Computer Networks 185 (2021) 107675

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Efficient and robust data availability solution for hybrid PLC/RF systems

Hassan N. Noura [a], Reem Melki [a,*], Ali Chehab [a], Javier Hernandez Fernandez [b,c]

[a] Department of Electrical and Computer Engineering, American University of Beirut, Beirut 1107 2020, Lebanon
[b] Iberdrola Innovation Middle East, Doha, Qatar
[c] Division of ICT, College of Science and Engineering, Hamad Bin Khalifa University, Qatar

ARTICLE INFO

Keywords:
PLC
PRIME
Radio frequency communications
Data availability
Availability attack
Information dispersal algorithm
RF
Data confidentiality

ABSTRACT

Recently, many efforts have been made by researchers, standardization organizations, telecommunication industries and energy providers, in order to realize robust and efficient communication in Smart Grids. To meet these requirements, a new generation of Narrow-Band Power Line Communication (NB-PLC) protocols have been proposed. The main advantage of PLC systems is that they require no additional costs in terms of deployment and wiring since the existing power cables can be used to transmit power and data, simultaneously. The PLC technology has many applications within Smart Grids such as advanced metering infrastructure, distributed automation, street light control, and public charging. One popular PLC standard is the Power-line Related Intelligent Metering Evolution (PRIME). This technology enables different types of Smart Grid services over electricity distribution networks. In PRIME, the security is implemented at the data link layer and it targets data confidentiality and message authentication, but not data availability. In this paper, we propose an efficient and robust security solution to guard against availability and confidentiality attacks. The main goal of the proposed solution is to enhance the security and reliability of PLC communication among end nodes with minimum overhead in terms of resources, complexity and latency. The solution is based on the Information Dispersal Algorithm (IDA) and the physical characteristics of PLC channels. In particular, end nodes benefit from a shared working key and the random channel parameters to derive the cryptographic primitives that are used in the proposed scheme. The security and performance tests showed that the proposed solution introduces a low overhead while offering a high degree of availability and reinforcing confidentiality.

1. Introduction

Power-Line Communication (PLC) is a well established technology that enables the simultaneous transmission of data and Alternating Current (AC) electric power through wired conductor links. It has attracted the attention of many researchers and telecommunication industries due to its high efficiency and low cost. Specifically, PLC requires no additional wiring, materials, installation or maintenance, since it exploits the existing electrical links to relay information. Like any communication system, the security of the transmitted data in PLC is of utmost importance and it is vulnerable to multiple active and passive attacks [1,2].

PRIME is an open and complete OFDM-based (Orthogonal Frequency Division Multiplexing) specification for NB-PLC tailored for Smart Grids [3]. This technology is already deployed in many countries in tens of millions of units [4,5]. The PRIME specification tackles the issue of security (data confidentiality and message authentication), however, it has no specific mechanism for data availability [6,7]. One of the main advantages of PRIME is that the utilized bandwidth is divided into eight independent channels (wired) that can be combined with wireless communication, creating hybrid PLC/RF networks [8–10]. Consequently, one could exploit these facts to enhance the reliability of existing links and to ensure the availability of data at all times. Typically, having one channel connecting the devices to the infrastructure makes it prone to different attacks such as Denial-of-Service (DoS) attacks, which renders data and devices inaccessible (if this single link fails, no communication is possible).

In this paper, we propose a new network model to guard against availability attacks at the end-node level. A hybrid PLC/RF system is considered consisting of multiple independent, wired and wireless links. On top of this network model, a novel cryptographic solution is proposed to achieve a high level of system reliability, in addition to reinforcing confidentiality with minimum latency and required resources. The proposed security solution is based on a joint encryption/encoding scheme, along with the characteristics of the physical channel. In

* Corresponding author.
E-mail addresses: hn49@aub.edu.lb (H.N. Noura), rmm71@aub.edu.lb (R. Melki), chehab@aub.edu.lb (A. Chehab), j.hernandezf@iberdrola.com (J.H. Fernandez).

https://doi.org/10.1016/j.comnet.2020.107675
Received 22 July 2020; Received in revised form 7 October 2020; Accepted 2 November 2020
Available online 7 November 2020

Table 1
Table of notations.

Symbol	Definition
M	The original message matrix
C	The coded message matrix
G	A set of N invertible coding matrices
z	Matrix coefficient
N	Number of generated coding matrices in G
L_P	Length of packet in bytes
SG	A selection table for G
π_{SG}	Update permutation table for SG
SL	A link selection vector
π_{SL}	Update permutation table for SL
n	Total number of possible channels
t	Threshold ($t \leq n$)
P_A	Number of affected packets for every T
T	Duration of one packet
Ch	Number of affected channels
$(\cdot)^{-1}$	Inverse operation

particular, the end nodes exploit the randomness and dynamicity of the physical layer to generate a dynamic key. Then, this key is used to derive multiple cryptographic primitives to be used to ensure confidentiality and availability, simultaneously. In this scheme, we modify the conventional IDA mechanism so that the coding matrices are only known to the communicating nodes. Accordingly, only a sub-set of the received data is needed to recover the original message. It should also be noted that the proposed solution accounts for impulsive noise that is dominant in PLC channels. Several security and performance tests are conducted to prove the robustness and efficiency of the proposed scheme.

The rest of the paper is structured as follows. Section 2 presents the necessary background information related to the PRIME specification and the IDA process. Section 3 reviews the existing availability attacks and the schemes presented in the literature. Section 4 describes the dynamic key derivation scheme and the generation process of the cryptographic primitives. Next, the proposed solution, which is based on IDA and the physical channel characteristics, is described in Section 5. Section 6 assesses the robustness of the proposed solution and its resistance against several security attacks. The performance of the proposed scheme is analyzed in Section 7. Finally, Section 8 summaries the contributions of this work and provides future prospects.

2. Background

In this section, we review the main concepts that the proposed solution relies on. Mainly, we describe the PRIME protocol and the IDA process. Note that Table 1 presents the notations used in this paper.

Fig. 1 illustrates the system model that is considered in this paper. We consider a hybrid PLC system in smart grids, where multiple wired and wireless connections are available. This type of systems is referred to as a hybrid PLC/RF system. More specifically, it is a general PLC system that is able to support a wireless connection such as the case of the LoRaWAN technology (which can have a Wi-Fi and LoRa connection) in addition to the existing wired PLC links. This system is general in the sense that the number of channels is flexible, that is we can only have wired connections and a variable number of channels (not necessarily 8 wired connections; we can dynamically choose a subset of them based on the conditions of the channels). The proposed solution is independent of the number and type of channels since users can benefit and exploit the physical characteristics of wired and wireless channels to generate a dynamic key. Although wired and wireless channels have different characteristics and unique features, users can extract random nonce values from each channel, combine them and perform the proposed scheme.

2.1. PRIME protocol

In [7], the detailed description of the PRIME specification is presented. For a better understanding of the proposed solution, we describe, next, the most important concepts of this protocol.

2.1.1. PRIME physical layer processes

The PRIME physical layer process is mainly divided into eight steps (Fig. 1):

- **Cyclic Redundancy Check (CRC)**: Upon the reception of the MAC (Media Access Control) data unit from the MAC layer, the CRC is appended to the PHY (physical-layer data unit) header. As for the PHY payload, the corresponding CRC is inserted at MAC layer.
- **Convolution coding**: In order to overcome channel errors, convolution coding with a rate of $1/2$ is applied. This step is not enabled by default and will only be activated in those cases with poor channel conditions.
- **Scrambling**: Unlike convolution coding, scrambling is mandatory to avoid having long sequences of 0's and 1's. Specifically, the header and the payload are both combined (mixed) with a pseudo-noise sequence that is generated using a linear feedback shift register. The feedback taps can be represented using a polynomial modulo 2 in finite field arithmetic.
- **Repetition**: When convolution coding is enabled, repetition (factor of four) is applied in order to ensure frequency and time diversity.
- **Interleaving**: Similar to repetition, this step is realized only when convolution coding is enabled. Here, an interleaver block is utilized to shuffle the bits in one symbol and consequently, randomize the occurrence of erroneous bits (avoid long bursts of bit errors).
- **Sub-carrier modulation**: This step is realized using one of three modulation techniques: DBPSK (Differential Binary Phase Shift Keying), DQPSK (Differential Quadrature Phase Shift Keying) or D8PSK (Differential Eight Phase Shift Keying).
- **Inverse Fast Fourier Transform (IFFT)**: IFFT is used to convert the modulated data from the frequency domain to the time domain.
- **Cyclic prefix**: Finally, a cyclic prefix is inserted to overcome inter-symbol interference. The cyclic prefix acts as a guard band between successive OFDM symbols. This step is crucial for OFDM systems.

Using uncoded D8PSK, PRIME is able to transfer per channel, at most 2268 bytes per packet and hence, it is able to achieve a rate equal to 128.6 kbps. On the other hand, using the robust mode (coded DBPSK), PRIME is able to attain 21.4 kbps with 377 bytes per packet [11].

2.1.2. Channel properties and types of noise

The PRIME protocol utilizes a frequency band ranging from 41.992 kHz to 471.6796875 kHz. It is divided into eight channels that can be used as a single channel or as multiple independent channels. A subset of these channels can also be used for the transmission and reception of data. Each channel employs OFDM modulation and it is composed of 97 equally spaced sub-carriers. Adjacent channels are separated by guard intervals of fifteen sub-carriers [7].

Similar to all band-limited channels, the PLC channel exhibits the following properties and specifications: RMS-DS (Root Mean Square Delay Spread) which represents the energy dispersion of the channel impulse response, the coherence bandwidth and the average channel gain (the average attenuation of the channel) [12]. Other parameters that are also taken into account when analyzing the PLC channel are the maximum achievable rate and the spacial correlation.

