# CypherChain: A Privacy-Preserving Data Aggregation Framework for Blockchain-Based DR Programs

***Abstract*—The integration of Distributed Energy Resources (DERs) into smart grids has significantly enhanced power capabilities but introduced complexities and potential voltage instability issues. Demand Response (DR) programs are crucial for efficient DER management in smart grids, yet they face privacy, transparency, and trust challenges in data management and energy transactions. Although blockchain technology promises decentralized trust and immutable transparency, it does not address privacy challenges due to its inherent transparency feature. This paper introduces 'CypherChain', a novel framework designed to preserve privacy in blockchain-based DR programs. CypherChain highlights two major innovations: the Clustered Hypergraph Aggregation and Interaction (CHAIN) protocol and the Collaborative Cyphertext Processing (CYPHER) protocol. The CHAIN protocol, leveraging Homomorphic Encryption and Hypergraph Coloring, ensures efficient and confidential data aggregation. Concurrently, the CYPHER protocol, rooted in Secure Multi-Party Computation (SMPC), facilitates cooperative encrypted data (Cyphertext) processing while preserving the privacy of individual data points. Our extensive evaluation, utilizing a real-world dataset from a smart building in Newcastle, Australia, demonstrates CypherChain's remarkable performance, achieving a 40% increase in data aggregation speed and a 30% reduction in computational costs compared to existing systems. These results highlight CypherChain's effectiveness in revolutionizing privacy-preserving mechanisms in smart grid environments, offering a comprehensive solution to privacy and transparency issues in blockchain-based DR programs.**

***Index Terms*—Blockchain Technology, Data Privacy, Demand Response Programs, Secure Multiparty Computation, Homomorphic Encryption, Hypergraph Coloring**

## I. Introduction

Demand Response (DR) programs have become integral to the modernization of smart grids, facilitating dynamic energy management and the integration of Distributed Energy Resources (DERs) [1]. These programs play a pivotal role in balancing energy supply and demand by actively involving prosumers in energy consumption and production decisions [2]. Prosumers are entities that both produce and consume electricity using small-scale DERs such as roof-top solar panels, windmills, energy storage systems, and controllable loads [3]. However, implementing DR programs is challenging, particularly in maintaining data privacy and ensuring transparency in transactions and interactions. To illustrate the severity of these challenges, consider the scenario of a large metropolitan area deploying DR strategies to manage peak loads depicted in Figure 1. Here, sensitive data, such as energy generation and usage patterns and DER operational data (control signals), handled by an aggregator are vulnerable to privacy breaches,

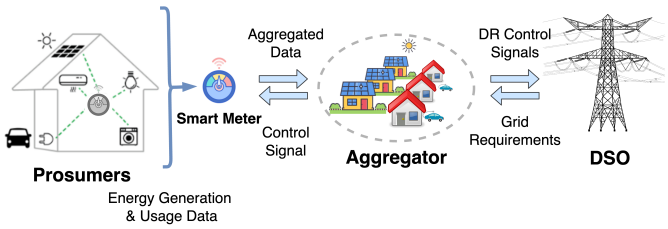


Fig. 1. A large metropolitan area deploying DR strategies to manage peak loads in a modern smart grid.

raising concerns over user confidentiality [4]. The aggregator acts as an intermediary, consolidating the energy contributions from multiple prosumers on behalf of the Distributed System Operator (DSO) [5]. Moreover, the transparency necessary for efficient market operations often conflicts with the need for privacy, underscoring the need for a solution that reconciles privacy with operational transparency [6].

Blockchain technology, a distributed ledger system, has gained prominence in smart grid applications, offering a new paradigm for managing energy data and transactions in DR programs [7]. The decentralization characteristic of blockchain facilitates a trustless environment, eliminating the need for central authority and thereby enhancing the resilience of energy systems [8]. Additionally, blockchain's inherent transparency and security features support the validation of transactions and the authentication of data, crucial for the integrity of DR programs [9]. However, while blockchain offers numerous advantages, it also presents specific challenges, particularly in the context of privacy. The inherent transparency of blockchain leads to a trade-off between privacy and transparency, wherein every transaction logged on the blockchain is accessible to all participants in the network. This visibility could inadvertently reveal sensitive information such as energy usage habits or operational specifics of DERs [3]. This tradeoff poses a significant challenge, as it can impact the willingness of prosumers to participate in DR programs and affect the overall effectiveness of these initiatives [9]. The need to balance privacy with transparency in blockchain applications is thus a critical research area that requires innovative solutions to ensure the successful deployment of blockchain in smart grid environments.

Recent advancements in blockchain-based Demand Re-

sponse (DR) programs and smart grid systems have led to various privacy-preserving solutions, each with its own trade-offs. For example, private or permissioned blockchains, as suggested by Androulaki et al. [10], enhance privacy but risk revealing sensitive competitive data, such as prosumer bids, in market-driven environments. Similarly, off-chain storage methods, explored by Dorri et al. [11], protect data but can fragment the system, compromising transparency and data aggregation efficiency. Furthermore, Zero-Knowledge Proofs (ZKPs), highlighted for their confidentiality preservation, introduce increased computational complexity, as pointed out by Goldreich [12]. This complexity negatively impacts system efficiency. Additionally, Homomorphic Encryption (HE), enabling computation on encrypted data, preserves privacy but can limit transparency in multi-party scenarios as encrypted outputs are not directly interpretable by all parties [13]. Lastly, Secure Multi-Party Computation (SMPC), facilitating joint computation over distributed datasets, addresses privacy concerns but suffers from computational and implementation challenges in DR contexts, as noted by Wang et al. [14]. Collectively, these approaches progress blockchain privacy, but often at the expense of either transparency or efficiency, underscoring the necessity for a more integrated solution that adeptly balances all three aspects of smart grid operations.

