DISTANCIAS.

Dispositivos	Distancia	Enviados	Recibidos
UCG-UCG	200 m	100	100
	400 m	100	100
	600 m	100	98
	800 m	100	97
	1000 m	100	93

C. ETD

En esta subsección, se muestran resultados numéricos con el algoritmo propuesto para ETD en líneas de distribución.

La Fig. 13a muestra una gráfica de datos simulados de consumo de potencia eléctrica, estos datos se han aproximado a un entorno real de operación durante tres meses con tres gabinetes. También se muestra la suma de las potencias de los tres gabinetes, así como la potencia suministrada por el TD. Se han considerado pérdidas técnicas del 3 % (ver Fig. 13 b).

Después de sesenta días, se simula un robo de 100 W entre el gabinete 1 y el gabinete 2 (ver Fig. 8). Se seleccionó esta

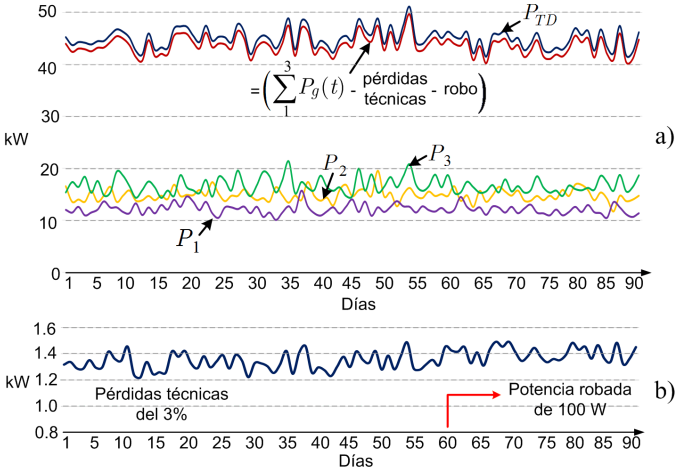


Fig. 13. Consumo durante tres meses. a) De arriba a abajo: potencia del TD, potencia de los tres gabinetes, menos las pérdidas técnicas, menos el robo, y potencia demandada de cada uno de los tres gabinetes. b) Pérdidas técnicas más robo.

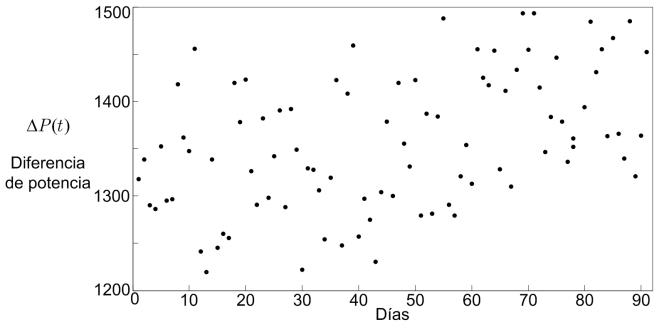


Fig. 14. Distribución de puntos de la $\triangle P(t)$ por día.

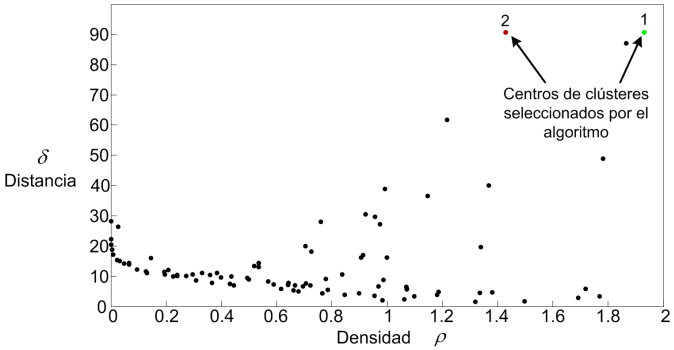


Fig. 15. Gráfico de decisión.