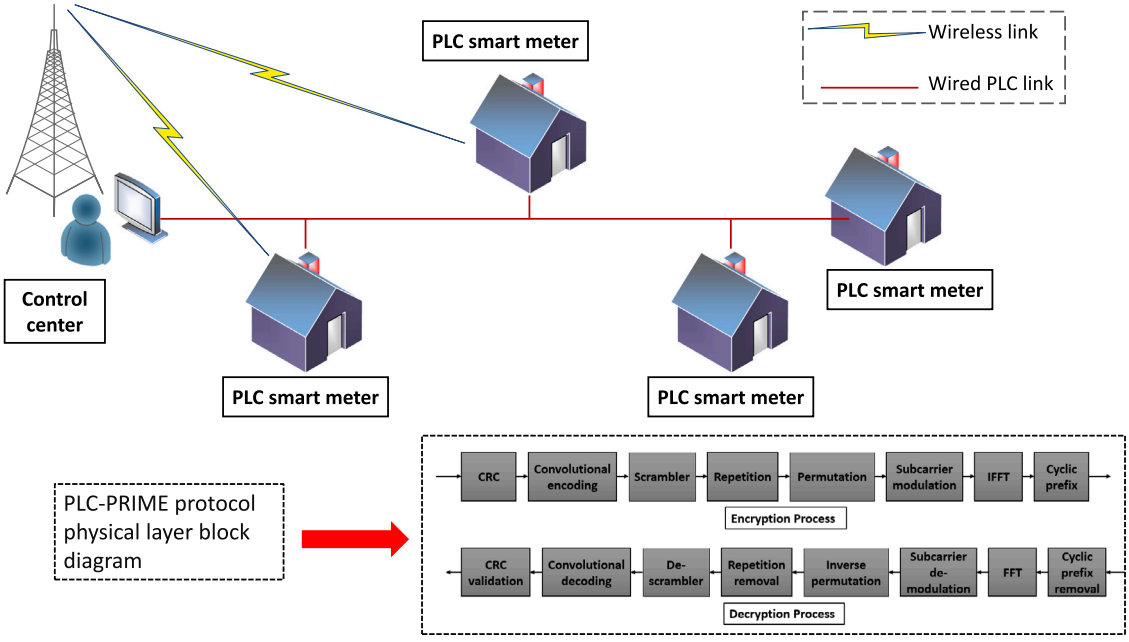


Fig. 1. An example of the considered system model: a smart grid utilizing wireless and wired PLC-PRIME connections.

In contrast to most communications channels, the PLC channel cannot be modeled in the context of Additive White Gaussian Noise (AWGN). Multipath effects which severely degrade the performance of a PLC channel are attributed to three main factors, namely: impedance mismatch, attenuation and noise. Noise is the main error source in PLC channels, especially narrow-band noise and the cyclic periodic impulsive noise synchronous to the mains [13]. Several research works in the literature have targeted the "cyclic periodic impulsive noise synchronous to the mains" due to its severity. In particular, the periodic impulsive noise affects/damages some packets, significantly, whereas other packets are received unaltered [14].

Generally, there are three types of noise in PLC [15]:

- Background noise: It originates from the environment.
- Narrow-band noise: It originates from radio amateur, wireless communication systems and other broadcasting systems.
- Impulsive noise: In PLC, this type of noise is predominant. The impulsive noise occurs periodically every half an AC (Alternating Current) cycle and it lasts between 10% and 30% of a period [14]. It is divided into three sub-types which are:
  - The cyclic periodic impulsive noise synchronous to the mains: It is composed of a train of impulses. The frequencies are a multiple of the frequency of the main electrical network. The most common sources of this type of impulsive noise include rectifiers, laptops, light dimmers, desktop computers and LCD monitors.
  - The cyclic periodic impulsive noise asynchronous to the mains: Unlike the synchronous impulsive noise, the asynchronous impulsive noise does not depend on the frequency of the electrical network. Typically, it is generated by switch mode power supplies that are connected to the network.
  - Aperiodic impulsive noise: This type is less frequent but more destructive than the previously discussed types of impulsive noise. It is mainly due to the On/Off and the plug/unplug activities in the power network.

It should be noted that the previous discussion applies to LV-PLC (Low Voltage) networks. A LV-PLC channel is not static and suffers from noise [16]. Differently, MV-PLC (Medium Voltage) is more static since it is based on three stable and non-varying parameters which are: line length, transmitter impedance and receiver termination [14]. The narrow-band PLC channel and the noise characteristics of an OFDM-based system (LV network) are the main focus of this work, hence, MV-PLC will not be discussed any further.

2.1.3. Security

Security in PRIME is realized at the MAC layer. In particular, the security mechanism, presented in the specification, achieves proper key generation and distribution, data confidentiality, authentication and integrity of packets and robustness against replay attacks.

Primarily, the base node and the service node have both a common shared key which is the Device Unique Key (DUK). The generation and distribution of this key is not indicated in the specification. It is only used for key derivation purposes using a key distribution function that is based on the AES-CMAC algorithm (Advanced Encryption Standard-Cipher-based Message Authentication Code). From DUK, two sub-keys are obtained: the Key Wrapping Key (KWK) and the REG Key (REGK). The former key is used to exchange keys between the base node and the service node, while the latter is used to encrypt some packets in the registration phase to ensure mutual authentication. More specifically, the main purpose of KWK is to encrypt and relay two keys that are used to encrypt/decrypt the transmitted data from the base node to the service node. These keys are the Working Key (WK), which is only known by the base node and the service node and the Sub-network Working Key (SWK), which is known by the entire sub-network.

The working key (and the SWK) is used to attain data confidentiality and message authentication (source authentication and message integrity) using the multi-round AES-128-CCM scheme (Counter with Cipher Block Chaining). The nonce that is used in AES-128-CCM along with WK, guard against replay attacks. Each service node should have a 4-byte counter that is incremented for every new message.

On the other hand, the defined security mechanism of the PLC-PRIME protocol does not account for link failure, data loss due to noise or data availability attacks. For this purpose, a data availability scheme is proposed in the following sections based on the IDA process.

Table 2
The comparison of two popular secret sharing variants.

	Communication & Storage overhead	Key management	Operation base	Security services	Additional operations
Shamir's secret sharing scheme	$(n-1)\times\|M\|$	Keyless	Polynomial	Secret sharing and data confidentiality	None
IDA	$(n-t)\times\frac{\|M\|}{t}$	Keyless	Matrix	Secret sharing, data confidentiality, source authentication and message integrity	None

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ G[t][0] & G[t][1] & \cdots & G[t][t-1] \\ \vdots & \vdots & \ddots & \vdots \\ G[n-1][0] & G[n-1][1] & \cdots & G[n-1][t-1] \end{bmatrix} \begin{bmatrix} M[0] \\ M[1] \\ \vdots \\ M[t-1] \end{bmatrix} = \begin{bmatrix} M[0] \\ M[1] \\ \vdots \\ M[t-1] \\ C[t] \\ \vdots \\ C[n-1] \end{bmatrix} \quad (1)$$

2.1.4. PLC applications

The applications of the PLC technology are mainly divided into two classes: applications of Broad-Band PLC (BB-PLC) and applications of Narrow-Band PLC (NB-PLC). The former class includes last-mile telecommunications, voice over IP and high-definition television [17, 18], whereas the latter is usually exploited for metering, lighting control, and management of energy, and grid [19–21]. Generally, the applications of PLC can be summarized into: home and industry automation, multi-media via PLC, smart grid applications (it is applied in the generation phase, transmission phase and distribution phase), in-vehicle PLC and backbone for RF (Radio Frequency), VLC (Visible Light Communication) and NFC (Near-Field Communications) [15].

Similar to PLC, the hybrid PLC/RF system has a major role in smart grids. Other applications include: industrial automation, Automatic Meter Reader (AMR), Automated Metering Infrastructure (AMI), indoor and vehicular applications and Machine-to-Machine (M2M) systems [22].

2.2. IDA process

The Information Dispersal Algorithm (IDA) is a coding technique that converts a digital source file into n small digital files (shadows), where only t out of n shadows are needed to reconstruct the initial source file, correctly [23]. This is similar to the concept of network coding, except that network coding requires re-coding at intermediate nodes, whereas IDA is only performed at the end nodes. The Plank and Resch systematic (t, n)-IDA, which is a variant of Rabin's IDA [24], consists of two functions: $Share$ and $Recover$. The $Share$ function takes as input a message M and outputs an n-vector. The $Recover$ function inputs the elements of the n-vector, however, it utilizes only t elements since these elements are sufficient to recover the original message [25].

To perform the coding, an $(n \times t;\ t \leq n)$ publicly known binary matrix, G, is required. The first t rows of G form a $(t \times t)$ identity matrix and the remaining $(n-t)$ rows consist of bits such that any t rows of n are linearly independent. The message is parsed into a t-vector message M. Afterwards, the message and the coding matrix are multiplied to generate an n-vector coded message C. Since the first t rows of G constitute an identity matrix, the first t rows of C will be identical to M. The matrix multiplication is shown in Eq. (1).

For the recovery of the original message, t rows are selected from the n-vector coded message C to obtain a new t-vector C'. Next, a $(t \times t)$ matrix G' is derived from the $(n \times t)$ matrix, G. Here, it should be noted that any $t \times t$ sub-matrix G' can be extracted from G since the rows of G are linearly independent, therefore, having an inverse matrix is always guaranteed. Finally, G' is inverted and multiplied with C' to recover M: $M = G'^{-1} \cdot C'$ [25].

Another secret sharing scheme is Shamir's technique [26]. The input to this scheme is a message of size M, which is divided into n encoded fragments such that t encoded fragments are needed for initial data recovery (at least t). The Shamir's secret sharing algorithm provides a high level of confidentiality, but it has quadratic complexity with respect to t. It also exhibits a high memory overhead since the size of each fragment is as large as the size of the initial data. Therefore, this method is usually applied to ensure the protection of small or important data like encryption keys. For larger data, Shamir's scheme introduces high overhead in terms of computations, communication and storage. According to Table 2, Shamir's secret sharing scheme suffers from high communication, computational and storage overhead. In contrast, IDA is a more efficient solution that significantly reduces this overhead. In addition, it offers more security services such as data confidentiality, source authentication and message integrity and data availability.

For the proposed solution, we apply the same IDA coding scheme, however, we design and employ different coding matrices that are only known to the communicating nodes.

3. Related work

In this section, we review the research works related to (1) network coding, (2) physical layer security to overcome the damaging effects of impulsive noise and to enhance the security level in PLC systems, and (3) the existing security attacks that target and affect the availability of such systems.

3.1. Network coding schemes in the literature

Due to the harsh environment of NB-PLC channels and the properties of the cyclostationary noise (periodic), the reliability and availability of transmitted data is compromised and highly affected in this type of wired channels. In particular, some packets are extremely damaged while others are received intact. To overcome this weakness, several schemes have been presented in the literature based on the concept of "coding"; the end nodes combine packets in such a way that only a sub-set of the transmitted data (based on a certain threshold) is adequate to recover the full message. The contributions of the work presented in [13] are divided into three parts: first, the authors characterize the transmission error correlation, then they propose a new OSI-Layer2 (data link layer) scheme based on network coding to attain reliable communication and finally, they validate the advantages of the proposed scheme using real implementation.

Similarly, the authors in [27] evaluate the performance of network coding at the data link layer in PLC systems. The presented work mainly focuses on the combination of network coding and the ARQ (Automatic Repeat Request) mechanism, and the advantages of such a technique in terms of data rate and latency. Authors in [28] also exploit network coding to enhance the resilience, efficiency, security and reliability of smart grid communication. In particular, intermediate nodes store and send linear combinations of received or overheard data packets, so that critical data is relayed reliably in case of any connection breakdown.