In response to these challenges, we propose 'CypherChain', an innovative framework designed to address the limitations of existing blockchain-based solutions in DR programs and smart grids. CypherChain introduces two major components: the Clustered Hypergraph Aggregation and Interaction (CHAIN) protocol and the Collaborative Cyphertext Processing (CYPHER) protocol. The CHAIN protocol, leveraging HE, ensures efficient and confidential data aggregation. Additionally, the CHAIN protocol leverages Hypergraph Coloring to effectively segment and store data, significantly improving the framework's scalability and operational efficiency [4]. On the other hand, the CYPHER protocol, rooted in SMPC, facilitates transparent data processing while preserving the privacy of individual data points. Together, these components address the critical balance between privacy and transparency in blockchain-based DR systems, providing a comprehensive solution for privacy-preserving data management in smart grid environments. The specific contributions of CypherChain are as follows:

1) We propose a Clustered Hypergraph Aggregation and Interaction (CHAIN) protocol that utilizes HE and Hypergraph Coloring. This protocol addresses the specific challenge of private data aggregation from multiple sources while maintaining efficient data management. CYPHER ensures that encrypted data computations safeguard user confidentiality, and its adaptive segmentation algorithm optimizes data handling for scalability and operational efficiency.
2) We propose the Collaborative Hypergraph Aggregation and Interaction (CHAIN) protocol, based on SMPC. This protocol addresses the gap in cooperative data processing among multiple parties without revealing individual data inputs. CHAIN maintains both privacy and data integrity, ensuring transparent and secure data sharing in DR programs.
3) We present a comprehensive analysis and evaluation of CypherChain. Utilizing real-world smart building data, our evaluation demonstrates that CypherChain outperforms conventional systems, achieving a 40% faster data aggregation speed and a 30% reduction in computational costs. This substantial operational efficiency, demonstrated in a practical smart building context, highlights CypherChain's effectiveness in real-world DR applications.

The rest of the paper is organized as follows: Section II provides the background and related work, laying the foundation for the terminology and concepts discussed. Section III details the proposed CypherChain framework, including its unique contributions. Section IV presents a comprehensive implementation and evaluations of CypherChain of the framework. Finally, Section V concludes the paper and outlines future directions.

## II. Background and Related Work

This section examines the integration of blockchain technology into DR programs, emphasizing its critical role in managing DERs within smart grids. The discussion includes an analysis of the privacy challenges inherent in blockchain-based systems, a review of current privacy preservation solutions, and an identification of their limitations. Additionally, this section offers a succinct overview of HE, SMPC, and Hypergraph Coloring, detailing their significance and implementation in mitigating these challenges, thereby improving the efficacy and privacy of DR programs.

### *A. Blockchain in Demand Response Programs*

The role of blockchain in DR programs extends beyond mere technological innovation; it revolutionizes how energy systems interact and operate. Blockchain's decentralized nature inherently aligns with the distributed framework of DR programs, where it facilitates a more dynamic and transparent energy distribution process. Samadi et al. emphasize blockchain's ability to reduce demand efficiently, save on DERs surplus generation, and incentivize customers through monetary rewards, thereby enhancing customer engagement in DR programs [22]. Integrating DERs, such as solar panels and energy storage systems, into smart grids is crucial for sustainable energy management. Blockchain technology provides a robust platform for this integration, offering features like traceability, decentralization, and immutability. Deshpande et al. highlight how blockchain not only brings transparency to DR marketplaces but also balances the metrics of smart grid management, ensuring efficient operation and control [23].

### *B. Privacy Challenges in Blockchain-based DR*

Blockchain technology's application in DR programs, while advantageous for energy management, introduces specific privacy issues that must be addressed. One of the primary

TABLE I
SUMMARY OF PRIVACY SOLUTIONS IN BLOCKCHAIN-BASED DR PROGRAMS

| Reference | Technology | Application in DR | Key Contributions | Identified Gaps |
|---|---|---|---|---|
| Li et al. [15] | Homomorphic Encryption (HE) | Smart grid communications | Introduced EPPDR scheme combining HE with adaptive key evolution for enhanced privacy | Computationally intensive, impacting processing speed and resource demands |
| Bos et al. [16] | Partially Homomorphic Encryption (PHE) | Privacy-friendly forecasting in smart grids | Explored PHE schemes for privacy while enabling data computations | Balancing security with computational efficiency |
| He et al. [17] | Partially Homomorphic Encryption (PHE) | Secure data exchange in smart grids | Demonstrated PHE for secure and efficient data exchange | Need for improved processing speed in PHE applications |
| Alexandru et al. [18] | Secure Multiparty Computation (SMPC) | Cloud-based control systems in smart grids | Enhanced privacy in multi-party environments using SMPC | Challenges in coordination among multiple parties |
| Vogelsang et al. [19] | Garbled Circuits, SPDZ | Time-to-event analyses | Suggested potential of garbled circuits and SPDZ for privacy-preserving solutions | Resource-intensive management of complex SMPC systems |
| Cao et al. [20] | Garbled Circuits | Improvement in secure computations | Studied the efficiency of garbled circuits in secure computations | Enhancing computation and security in SMPC techniques |
| Karumba et al. [21] | Hypergraph Coloring | Decentralized energy trading systems | HARB framework addressing scalability, privacy, interoperability in blockchain systems | Adapting hypergraph techniques for DR program applications |

concerns is the exposure of sensitive information regarding users' energy consumption patterns. Bracciale et al. emphasize the risk of publishing such sensitive data within blockchain-based energy trading systems, highlighting the potential for privacy violations [24]. Moreover, Zhou et al. identify privacy leakage and security threats as major challenges in blockchain-based DR programs, particularly in systems involving the Internet of Electric Vehicles. These challenges arise from the detailed data required for effective DR management, which, if not properly secured, can compromise user privacy [9].

Another significant aspect is the balance between data transparency and user privacy. As blockchain inherently promotes transparency, ensuring user privacy within this transparent ecosystem is complex. Ghasemkhani et al. highlights this issue in incentive-based DR programs, where fine-grained power consumption data can inadvertently reveal user behaviour patterns [25]. Furthermore, implementing secured data mechanisms that allow for on-demand data disclosure is essential to address these privacy concerns. Aslam et al. [26] and Pei et al. [27] both discuss the need for blockchain-based systems in DR to have robust cryptographic algorithms and proper authentication controls to maintain a balance between transparency and privacy.