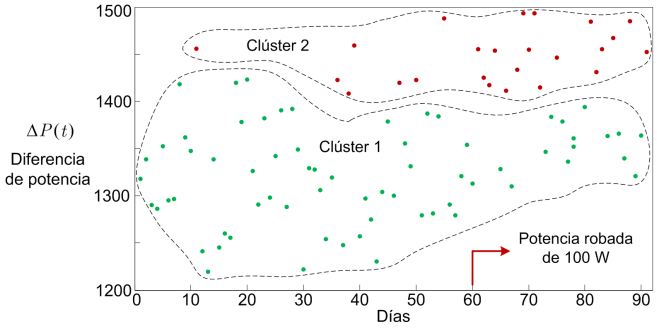


Fig. 16. Clústeres generados.

carga porque es la mínima conectada de manera ilegal.

La Fig. 14 muestra la distribución de datos $\triangle P(t)$ por día, representados en un espacio bidimensional. Con estos datos, se ejecutó el algoritmo de DensityClust con el kernel Gaussiano, y cada punto (ρ_i, δ_i) fue trazado en el sistema de coordenadas, para obtener el gráfico de decisión que se muestra en la Fig. 15. En este caso, el algoritmo se programó para seleccionar dos centros de clústeres, siendo estos los dos valores con mayor δ. Para esto, el algoritmo considera los δ_i que tienen un valor mayor que la distancia típica del vecino más cercano (9).

La Fig. 16, muestra los dos clústeres generados por el algoritmo. Dependiendo del número de días "anormales" (umbral 1, clúster 2) y dependiendo de los límites del umbral 2 (un mes), el algoritmo evalúa si existe robo o no, en este caso, puesto que se cumplen los dos umbrales en el último mes, el algoritmo confirma robo. De este modo, se demuestra que el algoritmo es capaz de realizar una clasificación aceptable de los consumos "normales" con respecto a los considerados "anormales" (robo), en este caso, con una ACC de 0.96.

Con este algoritmo, también es posible identificar, cuantificar y penalizar al mismo tiempo los casos de robo eventuales por semana o por mes, principalmente para picos de robo.

Evidentemente, si se incrementa la magnitud de la potencia robada, el algoritmo incrementará la precisión de detección. Además, incrementará la capacidad de identificar y cuantificar las pérdidas técnicas frente a las pérdidas no técnicas (NTLs).

En el caso de confirmarse el robo de energía después de algunos días (en este caso, 1 mes) en una determinada área alimentada por el TD, se activan las alarmas en el sistema de gestión de medición y se ejecutan las acciones determinadas para tales casos, que pueden ser, entre otros, vigilancia física del área atacada, o dependiendo de la cantidad de energía sustraída, se podrá realizar un cobro porcentual a los gabinetes bajo ataque según la resolución [29] para reducir las NTLs.

VI. Conclusiones

Se presentó el desarrollo y la implementación de un sistema AMI-Modular, diseñado para incrementar la seguridad física de los SMs, y consta principalmente de SMs de diseño especial, para que puedan ser conectados fácilmente (*plug and play*) por el personal de CFE dentro de un gabinete hermético. Hasta 24 SMs por gabinete son gestionados por un dispositivo electrónico también desarrollado para esta AMI, denominado UCG. La UCG tiene la capacidad de comunicarse de manera inalámbrica con otras UCGs con elevada ciberseguridad, para formar una red mesh y reenviar datos a la UCD más cercana, y esta a su vez, los envía a la HES de la CFE.

El protocolo de comunicación implementado es bidireccional y realiza la conexión/desconexión del relé del SM en un tiempo de 5 s, codifica y decodifica mensajes encriptados de acuerdo con el estándar de la CFE.

Además, se ha propuesto una estrategia de detección de cargas eléctricas ilegales conectadas a las líneas de distribución, el algoritmo propuesto puede operar con los datos generados por los SMs, y con los obtenidos desde un medidor de potencia a la salida del TD. El algoritmo es capaz de identificar el robo de energía desde bajos niveles, y a mayor energía robada, incrementa su precisión. Sin embargo, se debe considerar que en un entorno de operación real, los datos medidos pueden ser muy variables, pueden incluir ruido, y factores no maliciosos. No obstante, la magnitud de la potencia robada por lo común es mucho mayor a la aquí evaluada. La implementación del algoritmo propuesto en la HES de la CFE, será el siguiente tema de desarrollo.