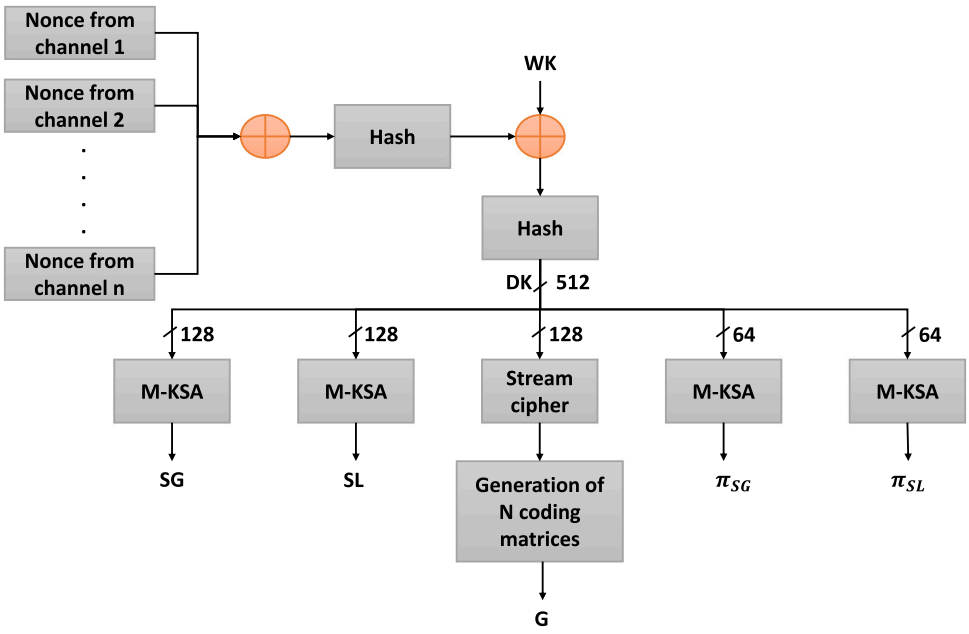


Fig. 2. The proposed dynamic key derivation function and the generation process of the corresponding primitives.

Authors in [29] employ a relay between the transmitter and receiver to overcome the effect of attenuation for long transmission distances between two nodes. This is mainly attributed to frequency selectivity in PLC channels. Moreover, network coding is utilized to improve data rate at the MAC layer.

On the other hand, authors in [30] present a slightly different approach, which is tunable sparse network coding. This method requires transmitting sparsely coded packets at the beginning of the transmission process but then, it converts to a denser coding structure towards the end. The proposed sparse structure is maintained in the recombination process at the intermediate nodes.

In contrast, authors in [14] propose an LT-based (Luby Transform) technique to mitigate the effect of impulsive noise (packet level). Specifically, a small packet sequence is first transmitted. This sequence is not encoded and it is used to analyze the characteristics of the periodic noise and distinguish the packets that are affected by noise and those which are not. Moreover, it allows the communicating nodes to estimate and approximate the burst time and its duration by identifying the affected and unaffected packets. At the receiver, a synchronized clock is utilized to mark the packets that are more likely to be hit by periodic impulsive noise. By utilizing LT codes, the receiving node is able to recover the original data without using the damaged packets.

3.2. Physical layer security for PLC/VLC/RF systems

A different goal is targeted in [31], where the authors design and experimentally test a communication stack for PLC using open standards. All of the protocols of the communication stack are addressed fully, from the physical to the network layer, taking into account the routing mechanism. The authors prove that it is possible to create a heterogeneous PLC/RF network in a low power and lossy network based on open standard protocols.

In [32], the authors aim at mitigating active attacks, namely the impersonation attack and jamming attack in Visible Light Communication (VLC) systems. For the former attack, authentication is ensured using binary hypothesis testing along with the channel gain, which acts as the fingerprint of the transmit device. As for the latter attack, the receiver treats the adversary's interference as Gaussian noise to recover the data sent by the transmitter. Authors in [33] also focus on VLC systems and they study the feasibility of using VLC as a network technology. The implementation challenges related to VLC-based networks are also discussed and highlighted. Moreover, a hybrid VLC/RF system utilizing energy harvesting in the presence of an eavesdropper, is assumed in [34]. In this work, the secrecy outage performance is studied taking into account light energy harvesting and the randomness of the locations of the legitimate receiver and the eavesdropper.

Currently, very few papers in the literature address the security of hybrid PLC/VLC/RF systems using physical layer characteristics. This is the main motivation behind the work in this paper.

3.3. Attack types

The sinkhole attack is an availability attack that consists of redirecting the traffic into a malicious node (gateway) to drop it or modify it, in case the encryption key was disclosed [35]. A similar attack, the down-link routing attack, requires two steps: eavesdropping on network traffic and replaying a message to another network through a compromised gateway. Another type of availability attacks is the replay attack which combines messages in order to illegally connect to the network. This can be achieved by intercepting and jamming the initial authentication session (join procedure). In this case, the sender will not be able to receive the reply (join-accept) message. Then, the malicious attacker can replay the join message after a counter overflow is achieved. Hence, flooding the End Device (ED) by the replay messages resulting in a DoS attack [36–39]. The beacon desynchronization attack is when an attacker compromises a gateway and starts sending fake beacons to the EDs. Thus, the gateway will open unconfirmed windows, which can lead to high interference and consequently causing DoS attacks [40,41]. ACK (acknowledgment) spoofing is an availability attack caused by the lack of confirmation of the communicated data in the acknowledgment message. In this scenario, the attacker acknowledges an up-link message by a previous ACK message [36,37]. Finally, an unpreventable attack consists of physically damaging the gateways or EDs and causing a DoS attack. In all of these attacks, the gateway is mainly compromised or damaged to disable the PLC-PRIME network availability. While some solutions were considered to defend against some of these attacks, such as adding a timestamp to prevent replay attacks, the probability of compromising the PLC-PRIME network availability remains very high. In this context, a new network model is required to reduce such risks and the related attacks. Therefore, in this paper we propose the following:

- A new network model with multiple channels to guarantee network availability (wired and wireless).
- Employing IDA to encode and distribute data while minimizing the delay and network overhead.
- A dynamic key-based derivation scheme of the encryption keys and the cryptographic primitives.

4. The proposed derivation scheme of the channel-based key and the dynamic cryptographic primitives

The first step in the proposed security solution is deriving a secret session key (referred to as the dynamic key) from the pre-shared working key (WK) and a channel-based nonce. The utilized nonce is obtained from the physical characteristics and properties of the available channels (wired and wireless). Using the dynamic key, the required cryptographic primitives are generated and are used to ensure data availability, in an efficient manner. The cryptographic primitives include a set of invertible coding matrices for the IDA process, a matrix selection table for each packet, a channel selection vector, and two update permutation tables for the selection table and the selection vector. Since the utilized primitives are generated from a key that is only known to the base node and service node, data confidentiality is also achieved along with data availability. The proposed scheme is robust, lightweight and efficient. It is based on IDA and it is applied on packets containing a specific number of bytes at the data link layer. The proposed scheme is generic since it can be applied when having wired links only, or when having both: wired and wireless links.

It should be noted that having multiple links is crucial for the proposed scheme. Specifically, we exploit the presence of multiple wired and wireless links to secure data and ensure its availability when one or several links/channels have been compromised (link failure, loss or availability attack). The proposed solution cannot be applied when having a single channel since it is based on a cryptographic approach (it employs encryption). The main motivation behind the proposed scheme is to be able to recover the original message correctly, using a subset of the received data. This is achievable using an efficient coding scheme and a number of parallel independent channels, both of which are considered in this paper. In a hybrid system, it is difficult for the attacker to attack all of the utilized links, especially when different technologies and channels are used (for example PLC and LoRa technologies).

4.1. Dynamic key derivation

Throughout the paper, we consider that there are n active channels (wireless and wired). In the absence of wireless channels (only wired), the maximum number of possible channels is 8, as indicated in the PRIME specification. From each channel (wired and wireless), communicating nodes extract a channel nonce based on the unique physical characteristics and properties of wireless and wired channels. Wired channels are assumed to be static (not varying) and ideal, however, several studies have shown that wired channels have a similar, but less aggressive behavior compared to wireless channels. This is attributed to the fact that wired channels experience destructive levels of noise, fading, frequency selectivity and multi-path propagation. Next, the obtained nonces are combined using the Exclusive OR operation (XOR) and hashed using the SHA-512 scheme to derive the channel-based parameter (N_0) that will be used to generate the Dynamic Key (DK). In particular, N_0 is combined (XORed) with the working key (WK), which is relayed by the base node to the service node using the wrapping key KWK. The KWK is derived from the pre-shared device unique key DUK. The resultant is again hashed to produce DK (512 bits). While performing the XOR operation if the two inputs do not have the same length in bits, zero padding is applied. For every new session, new channel parameters are extracted from the available channels and a new dynamic key is generated (new cryptographic primitives). In case of channel non-reciprocity, that is if the end points are not able to extract equal channel parameters, then one node extracts the needed nonce values and relays them to the other node using KWK.

Key generation in wireless is a well-studied topic and given the symmetry and randomness provided inherently by the channel there is a considerable number of techniques available [42]. On the other hand, the PLC channel is reciprocal but not strictly symmetric [12,43], which limits the options available for common channel properties that can be used in the dynamic key generation when compared with wireless networks. Symmetry can be exploited in the time domain using channel impulse response [44], while reciprocity techniques can be used on the frequency response and transfer function as proposed in [44,45].

4.2. Construction of dynamic cryptographic primitives

After generating the dynamic key, we derive the cryptographic primitives that are needed to achieve data availability along with data confidentiality. In particular, DK is divided into five sub-keys, each having a different functionality (Fig. 2).