### C. Review of Existing Privacy Solutions

The landscape of privacy-preserving solutions in DR programs, particularly those leveraging blockchain technology, has seen significant innovations. However, each comes with its unique set of strengths and limitations. We focus on the following three key approaches:

*1) Homomorphic Encryption (HE):* HE has emerged as a cornerstone in developing privacy-preserving solutions for smart grid DR programs. Li et al. introduced the EPPDR scheme, which combines HE with adaptive key evolution, offering enhanced user privacy and cybersecurity in smart grid communications. This scheme represents a significant stride in protecting user data within DR scenarios [28]. However, the computationally intensive nature of HE, particularly in fully homomorphic encryption (FHE) forms, presents challenges in terms of processing speed and resource demands. Variations of HE, such as the Fan-Vercauteren somewhat homomorphic encryption and the partially homomorphic encryption (PHE) schemes, have also been explored for their fit in smart grid applications. Bos et al. discuss using these schemes for privacy-friendly forecasting in smart grids, highlighting their potential to preserve privacy while allowing necessary data computations [16]. Similarly, He et al. demonstrates using PHE for secure and efficient data exchange in smart grids, emphasizing its balance between security and computational efficiency [17].

*2) Secure Multiparty Computation:* SMPC has been increasingly used in privacy-preserving solutions, especially in the context of smart grid DR programs. SMPC allows multiple parties to collaboratively compute a function over their inputs while keeping those inputs private. Alexandru and Pappas highlight using SMPC in cloud-based control systems, enhancing privacy in multi-party environments [18]. However, managing complex coordination among multiple parties in SMPC systems can be challenging and resource-intensive. Various SMPC techniques have been developed to address these issues, including garbled circuits, homomorphic secret sharing, and SPDZ (Speedy Data Zero). For instance, Vogelsang et al.discuss using garbled circuits and SPDZ in time-to-event analyses, suggesting their potential for privacy-preserving solutions in DR programs [19]. Additionally, Cao et al. study the improvement of computation and security in garbled circuits, emphasizing their efficiency in secure computations [20].

*3) Hypergraph Coloring:* Hypergraph Coloring is a technique used for data segmentation and distribution in complex network systems, such as those found in Demand Response (DR) programs. A hypergraph is a generalization of a graph where edges can connect any number of vertices, not just two. Coloring in this context refers to assigning colors to vertices in such a way that no two connected vertices share the same color, thereby managing data dependencies and interactions efficiently. While the specific application in DR programs is not extensively covered in the existing literature, the HARB framework by Karumba et al. illustrates the potential of hypergraphs in blockchain-based systems. The HARB framework enhances blockchain-based decentralized energy trading systems, addressing scalability, privacy, and interoperability challenges. This approach could be adapted for DR programs to optimize data handling, reduce computational demands, and enhance scalability, especially in scenarios involving complex interdependencies of data and operations [21].

## III. CypherChain Framework

This section presents the CypherChain framework, an innovative approach to enhancing privacy, transparency, and efficiency in blockchain-based DR programs. CypherChain addresses the inherent privacy challenges within transparent and decentralized blockchain systems, particularly in managing sensitive energy data. By re-imagining traditional blockchain architecture, CypherChain offers a tailored solution that balances the transparency benefits of blockchain with the critical need for individual data privacy. While its design is application-agnostic, we exemplify its application in residential DR programs.

### A. Architecture

The CypherChain architecture represents an innovative integration of HE, SMPC, Hypergraph Coloring, and smart contracts. These technologies are strategically layered across the blockchain architecture to enhance privacy and efficiency in DR programs. As illustrated in Figure 2, this architecture comprises four key layers, each contributing to the system's overall functionalities of decentralisation, security, privacy, and transparency.

- **Data Layer:** This layer forms the blockchain's foundation, managing data storage and organization. It includes a blockchain (chain-of-blocks), transactions, and a Merkle tree. CypherChain uses Homomorphic Encryption (HE) in this layer for encrypting transaction data, ensuring confidentiality on the blockchain.
- **Network Layer:** Responsible for node communication in the blockchain network, this layer involves nodes, peer-to-peer networks, and protocols. CypherChain uses a hypergraph network model, dividing the network into color-coded clusters for efficient data exchange and managing scalability. The CHAIN protocol is proposed to support private data aggregation across the network.
- **Consensus Layer:** This layer maintains blockchain consensus and transaction validation. It includes consensus algorithms, miners (in proof-of-work (PoW) consensus), endorsers, orderers, and committers (in Practical Byzantine Fault (PBFT) consensus). CypherChain introduces the SMPC-based CYPHER protocol, validating encrypted transactions without revealing private inputs.
- **Application Layer:** The Application Layer is where decentralized applications (DApps) and user interactions are built on top of the blockchain infrastructure. Key components in this layer include DApps, User Interfaces, APIs, and smart contracts. The smart contracts ensure that operations such as calculating incentives or validating energy-saving measures are executed efficiently and transparently without compromising privacy.

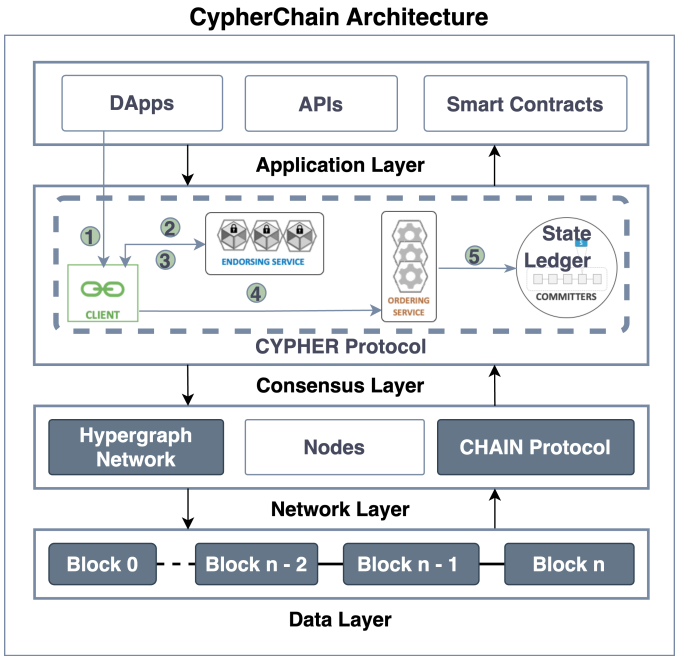


Fig. 2. Layered Architecture of the CypherChain Framework

Integrating these layers, the CypherChain architecture presents a privacy-centric blockchain solution designed for contemporary DR programs. It aims to tackle the privacy issues commonly associated with traditional blockchain systems, positioning CypherChain as a significant contribution to energy management.