Agradecimientos

El autor agradece a la empresa Tecnologías EOS, por el soporte técnico y administrativo, y al CONACyT (México) por el financiamiento. Proyecto PEI 00023455.

Referencias

[1] R. Jiménez, T. Serebrisky, and J. Mercado, "Power lost: sizing electricity losses in transmission and distribution systems in Latin America and the Caribbean," Inter-American Development Bank, 2014.

[2] P. Massaferro-Saquieres, "Detección de pérdidas no ténicas en redes eléctricas en un contexto de migración tecnológica y maximizando el retorno económico," PhD. Tesis, Universidad de la República, Montevideo, Uruguay, March 2022.

[3] Y. Guo, C. Ten, S. Hu and W. W. Weaver, "Preventive maintenance for advanced metering infrastructure against malware propagation," *IEEE Trans. Smart Grid,* vol. 7, no. 3, pp. 1314-1328, May 2016, DOI:10.1109/TSG.2015.2453342.

[4] X. Yu and Y. Xue, "Smart grids: A cyber–physical systems perspective," *In Proc. IEEE,* vol. 104, no. 5, pp. 1058-1070, May 2016, DOI:10.1109/JPROC.2015.2503119.

[5] B. Castro Inclán, "Control unit for systems used to supervise and monitor the electric power supply," Patent WO/2014/046531, March 2014.

[6] R. Binz, R. Bracho, A. Anderson, M. Coddington, E. Hale, M. Ingram, M. Martin, et al. "A report on the implementation of smart grids in Mexico". Golden, CO, USA. NREL/TP-7A40-72699, Jan. 2019.

[7] X. Xia, Y. Xiao, W. Liang and J. Cui, "Detection methods in smart meters for electricity thefts: A survey," *In Proc. IEEE,* vol. 110, no. 2, pp. 273-319, Feb. 2022, DOI: 10.1109/JPROC.2021.3139754.

[8] E. G. de Buda, "System for accurately detecting electricity theft," U.S. Patent 12/351,978, Jan. 14, 2010.

[9] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE J. Sel. Areas Commun.,* vol. 31, no. 7, pp. 1319-1330, July 2013, DOI: 10.1109/JSAC.2013.130714.

[10] K. Dineshkumar, et al., "Development of ARM processor based electricity theft control system using GSM network," *Int. Conf. Circuits, Power Comput. Tech. (ICCPCT),* Nagercoil, India, pp. 1-6, July 2015.

[11] A. Bin-Halabi, A. Nouh and M. Abouelela, "Remote detection and identification of illegal consumers in power grids," *IEEE Access,* vol. 7, pp. 71529-71540, May 2019, DOI:10.1109/ACCESS.2019.2920080.

[12] M. U. Saleem, et al., "Design, deployment and performance evaluation of an IoT based smart energy management system for demand side management in smart grid," *IEEE Access,* vol. 10, pp. 15261-15278, Feb. 2022, DOI: 10.1109/ACCESS.2022.3147484.

[13] S. Munawar et al., "Electricity theft detection in smart grids using a bybrid BiGRU–BiLSTM model with feature engineering-based preprocessing," *Sensors,* vol. 22, no. 20, p. 7818, Oct. 2022, DOI: 10.3390/s22207818.

[14] M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero and A. Gómez-Expósito, "Hybrid deep neural networks for detection of non-technical losses in electricity smart meters," *IEEE Trans. Power Syst.,* vol. 35, no. 2, pp. 1254-1263, March 2020, DOI: 10.1109/TPWRS.2019.2943115.