The first sub-key is used to construct a set of N invertible matrices, $G = G_1, G_2, G_3, \ldots, G_N$, using stream ciphering. Each coding matrix in G has n rows and n columns ($n \times n$), where n is the number of available channels. More specifically, the first sub-key (128 bits) is processed and the output is reshaped into a matrix of N rows and n columns. Each row contains n unique values $(z_1, z_2, \ldots, z_n)$. This set of values should not include zeros or repeated values (distinct values), otherwise stream ciphering is re-iterated to produce new values. These conditions are necessary to preserve the invertibility property of the obtained matrices, and thus, to recover the original data. Afterwards, each filtered row is used to obtain an $(n \times n)$ IDA Vandermonde matrix. The construction of the first coding matrix in the set G is shown below:

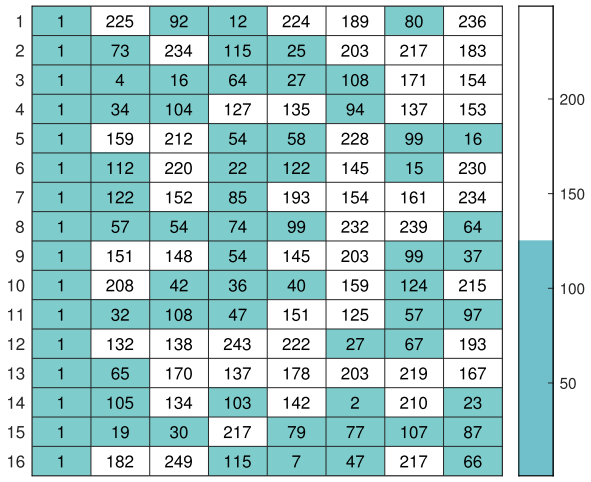
$$G_1 = \begin{bmatrix} 1 & z_1 & z_1^2 & \ldots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \ldots & z_2^{n-1} \\ 1 & z_3 & z_3^2 & \ldots & z_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_n & z_n^2 & \ldots & z_n^{n-1} \end{bmatrix} \quad (2)$$

where z_i is the matrix coefficient on the ith row ($i = 1, 2, \ldots, n$) and $z_i \neq z_j$ for $i = j = 1, 2, \ldots, n$. Any t rows and t columns of this IDA matrix form an invertible ($t \times t$) matrix, where t is a threshold value that is based on the quality of available channels ($t \leq n$). This method is repeated until all N coding matrices in the set G are generated. Fig. 3 represents one example of a Vandermonde matrix and its inverse.

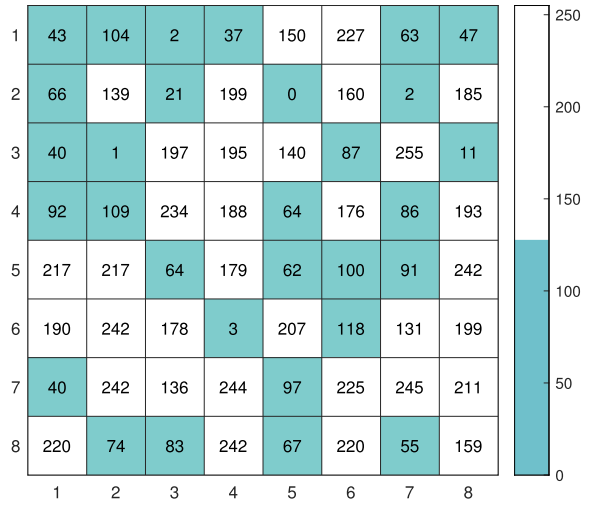
The second sub-key (128 bits) is used to generate a $(1 \times N)$ matrix selection vector SG (using the modified key scheduling algorithm M-KSA). This vector is used to choose one coding matrix out of N matrices ($G = G_1, G_2, \ldots, G_N$) for each new packet. It has random values between 1 and N. The third sub-key (128 bits) is used to derive a channel (link) selection vector SL ($n \times 1$) (M-KSA). It has random values between 1 and n and it is used to randomly shuffle the rows of each encrypted/encoded matrix before transmission. In other words, the packets that are transmitted on each channel are randomly permuted and then sent (order of packets is randomized). Finally, the fourth and fifth sub-keys (64 bits each) are used to obtain two update permutation tables π_{SG} and π_{SL}. The first permutation table π_{SG} has the same dimensions as SG and it is used to permute/update SG after N packets, whereas π_{SL} ($n \times 1$) is used to permute SL for every new packet.

5. Proposed security solution

The proposed model allows service nodes to connect and transmit data over multiple channels, simultaneously. In the PRIME specification the available bandwidth is divided into eight channels. These channels



(a) Vandermonde Matrix



(b) Inverse matrix

Fig. 3. A numeric example of a random Vandermonde matrix with $GF(2^8)$ and one of its corresponding inverse matrices.

can be used independently, or a subset of them can be utilized as a single transmission/reception band. For the proposed solution, we consider the case of 8 independent channels. Additionally, we assume that wireless transmission is also available. Therefore, a total of n independent channels (wired and wireless) are considered throughout the paper. This enhances the overall system performance in terms of increased end-to-end throughput, better resource utilization and smoother reaction to failures [46]. The proposed solution is applied at the packet level, where each packet contains L_P bytes. Data will be first encrypted and encoded at the data link layer, and then transmitted. In addition to data confidentiality and data availability, the proposed solution also accounts for impulsive noise and mitigates its effects on data. In the following, we describe the different steps of the proposed solution. Here, it should be noted that the security mechanism proposed in the PRIME specification can still be applied to increase the level of security even further.

5.1. Initialization

In order to overcome the effect of impulsive noise, we propose a simple channel estimation technique during the initialization phase (before communication). On each of the n channels, a small packet

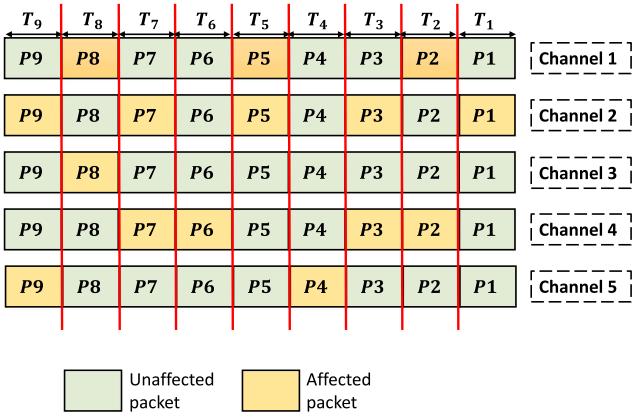


Fig. 4. The proposed estimation method assuming n = 5 channels and a 9 packet sequence over 9 time intervals.

sequence is transmitted as shown in Fig. 4 (duration of the packet is T). This sequence is publicly known by the transmitter and receiver and it is sent in plaintext, without encoding or encryption. Consequently, both of the transmitter and receiver will be able to identify the packets that are affected by impulsive noise and they will estimate the burst time occurrence and its duration. For example, at the first time interval, T_1, only one packet is affected among all channels (channel 2), whereas two packets are affected at T_2 on channels 1 and 4. On the other hand, the burst duration on channel 4 is $2{\times}T$ and it occurs for every 4 packets. This analysis is possible since the packet sequence is known, hence, the targeted packets can be easily distinguished. At the receiver side, a synchronized clock is employed. Using the simple estimation method, both nodes can approximate the number the packets that are expected to be hit by impulsive noise on each channel and at each time period T, and compensate for this loss using the proposed security scheme, which will be detailed next.

5.2. Data confidentiality: Joint encryption and encoding

Before applying the proposed scheme, communicating nodes can encrypt input data using the working key and the AES-128-CCM algorithm, as indicated in the PRIME specification. Then, input data (bytes) is grouped into multiple data matrices (bytes) and the proposed encryption/encoding scheme is performed at the data link layer. Note that the proposed scheme can also be used as an alternative to the multi-round AES algorithm (high complexity, latency and overhead) since it achieves data confidentiality and data availability at the data link layer, using one round of simple operations. It is simple, efficient and robust, and it is based on the dynamic channel properties of the physical layer. The proposed solution has a high randomness degree and security level, which makes it suitable for emerging communication systems and resource-limited devices.

5.2.1. IDA encoding and encryption

Following the channel estimation step, the number of affected packets, P_A, is stored at both of the communicating nodes for every time T. From Fig. 4, P_A is equal to 1, 2, 2, 1, 2, 1, 2, 2, 2 for T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8 and T_9, respectively (example). Moreover, based on the quality of the channels, an error threshold is assigned and it is equal to $\bar{t} = Ch + P_A$, where Ch is the maximum number of allowable link failures and $\bar{t}$ is the complement of t ($\bar{t} = n - t$, $t \leq n$). In other words, $Ch < n$ represents the threshold for recovering the original data, correctly. Beyond this value, data recovery is not possible. If P_A is ignored, then only data availability is attained with a minimum limit of $t = n - Ch$ channels. In case of any impulsive noise on any of the utilized $t = n - Ch$ channels, then data cannot be recovered. Since

periodic impulsive noise is the most dominant type of noise in PLC channels, we added another variable, P_A. Consequently, we will be able to recover the data in the presence of link failure and impulsive noise over $n-(P_A+Ch)$ channels using the proposed IDA scheme. It should be noted that: $0 \leq (\bar{t} = n-t = P_A+Ch) \leq n$. When $\bar{t} = n$, no transmission is carried out (all channels experience failure or noise). Differently, having $\bar{t} = 0$ means that none of the available channels is expected to experience any link failure or noise (perfect channel conditions).

For every time T, one input data matrix (containing a specific number of bytes) is processed, separately. The input data matrices have t rows and L_P columns. The value of t differs for every input matrix since P_A varies with time. In contrast, L_P is a fixed value for all input matrices, equal to the number of bytes in one packet. The value of Ch is also constant throughout the communication session.

The first data matrix M_1 (for T_1), having $t_1 = n - \bar{t}_1 = n - Ch - P_{A_1}$ rows and L_P columns ($t_1 \times L_P$ bytes/elements), is multiplied by the coding matrix $G_{SG(1)}$ ($n \times n$). This coding matrix is chosen based on the matrix selection table SG, which contains random values between 1 and N and it is used to randomly select a coding matrix for each input data matrix from the set $G = G_1, G_2, \ldots, G_N$. Since M_1 and $G_{SG(1)}$ have different dimensions, a sub-matrix, $\dot{G}_{SG(1)}$, is derived from $G_{SG(1)}$ and it has n rows and $t_1 < n$ columns $(1, z_i, z_i^2, z_i^3, \ldots, z_i^{t_1-1}$, where $i = 1, \ldots, n)$. Consequently, the resulting coded/encrypted matrix, C_1, will have n rows and L_P columns. Each row represents a packet having L_P bytes.

Next, the rows (packets) of C_1 are shuffled using the link selection vector SL ($n \times 1$) which contains random values between 1 and n.

At time T_1, each row/packet ($1 \times L_P$ bytes) of the permuted matrix C_1 is transmitted on one of the n available channels. The detailed procedure is illustrated in Fig. 5.

The same procedure is applied for all input matrices at Time $> T_1$. For every new input matrix, SL is updated using π_{SL}, whereas SG is updated after N input matrices (using π_{SG}), so that the order of the utilized coding matrices is changed randomly and frequently. In this way, adversaries will not be able to guess which coding matrix has been used at each time interval. Moreover, the coding/encryption matrices are derived from the dynamic key which is only known to the communicating nodes, hence, transmitted data will be secured from illegitimate nodes.