### B. Clustered Hypergraph Aggregation and Interaction (CHAIN)

Implementing the CHAIN protocol is a strategic process aimed at enhancing data exchange efficiency and privacy. This approach involves segmenting the blockchain network into distinct clusters using a hypergraph coloring algorithm. The development of this protocol can be structured as follows:

*1) Hypergraph-based Network Model:* In CypherChain, the hypergraph $H = (V, E)$ is defined where $V$ represents nodes or embedded computers owned by residential prosumers ($P_i$), aggregators ($A_j$), and the DSO. The hyperedges $E$ symbolize the connections or relationships among these nodes. These

TABLE II
NOTATIONS USED IN THE CYPHERCHAIN FRAMEWORK SECTION

| Notation | Description |
|---|---|
| $H$ | The hypergraph representing the blockchain network |
| $V$ | Set of vertices or nodes in the hypergraph |
| $E$ | Set of hyperedges in the hypergraph |
| $\mathcal{C}$ | Hypergraph coloring algorithm |
| $P_i$ | Prosumer $i$ in the network |
| $A_j$ | Aggregator $j$ in the network |
| $DSO$ | Distributed System Operator |
| $G_k$ | Geographical proximity criterion |
| $ECP_l$ | Energy consumption pattern criterion |
| $DRE_m$ | Demand Response Event type $m$ |
| $pk_i$ | Public key of prosumer $i$ |
| $sk_i$ | Private key of prosumer $i$ |
| $K_i$ | Cluster $i$ in the network |
| $D_i$ | Energy data of prosumer $i$ |
| $E_{pk_i}(D_i)$ | Encrypted data of prosumer $i$ |
| $s_{ij}$ | Share $j$ of prosumer $i$'s encrypted data |
| $t$ | Threshold in secret sharing scheme |
| $L_{\text{total}}$ | Total load reduction computed result |
| $\lambda$ | Security parameter for key generation |
| $C$ | Set of colors used in hypergraph coloring |
| $c_i$ | Specific color assigned to a cluster or node |
| $T$ | Coloring scheme type (e.g., vertex, hyperedge) |
| $h$ | Aggregation function applied to homomorphically computed results |
| $g_i$ | Homomorphic function applied by prosumer $i$ |
| $n$ | Total number of prosumers or nodes in a cluster |

relationships are formulated based on criteria such as geographic proximity ($G_k$), energy consumption patterns ($ECP_l$), or participation in specific DR events ($DRE_m$), creating a multi-dimensional network structure optimized for the DR context.

*2) Hypergraph Coloring Algorithm:* In the CHAIN protocol, the hypergraph coloring algorithm $\mathcal{C}$ is applied to the defined hypergraph $H = (V, E)$ network to assign colors to either the nodes $V$ or hyperedges $E$, depending on the chosen coloring scheme. This algorithm operates under the constraint that adjacent nodes $v_i, v_j \in V$ or intersecting hyperedges $e_k, e_l \in E$ must not have the same color. Formally, for vertex coloring, if $v_i$ and $v_j$ are adjacent, then $\mathcal{C}(v_i) \neq \mathcal{C}(v_j)$; for hyperedge coloring, if $e_k$ and $e_l$ intersect, then $\mathcal{C}(e_k) \neq \mathcal{C}(e_l)$. This coloring scheme ensures the segmentation of the network into distinct clusters based on specific DR program criteria such as demand response event participation, energy usage profiles, or geographic location. By doing so, the algorithm $\mathcal{C}$ effectively optimizes network segmentation to enhance operational efficiency and privacy within the DR program context. The algorithm $\mathcal{C}$, as described in Algorithm 1, plays a crucial role in the CypherChain architecture, facilitating the creation of a privacy-enhanced, efficient network tailored for DR applications.

*3) Segmentation and Cluster Formation:* Segmentation and Cluster Formation in CypherChain utilizes the hypergraph coloring algorithm $\mathcal{C}$ to create distinct clusters $K_i$ within the network. Each cluster $K_i$ comprises nodes $v \in V$ sharing the same color, facilitating localized data processing and enhancing privacy. For instance, nodes participating in the same DR event or connected to a specific aggregator are grouped, minimizing data exposure across the network. Clusters are defined based on DR criteria like event types $DRE_m$, geographic locations $G_k$, or aggregator associations $A_j$, as illustrated in Figure 3. This segmentation results in efficient, privacy-preserving sub-networks tailored to the operational requirements of DR programs.

**Algorithm 1** Hypergraph Coloring for Network Segmentation

**Require:** Hypergraph $H = (V, E)$, Coloring Scheme $T$
**Ensure:** Colored Hypergraph $H_c = (V, E, \mathcal{C})$

**function** COLORHYPERGRAPH($H, T$)
  Initialize color set $C$
  **if** $T$ is Vertex Coloring **then**
    **for** each vertex $v \in V$ **do**
      Assign color $\mathcal{C}(v) \in C$, ensuring no two adjacent vertices share the same color
    **end for**
  **else if** $T$ is Hyperedge Coloring **then**
    **for** each hyperedge $e \in E$ **do**
      Assign color $\mathcal{C}(e) \in C$, ensuring no two intersecting hyperedges share the same color
    **end for**
  **end if**
  **return** Colored Hypergraph $H_c = (V, E, \mathcal{C})$
**end function**

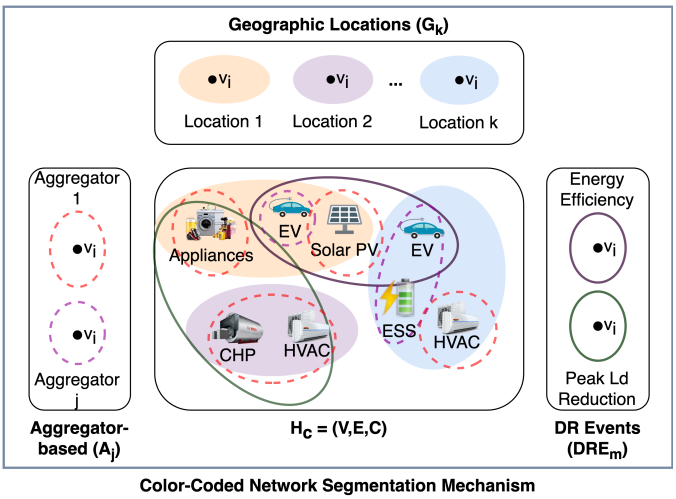


Fig. 3. Illustration of CypherChain's hypergraph coloring algorithm $\mathcal{C}$, forming clusters $K_i$ based on DR event types $DRE_m$, geographical locations $G_k$, or aggregator connections $A_j$, enhancing privacy and efficiency in DR program operations.