[15] E. U. Haq, J. Huang, H. Xu, K. Li and F. Ahmad, "A hybrid approach based on deep learning and support vector machine for the detection of electricity theft in power grids," *Energy Reports,* vol. 7, no. 6, pp. 349-356, Nov. 2021, DOI: 10.1016/j.egyr.2021.08.038.

[16] I. U. Khan, N. Javaid, C. J. Taylor and X. Ma, "Robust data driven analysis for electricity theft attack-resilient power grid,"*IEEE Trans. Power Syst.,* vol. 38, no. 1, pp. 537-548, Jan. 2023, DOI: 10.1109/TPWRS.2022.3162391.

[17] I. U. Khan, N. Javeid, C. J. Taylor, K. A. A. Gamage and X. Ma, "A stacked machine and deep learning-based approach for analysing electricity theft in smart grids," *IEEE Trans. Smart Grid,* vol. 13, no. 2, pp. 1633-1644, March 2022, DOI: 10.1109/TSG.2021.3134018.

[18] L. J. Lepolesa, S. Achari and L. Cheng, "Electricity theft detection in smart grids based on deep neural network," *IEEE Access,* vol. 10, pp. 39638-39655, April 2022, DOI: 10.1109/ACCESS.2022.3166146.

[19] A. Takiddin, M. Ismail and E. Serpedin, "Robust data-driven detection of electricity theft adversarial evasion attacks in smart grids," *IET Smart Grid,* vol. 14, no. 1, pp. 663-676, Jan. 2023, DOI: 10.1109/TSG.2022.3193989.

[20] X. Xia, J. Lin, Y. Xiao, J. Cui, Y. Peng and Y. Ma, "A control-chart-based detector for small-amount electricity theft (SET) attack in smart grids," *IEEE Internet Things J.,* vol. 9, no. 9, pp. 6745-6762, May 2022, DOI: 10.1109/JIOT.2021.3113348.

[21] P. Jokar, N. Arianpoo and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns," *IEEE Trans. Smart Grid,* vol. 7, no. 1, pp. 216-226, Jan. 2016, DOI: 10.1109/TSG.2015.2425222.

[22] K. Zheng, Y. Wang, Q. Chen and Y. Li, "Electricity theft detecting based on density-clustering method," Dec. 2017 *IEEE Innovative Smart Grid Techs-Asia* (ISGT-Asia), Auckland, NZ., 2017, pp. 1-6, DOI: 10.1109/ISGT-Asia.2017.8378347.

[23] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science,* vol. 344, no. 6191, pp. 1492-1496, Jun. 2014, DOI: 10.1126/science.1242072.

[24] C. H. Park and T. Kim, "Energy theft detection in advanced metering infrastructure based on anomaly pattern detection," *Energies,* vol. 13, no. 15, pp. 1-10, Jul. 2020, DOI: 10.3390/en13153832.

[25] Y. Peng, et al., "Electricity theft detection in AMI based on clustering and local outlier factor," *IEEE Access,* vol. 9, pp. 107250-107259, Aug. 2021, DOI: 10.1109/ACCESS.2021.3100980.

[26] S. Jain, et al., "Rule-based classification of energy theft and anomalies in consumers load demand profile," *IET Smart Grid,* vol. 2, no. 4, pp. 612-624, May 2022, DOI: 10.1049/iet-stg.2019.0081.

[27] N. Duan, C. Huang, C. -C. Sun and L. Min, "Smart meters enabling voltage monitoring and control: the last-mile voltage stability issue," *IEEE Trans. Industr. Inform.,* vol. 18, no. 1, pp. 677-687, Jan. 2022, DOI:10.1109/TII.2021.3062628.

[28] M. Faheem, et al., "Smart grid communication and information technologies in the perspective of industry 4.0: opportunities and challenges," *Computer Science Review,* vol. 30, pp. 1-30, Nov. 2018, DOI:10.1016/j.cosrev.2018.08.001.

[29] Energy Regulatory Commission. Resolution No. RES/999/2015.

[30] Mexican Official Standard NOM-001-CRE/SCFI-2019.