To ensure a maximum level of protection, the number of coding matrices in the set G should be equal to the number of packets to be transmitted. As a result, each input matrix will be encoded/encrypted using a unique encoding matrix, with additional overhead in terms of memory during the initialization phase.

5.2.2. Message authentication

After the encoding/encryption step, each packet is authenticated, as indicated in PLC-PRIME specification. Message authentication should be applied after encoding for the Network Server (NS) to check and validate the integrity and authenticity of the encoded/encrypted packets.

5.2.3. Decryption and decoding

Upon reception, the received packets are examined in order to validate their integrity and authenticate their source. The verification process is indicated in the PRIME specification. If the encoded/encrypted packets are authenticated, the proposed decryption/decoding process will be carried out, otherwise the packets will be discarded.

Next, each group of n packets will be decoded/decrypted, independently (one packet is received on each of the n channels at every interval T). At each time interval, T, the packets that are received on all channels are grouped to form a matrix, C, of size ($n \times L_P$). Next, the rows of each of these matrices are de-shuffled using the inverse selection vector SL^{-1} and the inverse permutation table π_{SL}^{-1}. Afterwards, the inverse IDA process is applied. The first group of encoded/encrypted packets which form the matrix, C_1 ($n \times L_P$), is received at T_{1+D}, where D refers to the transmission delay. From this matrix a ($t_1 \times L_P$) sub-matrix, $\dot{C}_1$, is derived. In other words, the t_1 rows/packets which correspond to the channels with the best conditions (t_1 channels out of n) are considered for the decryption/decoding process. This is possible since the proposed solution allows users to recover the original information using a subset of the received data (using t channels). Hence, one can overcome link failure, data availability attacks, and the effects of noise.

For the decryption/decoding process, an inverse coding matrix is required (Fig. 6). For the recovery of M_1, the inverse of $\dot{G}_{SG(1)}$ is first derived, and then multiplied by $\dot{C}_1$. However, $\dot{G}_{SG(1)}$ is not a square matrix (not invertible), thus, another sub-matrix, $\ddot{G}_{SG(1)}$, is derived from $\dot{G}_{SG(1)}$. The obtained matrix has t_1 rows and t_1 columns, and an inverse matrix, $\ddot{G}_{SG(1)}^{-1}$. The choice of $\ddot{G}_{SG(1)}$ is not random, where the t_1 rows corresponding to the t_1 best channels at T_{1+D} are selected from n (similar to the choice of $\dot{C}_1$). Here, it should be noted that any square sub-matrix of $G_{SG(1)}$ (or $\dot{G}_{SG(1)}$) is invertible. Finally, $\ddot{G}_{SG(1)}^{-1}$ ($t_1 \times t_1$) is multiplied with $\dot{C}_1$ ($t_1 \times L_P$) to recover M_1 ($t_1 \times L_P$). For subsequent packets, the same process is performed.

Using this technique, only a subset of the received packets is sufficient to correctly retrieve the original data.

5.3. Discussion

The proposed solution jointly achieves data confidentiality and data availability, using the proposed encoding/encryption process. In particular, end nodes exploit the random and dynamic nature of PLC channels, along with the working key (only known to the communicating nodes), to generate a channel-based dynamic key. From this key, multiple cryptographic primitives are derived: (1) a set of encoding/encryption matrices, (2) two selection tables and (3) two permutation tables to frequently update/change the selection tables. Since the utilized parameters (the working key and channel parameters) are only known to the communicating nodes, the obtained cryptographic primitives will also be only known to the communicating nodes. Therefore, robust data encryption can be successfully realized. This is mainly attributed to the fact that adversaries will not be able to acquire the used cryptographic primitives in the encoding/encryption process, thus, they will not be able to recover the transmitted data. In addition, we have proposed a simple updating mechanism based on the permutation scheme. By constantly and dynamically changing the utilized cryptographic primitives, illegitimate data retrieval becomes extremely hard. This increases significantly the security level.

In order to ensure data availability, we propose a simple and efficient encoding process. More specifically, each of the generated coding matrices is multiplied by a data sub-matrix, to obtain an encrypted/encoded matrix. Initially, all of the coding matrices have equal dimensions: n rows and n columns, where n is the number of active channels. However, for each new data matrix at each time interval T, only t rows are utilized out of n (dimensions become $n \times t$). The value of t mainly depends on two factors; Ch: maximum number of allowable link failures, and P_A: the estimated number of packets that are likely to be affected by impulsive noise during one time period T. Typically, t represents the minimum threshold for achieving reliable communication. Beyond this value, correct recovery of data is not possible, since the proposed solution (based on IDA) allows nodes to recover data using a subset of the received data based on a specific threshold (subset related to the threshold value t). When $t < n$, link reliability is achieved at the expense of a lower data rate, where $(n - t)$ channels can be used for data recovery in case of an attack, failure or poor channel conditions. In contrast, when $t = n$, only data confidentiality is realized. Therefore, the proposed scheme enhances not only the reliability and performance of communication channels, but also the security of transmitted data.

On the other hand, the values of t and n are flexible. The value of n depends on the number of available wireless and wired channels.

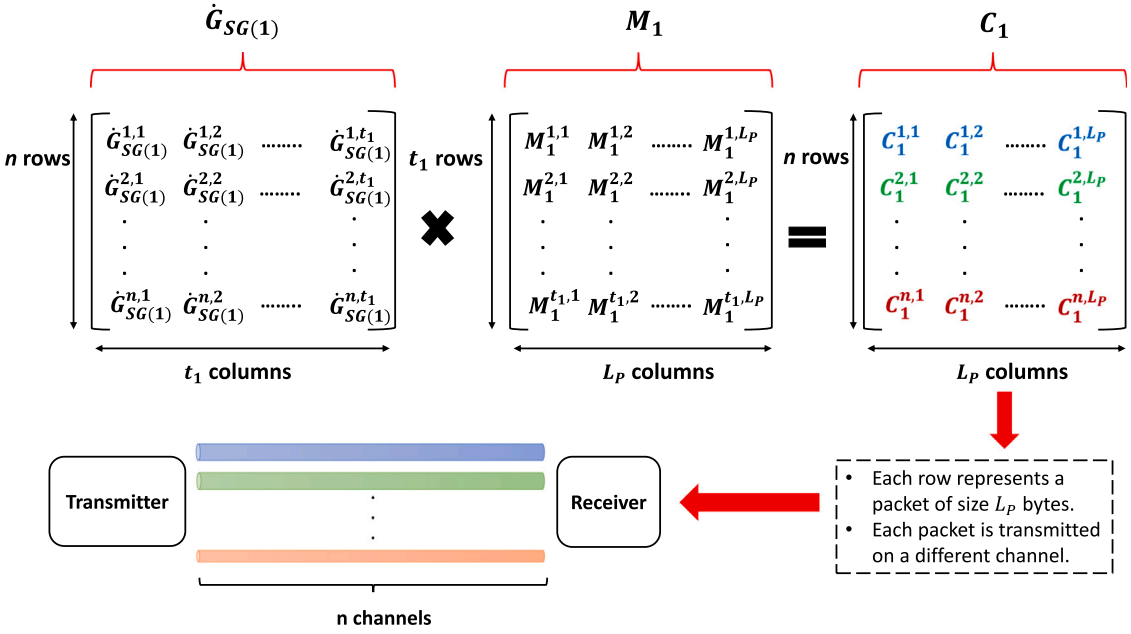


Fig. 5. PRIME encryption and encoding scheme using the proposed IDA process (at T_1).

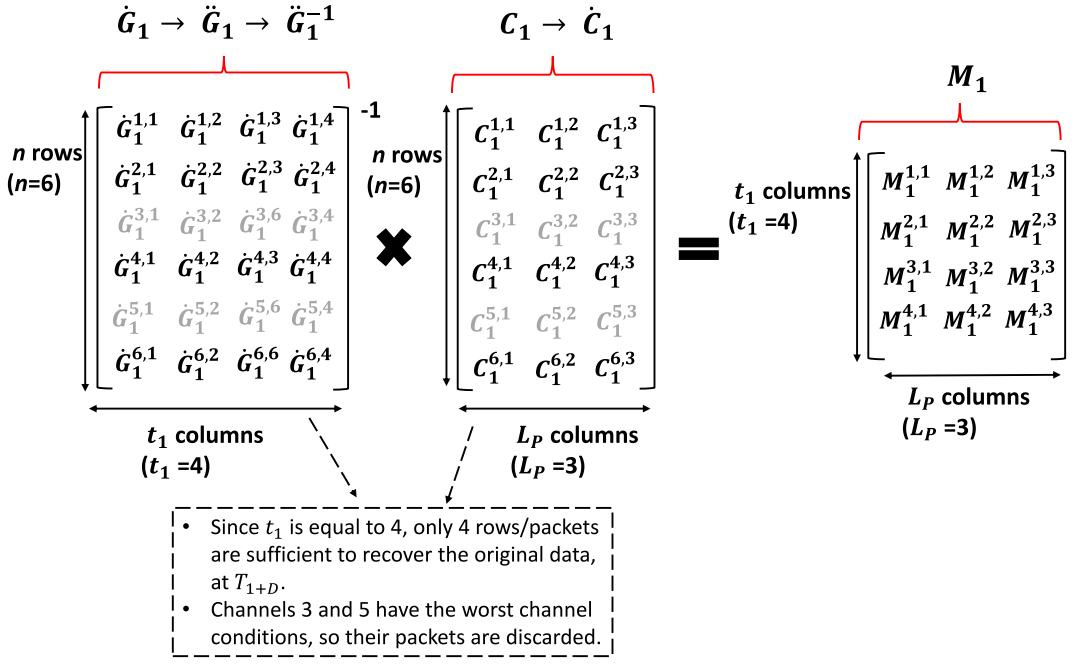


Fig. 6. An example of the PRIME decryption and decoding scheme with $n = 6$, $t_1 = 4$ and $L_P = 3$, at T_{1+D}.

Whereas, the value of t depends on the applications and users. Generally, high data-rate applications require a value of t equals to or close to n at the expense of low link reliability, however, low data rate and error-sensitive applications require a larger $(n - t)$ value.

The importance of the proposed solution is that it ensures link reliability and data privacy in an efficient and simple manner (low computational complexity). The reduction in the overhead is attributed to the fact that only t rows of C are used in the decryption process instead of using all n rows.

6. Security analysis

In this section, the security of the proposed solution is analyzed and validated. The proposed scheme mainly depends on a dynamic structure and a joint encryption/encoding mechanism, which in turn provides a high level of randomness, uniformity, and sensitivity of the encrypted data. The strength of the proposed cryptographic solution against several types of attacks is evaluated next.