### C. Collaborative Cyphertext Processing (CYPHER) protocol

The CYPHER protocol in the CypherChain framework is designed to aggregate data securely and privately using PHE and HSS schemes. Here is a detailed explanation of each step:

1: **Key Generation and Distribution:** In the initial network setup, each prosumer ($P_i$) creates a unique public-private key pair ($pk_i, sk_i$) for authentication and data encryption.

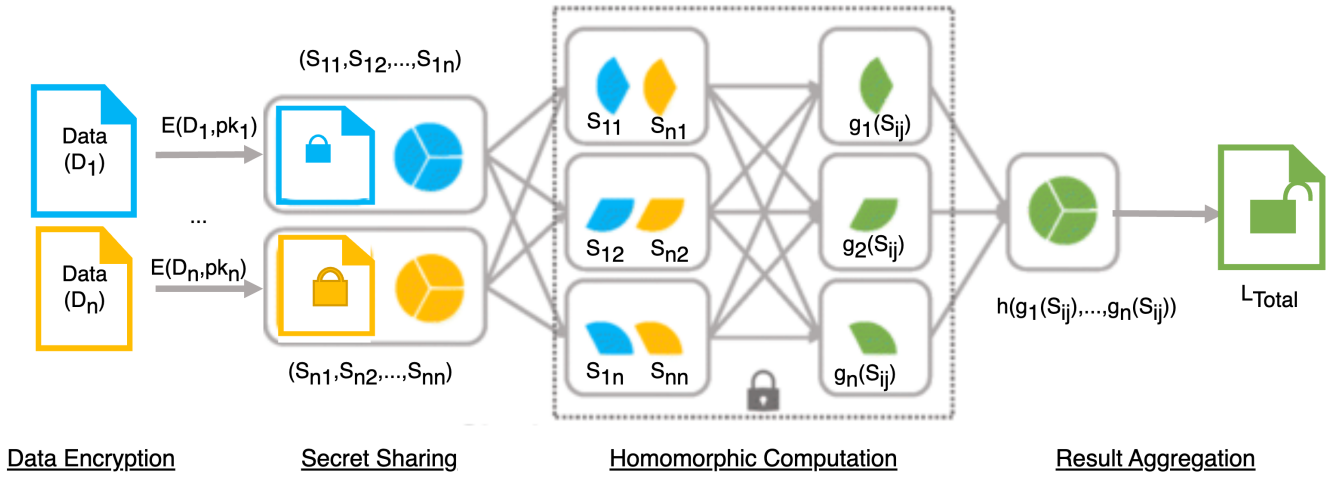


Fig. 4. The Privacy-Preserving Data Processing Protocol within the CypherChain framework.

Public keys ($pk_i$) are distributed network-wide, allowing nodes to encrypt data for $P_i$, while private keys ($sk_i$) are privately held for decryption. This process is vital for secure data encryption and aggregation. The mathematical formulation of key generation is: $(pk_i, sk_i) \leftarrow \text{KeyGen}(\lambda)$. This foundational step ensures data privacy and security, supporting non-repudiation in the CypherChain framework.

2: **Network Segmentation and Cluster Formation:** Using the Hypergraph Coloring Algorithm ( $\mathcal{C}$), CypherChain divides its network into clusters ($K_i$) based on DR events, location, and aggregator ties. Nodes or hyperedges in hypergraph $H = (V, E)$ are color-coded, forming clusters with similarly colored nodes. For instance, prosumers in a peak load reduction event ($DRE_{plr}$) form cluster $K_{plr}$. The cluster formula is $K_i = \{v \in V | \mathcal{C}(v) = c_i\}$. This segmentation boosts efficiency and privacy by localizing interactions, reducing wide data spread.

3: **Data Encryption:** Upon the formation of clusters, each prosumer within a cluster $K_i$ encrypts their individual energy data $D_i$ using their public key $pk_i$, resulting in encrypted data $E_{pk_i}(D_i)$. This encryption occurs at the prosumer's node, ensuring the confidentiality of data before it enters the CypherChain network. This is represented mathematically as: $E_{pk_i}(D_i) = \text{Encrypt}_{pk_i}(D_i)$

4: **Secret Sharing:** Following the encryption, each prosumer's data is split into secret shares using a secure sharing scheme. This step is fundamental for the HSS process. If the total number of prosumers in a cluster $K_i$ is $n$, and the encrypted data of a prosumer $P_i$ is $E_{pk_i}(D_i)$, the data is divided into $n$ shares $s_{i1}, s_{i2}, \ldots, s_{in}$. Each share $s_{ij}$ is then distributed to prosumer $P_j$ within the same cluster. The secret sharing is mathematically represented as: $E_{pk_i}(D_i) \rightarrow (s_{i1}, s_{i2}, \ldots, s_{in})$. This process ensures that no single prosumer or a subset of prosumers can reconstruct the original data unless a certain threshold of shares is combined, thereby preserving the confidentiality of individual prosumer data.

5: **Homomorphic Computation:** Each prosumer in the cluster $K_i$ engages in a homomorphic computation on their respective shares of encrypted data. For example, in a peak load reduction DR event, prosumers calculate their shares to contribute to the collective assessment of load reduction without revealing their individual data. Let $g_i$ be the homomorphic function applied by prosumer $P_i$ on their share $s_{ij}$. This function is aligned with the overall objective of calculating the total load reduction in the cluster. The computation by each prosumer is represented as: $g_i(s_{ij})$ for each $j = 1, 2, \ldots, n$. These homomorphic computations ensure that the operations on encrypted shares preserve the properties necessary for the final aggregation. The collective computation across all prosumers in $K_i$ forms the basis for deriving the total impact of the DR event while maintaining the confidentiality of individual contributions.