[31] Comisión Federal de Electricidad. Sistema de infraestructura avanzada de medición (AMI). Especificación CFE G0100-05, April 2015.

[32] R. Morales-Caporal, et al., "Development and implementation of a relay switch based on Wi-Fi technology," *17th Int. Conf. Elec. Eng., Comp. Sci. and Aut. Control,* (CCE), CDMX, México, Nov. 2020, DOI:10.1109/CCE50788.2020.9299142.

[33] IPC. Homepage. [Online] Available: https://www.ipc.org (2019/03/09).

[34] Texas Instruments:Implementation of a low-cost three-phase electronic watt-hour meter using the MSP430F67641. Application Report, SLAA621C, March 2014.

[35] A. Pérez-López, "Diseño de un medidor inteligente de electricidad," M.Sc. Tesis, TecNM-Ins. Tec. de Apizaco, Tlax., México, Aug. 2017.

[36] Microchip Technology, Inc. ATSAM4CMP32. Datasheet, Oct. 2016

[37] Microchip: ATSAM4C Series. Dual Arm® Cortex® M4 Core SOC with advanced security features for residential and C&I smart meters, 2021.

[38] Microchip: ATSENSE-301 Interfacing Guidelines, Jun. 2016.

[39] Tecnologías EOS. Presentacio_n_EOS_Eficiencia.pdf. [Online] Available: https://www.gob.mx/cms/uploads/attachment/file/344256/Presentacio_n_EOS_Eficiencia.pdf. (2022/11/07).

[40] J. E. Quiriz-López, "Diseño e implementación de un módulo de comunicación basado en RF y 6LoWPAN para el envío de datos a la CFE," M.Sc. Tesis, TecNM-Ins. Tec. de Apizaco, Tlax., México, Dec. 2017.

[41] Atmel Co., "Low power, 700/800/900MHz transceiver for ZigBee, IEEE 802.15.4, 6LoWPAN, and ISM applications," Rev.:42002E–MCU Wireless, Jun. 2017.

[42] Atmel Co.: SAM4S Series. Datasheet.pdf, Jun. 2017.

[43] Tecnologías EOS. Gabinete modular de medición. Ficha técnica. [Online] Available: https://tec-eos.com/v1/wp-content/uploads/2018/09/FICHA-TECNICA-GABINETE-PARA-MEDIDOR.pdf. (2022/11/07).

[44] N. Kushalnagar, G. Montenegro and C. Schumacher, "IPv6 over low-power wireless personal area networks (6LoWPANs): overview, assumptions, problem statement, and goals," *RFC 4919*, pp 1-12, 2007.

[45] D. Harkins, "Authentication of equals: a secure, password-based key exchange for mesh networks," *2nd Int. Conf. Sensor Techs. and Apps.,* Cap Esterel, France, 2008, DOI:10.1109/ SENSORCOMM.2008.131.

[46] QiQi Duan. Testing Framework, DensityClust, MATLAB Central File Exchange. Nov. 2015.

[47] Comisión Federal de Electricidad. Watthorímetros monofásicos y polifásicos electrónicos. Especificación CFE GWH00-78, Jan. 2006.

Roberto Morales Caporal (S'05–M'08–SM'14) received the B.Sc. degree in electromechanical engineering from the National Technological Institute of Mexico/Technological Institute of Apizaco (TecNM/ITA), Apizaco, México, in 1999, the M.Sc. degree in electrical engineering from the SEPI-Superior School of Mechanical and Electrical Engineering (ESIME), National Polytechnic Institute (IPN), Mexico City, Mexico, in 2001, and the Dr.-Ing. degree in electrical engineering from the University of Siegen, Siegen, Germany, in 2007. From 2001 to 2003, he was a Lecturer with the UPIITA, IPN. Since 2008, he is with TecNM/ITA. His areas of research interest include DSP-based digital control, control of power converters, hardware design and the IoT. He is a member of the National Research System (SNI) of Mexico.