In the proposed scheme, n data rows are spread over n different channels, where only t rows are required to retrieve the original message. Thus, the protection of the original message relies on the difficulty of collecting the t dispersed data rows from the n available channels. Specifically, the adversary should know the order of the encoded/encrypted data before collecting them from the corresponding channels. Moreover, the proposed solution is designed in such a way that the data-channel allocation is randomized based on the generated cryptographic primitives (dynamic and frequently updated). Therefore, the attacker must first guess the dynamic key, then predict the right

order of the received rows in each data matrix, and finally choose one encoding matrix from the pool of secret matrices, in order to correctly recover the original message.

Theoretically, well-known attacks such as statistical, differential, chosen/known plaintext/ciphertext, and brute-force attacks are prevented due to the dynamicity of the proposed approach [47]. However, we perform several security tests to validate this claim (multiple security metrics are considered). In this context, the proposed solution is presumed to be public, and it is assumed that the cryptanalyst has complete knowledge of the operations used, but none regarding the working secret key and the channel nonce.

6.1. Statistical attacks

In order to prove a scheme's robustness against statistical attacks, the following metrics should be examined: uniformity, high level of recurrence and maximum uncertainty (entropy). Consequently, we have plotted the Probability Density Function (PDF), recurrence and entropy of data that is encrypted/encoded using the proposed scheme.

6.1.1. Uniformity

The main requirement of any security scheme is having highly randomized output (ciphertext). Randomness can be assessed using multiple tests. One straight forward approach is plotting the distribution of encrypted data and evaluating it, visually. For this test, normally distributed data is generated and is fed to the proposed scheme. Fig. 7b shows the PDF of the resulting encrypted/encoded data. The presented results prove the uniformity of ciphertext, which confirms the desired randomness level.

6.1.2. Recurrence

Randomness can also be tested using the recurrence plot. Recurrence measures the relation of data at different time intervals (similarity between delayed versions). Typically, this plot should be extremely scattered for a robust security scheme. From Fig. 7a, it is evident that the recurrence graph is randomized since the plotted points span the entire available space. Instead of having all points within one area (normal distribution), the output is uniformly distributed over the whole graph. This validates that the proposed scheme achieves the desired obscurity level.

6.1.3. Entropy

Entropy calculates the level of data uncertainty. When entropy is equal to 1 (for data in bits), a maximum level of uncertainty is achieved, hence, the obtained data provides new information. In contrast, when entropy is 0, this means that no new information is inferred, therefore, the obtained data is predictable (not random). Fig. 7c shows the PDF of the entropy for 1000 iterations. The obtained graph proves that most of the entropy values range between 7.95 and 7.96, which is close to the desired uncertainty value at the byte level, 8.

6.2. Differential attacks

To overcome differential attacks, the avalanche effect should be satisfied. The avalanche effect depends on two important properties, confusion and diffusion. These properties mandate that a single bit change in the key or input (plaintext), should alter at least half of the bits at the output. For this purpose, we have conducted the difference and key sensitivity tests.

6.2.1. Independence

The independence property is validated using the difference test, in which a single change in the input should modify 50% of the bits at the output (resulting ciphertext). Fig. 8a displays the PDF of the difference values between the original and encrypted/encoded packets for 1000 iterations. It is evident that the desired independence property is achieved since the majority of the resulting values are close to 50%. This is also confirmed using Table 3.

Table 3
The percentage difference among encoded/encrypted packets for a random session key with $t = 4$ and $n = 8$.

	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6	EE_7	EE_8
E_1	48.80	48.63	49.88	49.56	49.88	50.10	50.20	50.63
E_2	50.49	49.93	49.76	50.42	49.56	49.49	49.98	50.56
E_3	49.54	49.95	50.07	49.37	49.78	50.49	50.24	50.05
E_4	49.02	49.19	51.03	50.42	50.49	49.10	48.27	49.88

Table 4
Key sensitivity test between two encoded/encrypted packets obtained for the same data but with two slightly different dynamic keys (DK and DK') with $t = 4$ and $n = 8$.

	EE_1	EE_2	EE_3	EE_4	EE_5	EE_6	EE_7	EE_8
EE'_1	50.56	49.98	50.88	48.56	50.34	50.20	50.29	48.85
EE'_2	49.90	51.66	50.37	50.15	50.02	50.81	50.42	50.78
EE'_3	49.54	51.	49.80	49.73	49.37	50.29	49.12	50.37
EE'_4	48.78	49.76	50.27	49.90	49.15	50.85	50.46	50.34
EE'_5	50.22	50.12	49.37	49.83	49.56	49.90	48.97	50.42
EE'_6	49.56	49.02	50.17	49.80	49.29	50.66	49.34	50.10
EE'_7	51.51	48.58	50.51	50.59	50.71	51.44	51.10	50.05
EE'_8	49.51	49.17	49.78	49.95	48.51	50.56	50.32	50.05

6.2.2. Key sensitivity

Fig. 8b and Table 4 show that a slight change in the dynamic key leads to a 50% difference in ciphertext, thus, attaining the key sensitivity property. Also, we have measured the sensitivities of the utilized cryptographic primitives upon changing a single bit in the dynamic key with $n = 8$ and $t = 4$, and for 1000 iterations. Fig. 9a quantifies the difference between two sets of coding matrices at the bit level. This value is very close to the ideal value, which is 50%. Therefore, the obtained results indicate that a different set of coding matrices is obtained for slightly modified keys. Figs. 9b and 9c measure the difference between matrix selection tables and the link selection tables at the packet level for a slight modification in the key, respectively. The results also indicate that completely different selection tables are produced (difference greater than 90%). Consequently, any key modification will lead to different cryptographic primitives and consequently, a better key avalanche effect is achieved.

6.3. Chosen/known plaintext/ciphertext attacks

Traditional cryptanalysis techniques, which encompass chosen/known plaintext/ciphertext attacks, rely on the relation between several plaintext/ciphertext pairs to recover the utilized encryption key and thus, obtain the original data. Due to the dynamicity of the proposed solution, this type of attacks is not possible. In particular, for every packet or set of packets, the utilized cryptographic primitives are updated/changed, in a dynamic and random manner. The same message would lead to two different ciphertexts at different time intervals using the proposed scheme since different selection tables and different coding matrices are utilized for every packet. As a result, any relation between different plaintext/ciphertext pairs will no longer exist.

6.4. Brute force attacks

In the proposed scheme, communicating nodes estimate the common shared channels and then combine the pre-shared working key with a channel-based nonce. After XORing the two parameters, the result is hashed using the SHA-512 algorithm to obtain a 512-bit dynamic key. The size of this key is greater than 128 and hence, it is sufficient to resist brute force attacks.

Next, we assess the security of the proposed solution against the attacks listed in Section 3.3. To perform the sinkhole attack, illegitimate users should have the utilized encryption key. Since the dynamic key is based on two factors, the secret working key and the common

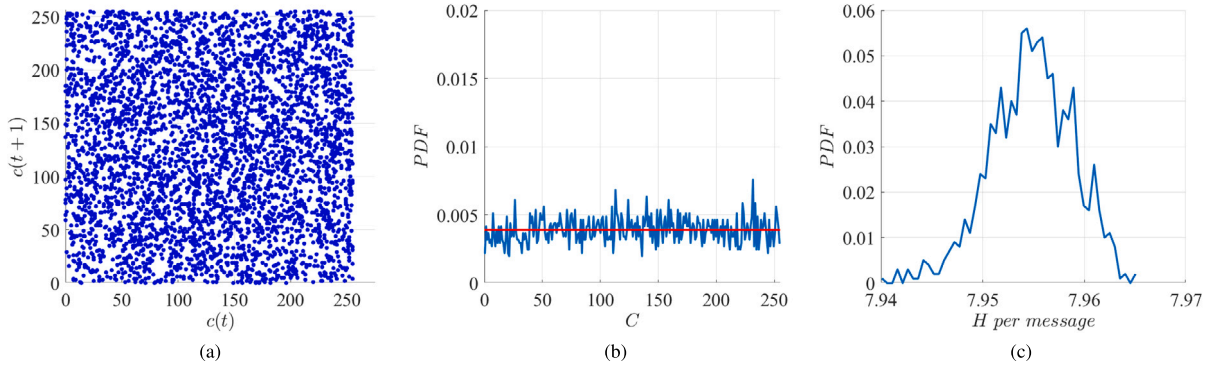


Fig. 7. (a) The recurrence plot of encoded/encrypted packets, (b) their corresponding PDF, and (c) the PDF of the entropy for 1000 iterations ($n = 8$ and $t = 4$).

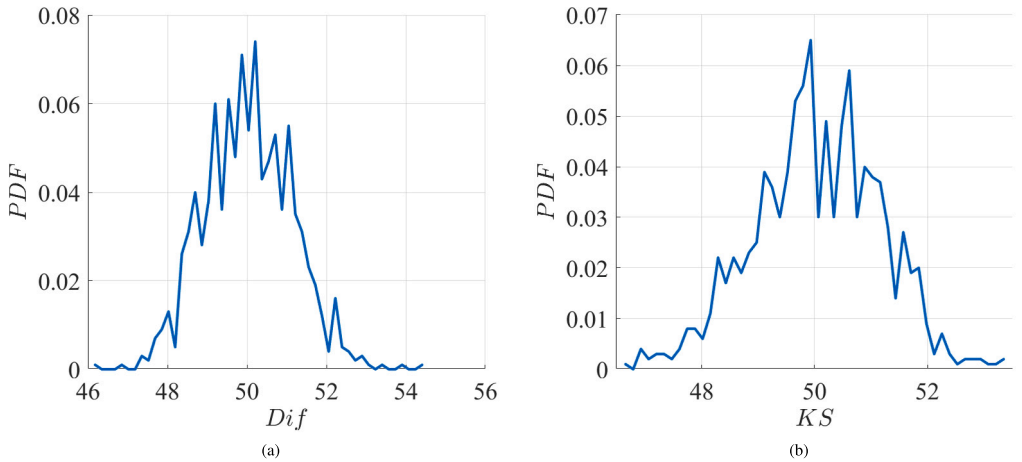


Fig. 8. (a) The PDF of the independence test which measures the bit difference between the original and encoded/encrypted packets for 1000 iterations ($n = 8$ and $t = 4$). (b) The PDF of key sensitivity test measuring the bit difference between two encoded/encrypted packets that are encrypted with two slightly different dynamic keys for 1000 iterations ($n = 8$ and $t = 4$).