6: **Result Aggregation:** The final step in the protocol is to aggregate the homomorphically computed results from each prosumer in cluster $K_i$ to derive the overall impact of the DR event, such as peak load reduction. This is accomplished by the aggregator, who collects the results $g_i(s_{ij})$ from all prosumers and applies an aggregation function $h$ to calculate the total load reduction. The aggregated result is symbolized as $L_{\text{total}}$ and is computed as follows: $L_{\text{total}} = h(g_1(s_{1j}), g_2(s_{2j}), \ldots, g_n(s_{nj}))$. Here, $L_{\text{total}}$ represents the total amount of load reduced during the DR event, and $j$ denotes the number of prosumers in the cluster. This aggregation ensures that the cumulative effect of the DR event is accurately calculated while preserving the confidentiality of individual prosumer data.

This protocol ensures that individual prosumer data remains confidential throughout the aggregation process. By leveraging PHE for data encryption and HSS for secure computation, the CypherChain framework addresses the challenge of maintaining privacy in a transparent blockchain environment.

## IV. Implementation and Evaluation

This section delves into the practical implementation and comprehensive evaluation of the CypherChain framework. It presents a proof-of-concept (PoC) implementation and a detailed analysis of the framework's security and privacy features. It concludes with a performance evaluation to assess its efficiency and scalability.

### A. Proof-of-Concept Implementation

In our PoC, Raspberry Pi 4s acted as network nodes, representing DERs in a residential DR context. Each Pi had a Quad-core Cortex-A72 SoC, 4GB RAM, and 32GB storage. A stronger 'computing node' representing a DSO node with an Intel Core i7, 16GB RAM, and 1TB HDD handled complex blockchain tasks, akin to a utility company's role in a smart grid. Aggregator nodes ran on Heroku's cloud, providing resources like multi-core CPUs, over 8GB RAM, SSDs, and fast networks for data processing. Hyperledger Fabric was used for blockchain operations, supporting private transactions for our DR scenario. Golang developed smart contracts for DR, and cryptographic libraries, PHE and HSS, were integrated for secure, private data aggregation in CypherChain.

| time | net | tamb | demand | DR |
|---|---|---|---|---|
| 2020-12-22 | 153.5 | 19.8 | 1.5 | 0.0 |
| 2020-12-22 | 233.0 | 19.9 | 61.6 | 0.0 |
| 2020-12-22 | 226.5 | 20.1 | 49.6 | 0.1 |
| 2020-12-22 | 244.0 | 20.4 | 66.9 | 0.1 |
| 2020-12-22 | 258.8 | 20.6 | 73.1 | 0.5 |
| 2020-12-22 | 262.3 | 20.9 | 83.0 | 0.5 |
| 2020-12-22 | 247.3 | 21.4 | 72.5 | 1.1 |
| 2020-12-22 | 199.8 | 22.2 | 50.1 | 0.8 |
| 2020-12-22 | 214.3 | 22.9 | 57.1 | 1.5 |
| 2020-12-22 | 225.8 | 23.5 | 57.2 | 1.6 |

Fig. 5. Enter Caption

We simulated a Demand Response (DR) event aimed at peak load reduction by leveraging real-world energy usage time series data from smart buildings. The dataset, exemplified in Figure 5, comprises timestamps, net energy consumption (net), ambient temperature (tamb), demand, and DR participation levels. Each row corresponds to an energy usage event, with the 'time' column indicating the event timestamp, 'net' showing the net energy consumed at that moment, 'tamb' reflecting the ambient temperature, 'demand' representing the energy demand, and 'DR' detailing the DR intervention at that time. These records were treated as individual transactions within CypherChain, encrypted and securely aggregated, showcasing the framework's ability to handle real-world energy data confidentially. The CypherChain protocol ensures that sensitive information, such as individual energy usage patterns, cannot be isolated or identified during aggregation, thereby preserving prosumer privacy. This is achieved by employing the CYPHER protocol, which allows for the secure and private computation of aggregated results necessary for DR decisions. Figure 5 visually represents this data, where each transaction reflects a snapshot of the energy usage and DR variables, illustrating the practical application of CypherChain in managing and securing smart building energy data.

### B. Security and Privacy Analysis

CypherChain employs a multifaceted approach to mitigate the identified privacy threats in blockchain-based DR programs, introduced in [14]. Here's how each threat is addressed:

- **Data Exposure:** To counter data exposure risks, CypherChain utilizes the HE function, denoted as $E_{pk_i}(D_i)$, for encrypting transactions. This encryption ensures that sensitive data, such as individual energy consumption patterns, remain confidential while being processed on the blockchain. By doing so, it prevents potential privacy violations that could arise from the blockchain's inherent transparency.
- **Linkage Attacks:** The framework mitigates the risk of linkage attacks through its Hypergraph Coloring mechanism ($\mathcal{C}$), described in Section III-B. This approach segments the network into distinct clusters, reducing the traceability of individual transactions and thereby obfuscating transaction patterns. This segmentation makes it more challenging for attackers to correlate transactions and infer private information about prosumers.
- **Insider Threats:** Insider threats, such as malicious nodes or compromised aggregators, are addressed through the decentralized nature of the blockchain combined with a robust consensus mechanism. This mechanism is fortified by integrating homomorphic computation, which enhances trust and security within the network by ensuring that no single entity has undue influence or access to sensitive data.
- **Sybil Attacks:** The network's robust identity management and verification protocols are key in countering Sybil attacks. By ensuring that each participant in the network is authenticated and verified, CypherChain minimizes the likelihood of adversaries creating multiple fake identities to influence network operations or consensus mechanisms.
- **Collusion Attacks:** To mitigate the risk of collusion attacks, the framework employs a secret sharing scheme ($s_{ij}$) along with a threshold policy in its SMPC implementation. This approach ensures that private data cannot be reconstructed or accessed unless a certain threshold of shares is combined, thereby preventing unauthorized data aggregation and maintaining the confidentiality of prosumer data.