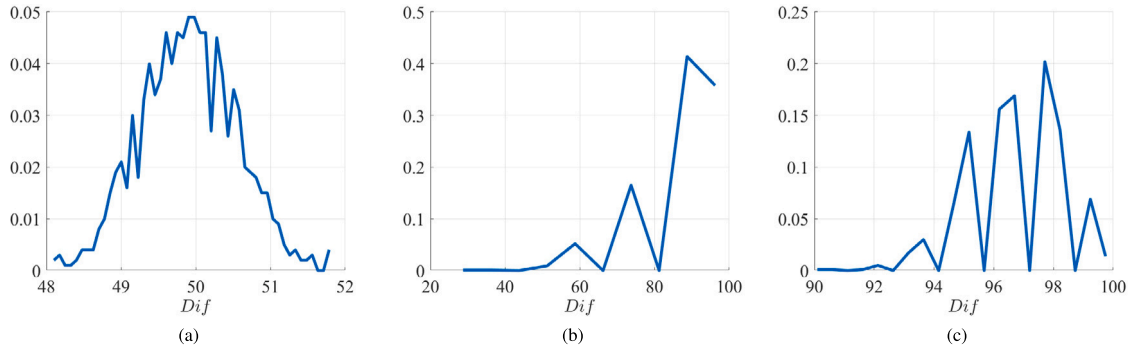


Fig. 9. The variation of the cryptographic primitive sensitivity test measuring the difference between the different cryptographic primitives obtained with two slightly different dynamic keys and for 1000 iterations ($n = 8$ and $t = 4$). (a) IDA matrices at the bit level, (b) matrix selection table and (c) link selection table, respectively.

parameters of the shared channels, acquiring and deriving the same dynamic key is very unlikely, especially that several nonce values are extracted from multiple channels and combined together (randomized nonce values). Regarding the downlink routing attack and the replay attack, the PLC-PRIME protocol already accounts for this type of attacks and it has been proven to resist this type of attacks. On the other hand, the beacon de-synchronization attack compromises a gateway and sends a large number of fake beacons (DoS attack). Since the proposed system model employs multiple wired and wireless links, users can discard the flooded link and data on that link, and recover the original data from the remaining channels. Here, it is very difficult for the attacker to jam all channels at the same time (wired and wireless). Finally, the ACK spoofing attack can be easily prevented by proper authentication, source authentication, message integrity and freshness of messages, all of which are accounted for in the PRIME protocol.

7. Performance analysis

In this section, the performance of the proposed scheme is evaluated in terms of computational complexity, execution time, communication overhead, transparency, flexibility, and error propagation.

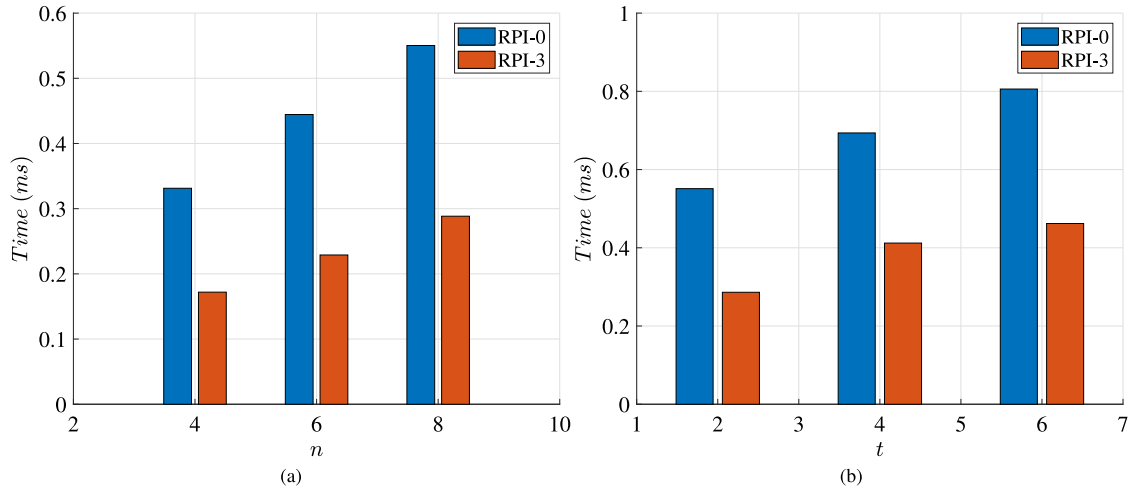


Fig. 10. Average execution time (ms) of the proposed encoding process (a) versus n for fixed $t = 2$ or (b) versus t with fixed $n = 8$.

7.1. Computational complexity

The proposed scheme is considered to have low computational complexity since it depends on simple operations in the encoding/decoding process. First, the input message is processed in the form of multiple matrices of flexible dimensions having t rows (variable, depending on channel conditions) and L_P columns (fixed) ($t \times L_P$). Each element in these matrices represents a different byte value. This step is performed after estimating the available channels and assigning the threshold values. Afterwards, each data matrix is multiplied by its corresponding encoding matrix of size ($n \times t$), separately. Here, parallelization can be employed to perform the matrix multiplication operation, which reduces the execution time and associated delays, significantly. This can be easily realized since each data matrix is encoded/decoded using unique cryptographic primitives, in an independent manner. Moreover, simple operations (mainly matrix multiplication) are required for the proposed confidentiality/availability scheme. Therefore, the computational complexity required to achieve data availability with IDA is relatively low [48].

7.2. Storage and communication overhead

The total size of the communicated matrices after applying the proposed solution is ($n \times L_P$), whereas the size of encoded matrices needed to recover the original message is ($t \times L_P$). Therefore, the encoding process introduces a data overhead equal to $(n - t) \times L_P \equiv \bar{t} \times L_P$ bytes, in order to benefit from IDA in terms of availability and resistance against channel errors. In other words, $(n - t)$ redundant encoded packets (rows) are produced and transmitted to prevent data loss due to damage or alteration. When the value of t is close to n, the communication overhead and availability level decrease. This means that the proposed solution exhibits a trade-off between communication overhead and availability level. In contrast, by decreasing the value of t beyond n, high data availability is achieved and high communication overhead is also produced. The value of n depends on the available wireless and wired channels. On the other hand, the value of t is related to impulsive noise and the number of allowable link failures.

7.3. Efficiency, transparency and flexibility

The proposed scheme utilizes simple operations, mainly, matrix multiplication, selection and permutation. These operations are only performed by the end nodes, and no further operations are needed at the gateways or the network servers (transparency). Therefore, in addition to maintaining existing PRIME security services such as data confidentiality, source authentication and data integrity, the proposed solution efficiently provides data and network availability, with appropriate computational complexity and communication overhead. Another significant feature of the proposed solution is that it is flexible and it adapts easily to any change in channel conditions as well as message length.

7.4. Execution time

Two experiments were carried out on a Raspberry Pie 0 (RPI-0) device and a Raspberry Pie 3 (RPI-3) device using C code. Fig. 10 presents the execution time of the proposed solution versus different values of n and t, for a packet length equal to 256. In the first experiment, the value of t is fixed and the value of n varies. In contrast, in the second experiment, the value of n is fixed and the value of t is varied. According to the results, it can be shown that if the values of n or t increase, the execution time also increases. This is logical since more time is required to process larger data matrices. The proposed solution can be seen to be very efficient since it introduces low computational overhead, and it allows fast encoding/decoding of data matrices. Another important observation is that the execution time decreases to half with RPI-3 compared to RPI-0. As a result, the proposed solution is practical for real-life implementations of PLC.

7.5. Resiliency against error propagation

Low error propagation is a mandatory requirement for the deployment of any security scheme. This is a hard challenge since these errors mainly result from interference and noise in transmission channels, leading to the destruction of transmitted data. In the proposed scheme, byte errors can be detected since they are limited to the erroneous encoded/encrypted packet. In order to overcome this issue, any t of the n available packets ($t \in n$ rows), that correspond to the channels having good performance can be selected in the decoding/decryption process. As a result, the proposed solution guarantees a high probability of correct data recovery with error less or equal to $(n - t)$ packets.

8. Conclusion

The intrinsic benefits of PLC along with the non-proprietary and open nature of the PRIME protocol have led to the deployment of millions of intelligent devices worldwide. The latest release of the standard included, among other features, an extension of the working band and an upgrade from one to a total of eight channels. The increase of available transmission options in PLC complements the efforts

combining PLC and RF into hybrid systems, offering yet an extra set of channels. Additional security mechanism could benefit from the new multi-channel paradigm as the current standard covers authenticity, confidentiality and integrity protection but has no specific mechanism for data availability. In this paper, we propose an effective scheme to prevent availability attacks on transmitted data, and overcome the destructive effects of impulsive noise. Particularly, we introduce an efficient solution to enhance the confidentiality and availability of data in PLC systems, thus, making the communication more robust, secure and reliable among end nodes. The solution is based on the Information Dispersal Algorithm (IDA) and the dynamic properties of physical channels. Security and performance tests have been performed, and the obtained results have confirmed the efficacy and robustness of the proposed solution. It has also been proven that the proposed method introduces a low overhead in terms of computations, complexity and resources, while ensuring a high level of availability and privacy.

CRediT authorship contribution statement

Hassan N. Noura: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing, Visualization, Project administration, Supervision, Investigation. **Reem Melki:** Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing, Visualization. **Ali Chehab:** Validation, Writing - original draft, Writing - review, Resources, Supervision, Project administration. **Javier Hernandez Fernandez:** Validation, Writing - original draft, Writing - Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This publication is supported by Iberdrola S.A., Qatar as part of its innovation department research studies. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of Iberdrola Group.

References

[1] Z. Sadowski, Comparison of PLC-PRIME and PLC-G3 protocols, in: International School on Nonsinusoidal Currents and Compensation (ISNCC), IEEE, 2015, pp. 1–6.
[2] M. Yan, J. Dai, P. Xu, The design and implementation of 128-bit AES encryption in PRIME, in: International Conference on Computer Science and Information Technology, Vol. 7, IEEE, 2010, pp. 345–348.
[3] ITU-T. Recommendation ITU-T G.9904; Narrowband orthogonal frequency division multiplexing power line communication transceivers for PRIME networks.
[4] A. Sanz, Evolution of PRIME to PLC-RF hybrid systems, in: Global Power, Energy and Communication Conference (GPECOM), 2019, pp. 74–79.
[5] G. Licciardo, et al., Evaluation of NB-PLC in railway environments, in: IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), IEEE, 2017, pp. 1–5.
[6] P. A. TWG. Specification for powerline intelligent metering evolution v.1.3.6. [Online]. Available: https://www.prime-alliance.org/wp-content/uploads/2013/04/PRIME-Spec_v1.3.6.pdf.
[7] P. A. TWG. Specification for powerline intelligent metering evolution v1.4. [Online]. Available: https://www.prime-alliance.org/wp-content/uploads/2014/10/PRIME-Spec_v1.4-20141031.pdf.
[8] P. Alliance. Virtual IEEE ISPLC 2020 – PRIME hybrid PLC-RF solution. [Online]. Available: https://www.prime-alliance.org/?p=7266&utm_source=rss&utm_medium=rss&utm_campaign=virtual-ieee-isplc-2020-prime-hybrid-plc-rf-solution.
[9] A. Sanz, Evolution of PRIME to PLC-RF hybrid systems, in: 2019 1st Global Power, Energy and Communication Conference (GPECOM), IEEE, 2019, pp. 74–79.
[10] A. Omri, J. Hernandez Fernandez, A. Sanz, M.R. Fliss, PLC channel selection schemes for OFDM-based NB-PLC systems, in: 2020 IEEE International Symposium on Power Line Communications and its Applications (ISPLC), 2020, pp. 1–6.
[11] M. Hoch, Comparison of PLC G3 and PRIME, in: Proc. IEEE International Symposium on Power Line Communications and its Applications, IEEE, 2011, pp. 165–169.
[12] A.M. Tonello, A. Pittolo, M. Girotto, Power line communications: Understanding the channel for physical layer evolution based on filter bank modulation, IEICE Trans. Commun. 97 (8) (2014) 1494–1503.
[13] J. Bilbao, et al., Reliable communications with network coding in narrowband powerline channel, in: Proc. IEEE International Symposium on Power Line Communications and Its Applications, IEEE, 2014, pp. 316–321.
[14] N. Andreadou, G. Fulli, NB-PLC channel: Estimation of periodic impulsive noise parameters and mitigation techniques, Int. J. Electr. Power Energy Syst. 103 (2018) 146–158.
[15] A. Ndjiongue, H. Ferreira, Power line communications (PLC) technology: More than 20 years of intense research, Trans. Emerg. Telecommun. Technol. 30 (7) (2019) e3575.
[16] M.R. Fliss, J. Hernandez Fernandez, A. Omri, G. Oligeri, NB-PLC successful transmission probability analysis, in: 2019 2nd International Conference on Smart Grid and Renewable Energy (SGRE), 2019, pp. 1–6.
[17] C. Cano, A. Pittolo, D. Malone, L. Lampe, A.M. Tonello, A.G. Dabak, State of the art in power line communications: From the applications to the medium, IEEE J. Sel. Areas Commun. 34 (7) (2016) 1935–1952.
[18] S. Galli, A. Scaglione, Z. Wang, For the grid and through the grid: The role of power line communications in the smart grid, Proc. IEEE 99 (6) (2011) 998–1027.
[19] A. Sendin, I. Peña, P. Angueira, Strategies for power line communications smart metering network deployment, Energies 7 (4) (2014) 2377–2420.
[20] M. Liserre, T. Sauter, J.Y. Hung, Future energy systems: Integrating renewable energy sources into the smart power grid through industrial electronics, IEEE Ind. Electron. Mag. 4 (1) (2010) 18–37.
[21] F. Mwasilu, J.J. Justo, E.-K. Kim, T.D. Do, J.-W. Jung, Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration, Renew. Sustain. Energy Rev. 34 (2014) 501–516.
[22] P. Kiedrowski, Toward more efficient and more secure last mile smart metering and smart lighting communication systems with the use of PLC/RF hybrid technology, Int. J. Distrib. Sens. Netw. 11 (10) (2015).
[23] S. Lin, W. Chung, An efficient (n, k) information dispersal algorithm based on fermat number transforms, IEEE Trans. Inf. Forensics Secur. 8 (8) (2013) 1371–1383.
[24] M. Rabin, Efficient dispersal of information for security, load balancing, and fault tolerance, J. ACM 36 (2) (1989) 335–348.
[25] L. Chen, T. Laing, K. Martin, Efficient, XOR-based, ideal (t, n)– threshold schemes, in: International Conference on Cryptology and Network Security, Springer, 2016, pp. 467–483.
[26] P. Singh, B. Raman, Reversible data hiding based on Shamir's secret sharing for color images over cloud, Inform. Sci. 422 (2018) 77–97.
[27] I. Tsokalo, R. Lehnert, F. Fitzek, Intraflow network coding on the data link layer for broadband PLC, in: International Symposium on Power Line Communications and its Applications (ISPLC), 2016, pp. 109–114.
[28] Y. Phulpin, J. Barros, D. Lucani, Network coding in smart grids, in: IEEE International Conference on Smart Grid Communications (SmartGridComm), 2011, pp. 49–54.
[29] S. Ezzine, F. Abdelkefi, J.P. Cances, V. Meghdadi, A. Bouallégue, Joint network coding and OFDMA based MAC-layer in PLC networks, in: 18th IEEE International Symposium on Power Line Communications and its Applications, 2014, pp. 311–315.
[30] R. Prior, D.E. Lucani, Y. Phulpin, M. Nistor, J. Barros, Network coding protocols for smart grid communications, IEEE Trans. Smart Grid 5 (3) (2014) 1523–1531.
[31] C. Chauvenet, et al., A communication stack over PLC for multi physical layer IPv6 networking, in: IEEE International Conference on Smart Grid Communications, IEEE, 2010, pp. 250–255.
[32] A. Ijaz, M. Rahman, O. Dobre, On safeguarding visible light communication systems against attacks by active adversaries, IEEE Photonics Technol. Lett. 32 (1) (2019) 11–14.
[33] A. Ndjiongue, T. Ngatched, O. Dobre, A. Armada, VLC-based networking: Feasibility and challenges, IEEE Netw. (2020).
[34] G. Pan, J. Ye, Z. Ding, Secure hybrid VLC-RF systems with light energy harvesting, IEEE Trans. Commun. 65 (10) (2017) 4348–4359.
[35] P. Grammatikis, P. Sarigiannidis, I. Moscholios, Securing the internet of things: Challenges, threats and solutions, Internet of Things (2018).
[36] S. Tomasin, S. Zulian, L. Vangelista, Security analysis of lorawan join procedure for internet of things networks, in: Wireless Communications and Networking Conference Workshops (WCNCW), IEEE, 2017, pp. 1–6.
[37] C. Salinesi, R. Mazo, O. Djebbi, D. Diaz, A. Lora-Michiels, Constraints: The core of product line engineering, in: International Conference on Research Challenges in Information Science (RCIS), IEEE, 2011, pp. 1–10.

[38] S. Na, D. Hwang, W. Shin, K. Kim, Scenario and countermeasure for replay attack using join request messages in LoRaWAN, in: International Conference on Information Networking (ICOIN), IEEE, 2017, pp. 718–720.
[39] J. Kim, J. Song, A simple and efficient replay attack prevention scheme for lorawan, in: Proceedings of the 2017 the 7th International Conference on Communication and Network Security, in: ICCNS 2017, ACM, New York, NY, USA, 2017, pp. 32–36, [Online]. Available: http://doi.acm.org/10.1145/3163058.3163064.
[40] E.V. Es, H. Vranken, A. Hommersom, Denial-of-service attacks on LoRaWAN, in: International Conference on Availability, Reliability and Security, 2018, pp. 1–6.
[41] A. Martínez, et al., Beacon frame spoofing attack detection in IEEE 802.11 networks, in: Third International Conference on Availability, Reliability and Security, IEEE, 2008, pp. 520–525.
[42] J. Zhang, T.Q. Duong, A. Marshall, R. Woods, Key generation from wireless channels: A review, IEEE Access 4 (2016) 614–626.
[43] M. De Piante, A.M. Tonello, Characteristics of the PLC channel: Reciprocity, symmetry and port decoupling for impedance matching, in: 2016 International Symposium on Power Line Communications and its Applications (ISPLC), 2016, pp. 93–97.
[44] F. Passerini, A.M. Tonello, Secure PHY layer key generation in the asymmetric power line communication channel, Electronics 9 (4) (2020) 605.
[45] W. Henkel, O.A. Graur, N.S. Islam, U. Pagel, N. Manak, O. Can, Reciprocity for physical layer security with wireless FDD and in wireline communications, in: 2018 IEEE Globecom Workshops (GC Wkshps), 2018, pp. 1–6.
[46] R. Melki, M. Mansour, A. Chehab, A fairness-based congestion control algorithm for multipath TCP, in: Proc. IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2018, pp. 1–6.
[47] H. Noura, A. Chehab, L. Sleem, M. Noura, R. Couturier, M. Mansour, One round cipher algorithm for multimedia IoT devices, Multimedia Tools Appl. 77 (14) (2018) 18383–18413.
[48] H. Noura, O. Salman, A. Chehab, R. Couturier, Preserving data security in distributed fog computing, Ad Hoc Netw. 94 (2019) 101937.

Hassan Noura received his degree in Computer and Communication Engineering from the IUL University in 2008, Lebanon, and his Ph.D. degree from Polytech'Nantes in 2012, France. In 2013, Noura joined Paris-Sud XI University, as a postdoctoral researcher, after that as research engineering at CEA, Grenoble. After that, he joined QMIC in 2015 and Telecom ParisTech in 2016. Recently, he gets his HDR in 2016 and joins AUB in 2017. His research interests include cryptography, network security, secure network coding, secure multimedia and secure distributed system.

Reem Melki received her B.S. degree in Electrical and Computer Engineering from the Rafik Hariri University in 2013, and her M.S. degree in 2015. She has received her Ph.D. degree in August 2020 from the American University of Beirut (AUB). Her main areas of research interests include security and privacy in wireless and mobile communication.

Ali Chehab received his Bachelor degree in EE from AUB in 1987, the Master's degree in EE from Syracuse University in 1989, and the Ph.D. degree in ECE from the University of North Carolina at Charlotte, in 2002. From 1989 to 1998, he was a lecturer in the ECE Department at AUB. He rejoined the ECE Department at AUB as an Assistant Professor in 2002, became Full Professor in 2014. He received the AUB Teaching Excellence Award in 2007. He teaches courses in Programming, Electronics, Digital Systems Design, Computer Organization, Cryptography, and Digital Systems Testing. His research interests include: Wireless Communications Security, Cloud Computing Security, Multimedia Security, Trust in Distributed Computing, Low Energy VLSI Design, and VLSI Testing. He has about 210 publications. He is a senior member of IEEE and a senior member of ACM.

Javier Hernandez Fernandez Mr. Javier Hernandez is the Technical Director of Iberdrola Innovation Middle East and has over 15 years' experience in telecoms, utilities and engineering companies specializing in the areas of smart grids, communications, bespoke software development and project management. Javier holds a B.Sc. in Computer Science from the Faculty of Engineering of the University of Ottawa (Canada), along with two Master Degrees in Energy Management (University of Zaragoza - Spain) and Project Management (Business School of the University San Pablo CEU / IEP - Spain). Along with his experience in the private sector, he has an extensive record in research including positions in the University of Ottawa and Altran/Telefonica R&D. Now he leads research and technical consulting projects as part of the Innovation team of Iberdrola.