Through these strategies, CypherChain effectively addresses the complex privacy challenges inherent in blockchain-based DR programs. The combination of extended blockchain architecture and hypergraph-based network segmentation mechanism, along with robust privacy-preserving data aggregation protocols, provides a comprehensive solution to safeguarding prosumer privacy while maintaining the operational efficiency and transparency of the blockchain network.

### C. Performance Evaluation

The performance evaluation of the CypherChain framework utilizes the Hyperledger Caliper benchmarking tool to quantify its efficiency and scalability in the context of privacy-preserving demand response programs. The evaluation focuses on key performance metrics, including throughput, latency, and resource utilization.

*1) Throughput and Latency:* Throughput and latency form crucial metrics to assess the transaction processing capabilities of the system. We varied the transaction send rate ($T_{xcr}$ and $T_{xtr}$) from 10 to 100 transactions per second (tps) over a simulation interval of 100 seconds to measure the number of successfully processed transactions per unit of time (throughput) and the time taken for a transaction to be processed (latency). These evaluations were averaged over 10 runs to yield dependable performance measurements.

The evaluation results, shown in Figure 6, compare the CypherChain framework against a baseline system lacking the privacy-preserving features of CypherChain.

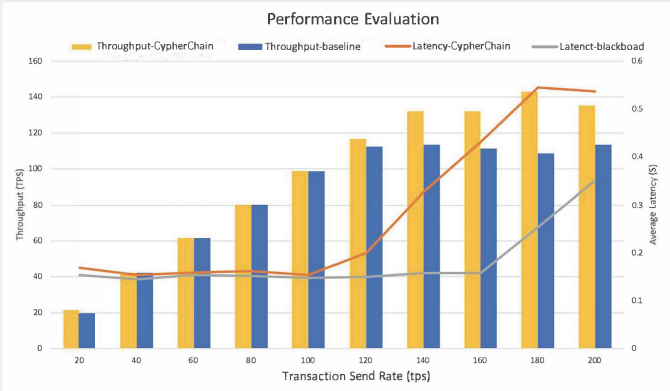


Fig. 6. Throughput and latency of the CypherChain framework compared to the baseline.

The CypherChain framework exhibits a throughput performance comparable to the baseline system's, suggesting the added privacy-preserving features do not significantly impact the transaction processing capabilities. Regarding latency, CypherChain posts marginally higher values than the baseline, especially at increased transaction rates. This added latency is attributed to the increased computational complexity induced by privacy-preserving operations such as secure multi-party computation and homomorphic encryption. However, the increased latency is within acceptable limits, ensuring the system's responsiveness remains largely unaffected by the privacy enhancements.

*2) Communication Overhead:* We evaluated the impact of Hypergraph clustering on optimizing Secure Multi-Party Computation (SMPC) with HE by assessing the communication overhead. The overarching goal was to reduce data exchange between clusters during cryptographic operations. By measuring the amount of data transferred or the number of communication rounds required between clusters, we gauged the effectiveness of Hypergraph clustering in decreasing communication overhead, thereby enhancing the overall performance. We determined the partitioning into $k$ clusters using a threshold $\theta$ that ranged from 0 to 1, aiming to minimize inter-cluster communication and enhance network efficiency and privacy.

*3) Resource Utilization:* To further understand the resource implications of the CypherChain framework, we compared its memory usage against that of the baseline system.

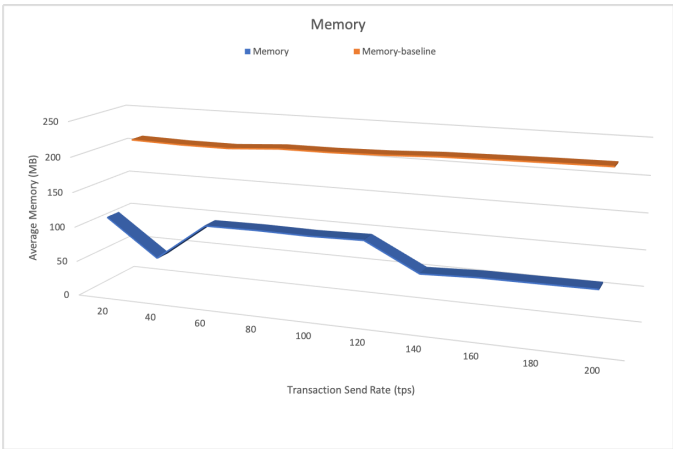


Fig. 7. Memory utilization of the CypherChain framework compared to the baseline.

Figure 7 shows that CypherChain demands marginally more memory resources than the baseline system. The increased memory usage can be traced back to the additional storage requirements for cryptographic keys and computations necessitated by privacy-preserving mechanisms. However, despite this increase, the memory usage remains within acceptable limits, affirming the framework's practicality for deployment on standard hardware configurations.

The CypherChain framework efficiently preserves privacy in demand response programs, performing comparably to conventional systems with minimal latency and resource usage increase, showcasing its practicality and potential for future enhancements and broader application.

## V. Conclusion

This research introduced CypherChain, an innovative framework aimed at balancing privacy with transparency in blockchain-based DR systems. Its unique approach, blending Homomorphic Encryption, Secure Multiparty Computation, and Hypergraph Coloring, has demonstrated significant potential for managing energy data securely. The empirical evaluation using smart building data underscores its effectiveness in real-world applications. This pioneering work opens new avenues for energy management research, emphasizing the need for continued advancements in privacy-preserving blockchain solutions. Moving forward, adapting CypherChain for diverse energy scenarios, including renewable energy sources and large-scale industrial applications, would be a valuable extension of this work.

## REFERENCES


[1] K. Kalsi, M. Elizondo, J. Fuller, S. Lu, and D. Chassin, "Development and Validation of Aggregated Models for Thermostatic Controlled Loads with Demand Response," *2012 45th Hawaii International Conference on System Sciences*, pp. 1959–1966, 2012.

[2] Y. Chen, M. Olama, X. Kou, K. Amasyali, J. Dong, and Y. Xue, "Distributed Solution Approach for a Stackelberg Pricing Game of Aggregated Demand Response," *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, 2020.

[3] S. Karumba, V. Dedeoglu, A. Dorri, R. Jurdak, and S. S. Kanhere, "Utilizing Blockchain as a Citizen-Utility for Future Smart Grids," *Wireless Blockchain: Principles, Technologies and Applications*, pp. 201–224, 2021.

[4] H. Fan and F. Li, "A privacy-preserving data aggregation scheme with fault tolerance for smart grid based on blockchain," vol. 12249, pp. 122 491A – 122 491A–8, 2022.

[5] V. Rigoni, D. Flynn, and A. Keane, "Coordinating Demand Response Aggregation With LV Network Operational Constraints," *IEEE Transactions on Power Systems*, vol. 36, pp. 979–990, 2021.

[6] S. Karumba, S. Sethuvenkatraman, V. Dedeoglu, R. Jurdak, and S. S. Kanhere, "Barriers to blockchain-based decentralised energy trading: a systematic review," pp. 41–71, 12 2023. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/14786451.2023.2171417

[7] E. Mengelkamp, B. Notheisen, C. Beer, D. Dauer, and C. Weinhardt, "A blockchain-based smart grid: towards sustainable local energy markets," *Computer Science - Research and Development*, vol. 33, no. 1-2, pp. 207–214, 2018.

[8] S. Karumba, S. S. Kanhere, R. Jurdak, and S. Sethuvenkatraman, "HARB: A Hypergraph-Based Adaptive Consortium Blockchain for Decentralized Energy Trading," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14 216–14 227, 8 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9187278/

[9] Z. Zhou, B. Wang, Y. Guo, and Y. Zhang, "Blockchain and Computational Intelligence Inspired Incentive-Compatible Demand Response in Internet of Electric Vehicles," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 3, pp. 205–216, 2019. [Online]. Available: https://dx.doi.org/10.1109/tetci.2018.2880693

[10] E. Androulaki, A. Barger, V. Bortnikov, S. Muralidharan, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Murthy, C. Ferris, G. Laventman, Y. Manevich, B. Nguyen, M. Sethi, G. Singh, K. Smith, A. Sorniotti, C. Stathakopoulou, M. Vukolić, S. W. Cocco, and J. Yellick, "Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains," *Proceedings of the 13th EuroSys Conference, EuroSys 2018*, vol. 2018-January, 4 2018. [Online]. Available: https://doi.org/10.1145/3190508.3190538

[11] A. Dorri, F. Luo, S. Karumba, S. Kanhere, R. Jurdak, and Z. Y. Dong, "Temporary immutability: A removable blockchain solution for prosumer-side energy trading," *Journal of Network and Computer Applications*, vol. 180, no. February, p. 103018, 2021. [Online]. Available: https://doi.org/10.1016/j.jnca.2021.103018

[12] O. Goldreich, "Secure multi-party computation," *Manuscript. Preliminary version*, 1998.

[13] C. Gentry, "Fully homomorphic encryption using ideal lattices," *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 169–178, 2009.

[14] B. Wang, L. Xu, and J. Wang, "A privacy-preserving trading strategy for blockchain-based P2P electricity transactions," *Applied Energy*, vol. 335, p. 120664, 4 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0306261923000284

[15] D. Li, Z. Aung, J. R. Williams, and A. Sanchez, "No peeking: privacy-preserving demand response system in smart grids," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 29, pp. 290–315, 2014.

[16] J. W. Bos, W. Castryck, I. Iliashenko, and F. Vercauteren, "Privacy-friendly Forecasting for the Smart Grid using Homomorphic Encryption and the Group Method of Data Handling," pp. 184–201, 2017.

[17] X. He, M.-O. Pun, and C.-C. J. Kuo, "Secure and efficient cryptosystem for smart grid using homomorphic encryption," *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1–8, 2012.

[18] A. Alexandru and G. J. Pappas, "Secure Multi-party Computation for Cloud-based Control," *ArXiv*, vol. abs/1906.0, 2019.

[19] L. Vogelsang, M. Lehne, P. Schoppmann, F. Prasser, S. Thun, B. Scheuermann, and J. Schepers, "A Secure Multi-Party Computation Protocol for Time-To-Event Analyses," *Studies in health technology and informatics*, vol. 270, pp. 8–12, 2020.

[20] Z. Cao, C. Huang, and Y. Li, "A Study on the improvement of Computation, Communication and Security in Garbled Circuits," *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 609–617, 2021.

[21] S. Karumba, S. Kanhere, R. Jurdak, and S. Sethuvenkatraman, "HARB: A Hypergraph-Based Adaptive Consortium Blockchain for Decentralized Energy Trading," *IEEE Internet of Things Journal*, vol. 9, pp. 14 216–14 227, 2022.

[22] M. Samadi, H. Schriemer, S. Ruj, and M. Erol-Kantarci, "Energy Blockchain for Demand Response and Distributed Energy Resource Management," *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 425–431, 2021.

[23] V. Deshpande, L. George, H. Badis, and A. Desta, "Blockchain Based Decentralized Framework for Energy Demand Response Marketplace," *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, pp. 1–9, 2020.

[24] L. Bracciale, P. Loreti, E. Raso, G. Bianchi, P. Gallo, and E. R. Sanseverino, "A Privacy-Preserving Blockchain Solution to Support Demand Response in Energy Trading," *2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON)*, pp. 677–682, 2022.

[25] A. Ghasemkhani, L. Yang, and J. Zhang, "Learning-Based Demand Response for Privacy-Preserving Users," *IEEE Transactions on Industrial Informatics*, vol. 15, pp. 4988–4998, 2019.

[26] S. Aslam, A. Tošić, and M. Mrissa, "Secure and Privacy-Aware Blockchain Design: Requirements, Challenges and Solutions," *Journal of Cybersecurity and Privacy*, 2021.

[27] X. Pei, X. Li, X. Wu, K. Zheng, B. Zhu, and Y. Cao, "Assured Delegation on Data Storage and Computation via Blockchain System," *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 55–61, 2019.

[28] D. Li, Z. Aung, J. R. Williams, and A. Sanchez, "P2DR: Privacy-Preserving Demand Response system in smart grids," *2014 International Conference on Computing, Networking and Communications (ICNC)*, pp. 41–47, 2014.