PureChain-Enabled Asymmetric Temporal Knowledge Transfer for Scalable Energy Forecasting

Nanteza Adah Lubwama [1], Love Allen Chijioke Ahakonye [2], Dong-Seong Kim [1] *, Jae Min Lee [1] †

[1] IT-Convergence Engineering, *Kumoh National Institute of Technology*, Gumi, South Korea

* NSLab Co. Ltd., Gumi, South Korea, *Kumoh National Institute of Technology*, Gumi, South Korea

[2] ICT Convergence Research Center, *Kumoh National of Technology*, Gumi, South Korea

(adah, loveahakonye, dskim, ljmpaul)@kumoh.ac.kr

***Abstract*—Energy consumption patterns display temporal asymmetry: peak hours (6–8 AM, 6–10 PM) feature dense, overlapping appliance signatures, whereas off-peak periods are cleaner but less diverse. Conventional Non-Intrusive Load Monitoring (NILM) methods treat these periods uniformly, leading to poor cross-temporal generalization and accuracy losses of up to 65% off-peak. We propose Asymmetric Temporal Knowledge Transfer (ATKT), a semi-supervised framework that leverages peak-to-off-peak knowledge transfer. ATKT combines multi-head attention for complex peak-hour disaggregation with LSTM-based modeling for off-peak sequences, unified under an asymmetric loss function that prioritizes peak periods. To ensure secure and distributed deployment, the framework integrates PureChain blockchain with a Proof-of-Authority-and-Association (PoA^2) consensus. Experiments on Energy Forecasting and Load Forecasting datasets show significant gains, with R^2 scores of 0.9324–0.9647 for load data and 0.997 for energy data, outperforming standard LSTM baselines. The blockchain layer delivers low latency (0.0317–0.0340s) and stable throughput (29.6–32.4 TPS), ensuring scalability. Overall, ATKT provides a comprehensive solution for intelligent energy management, enhancing grid stability and facilitating the integration of renewable energy sources.**

***Index Terms*—Blockchain, Energy Forecasting, LSTM, PoA^2, PureChain.**

I. Introduction

Energy consumption follows an apparent temporal asymmetry that existing Non-Intrusive Load Monitoring (NILM) techniques often overlook [1]. Peak periods (6-8 AM, 6-10 PM) have overlapping appliance operations, making disaggregation challenging despite their rich interaction signatures [2]. Off-peak times, though offering cleaner appliance patterns, lack the diversity needed for robust training [3]. This asymmetry reflects residential energy use, where morning and evening peaks involve simultaneous use of HVAC, water heaters, lighting, and electronics, creating complex, overlapping signatures. Off-peak periods, in contrast, typically feature individual appliance cycles, like refrigeration, offering more precise but less diverse training data.

Traditional semi-supervised disaggregation methods rely on large labeled datasets [4]. While recent approaches show promise, they treat all temporal periods uniformly [5], [6], overlooking the varying information content across different periods. Peak hours, with rich multi-appliance interaction patterns, offer insights that could aid off-peak disaggregation, which, while easier to identify, has limited training data. Existing methods fail to leverage cross-temporal knowledge transfer, missing opportunities to enhance disaggregation performance.

Temporal uniformity in NILM systems causes significant performance gaps. While baseline models perform well during peak periods, their accuracy drops by up to 65% during off-peak times, revealing a misalignment between training data and real-world conditions [7]. Models trained on peak data struggle with simpler off-peak patterns, while those trained on off-peak data miss crucial interactions during high-demand periods. Additionally, NILM systems face challenges in maintaining data integrity, verifying models, and ensuring secure deployment in distributed smart grids. These challenges include ensuring model authenticity, maintaining auditability for regulatory compliance, and enabling secure collaborative learning without compromising sensitive data consumption [8].

We introduce Asymmetric Temporal Knowledge Transfer (ATKT). This novel semi-supervised framework explicitly leverages peak-to-off-peak knowledge transfer for enhanced load disaggregation while addressing trust and verification challenges through blockchain integration [9]. ATKT recognizes that peak periods contain comprehensive appliance signature information that can be transferred to improve off-peak disaggregation accuracy, while the blockchain infrastructure ensures cryptographic verification of model integrity. This paper makes the following contributions:

- We introduce the semi-supervised approach that explicitly models temporal asymmetry between peak and off-peak consumption periods, using distinct neural architectures for each temporal context.
- Demonstration of the benefits of the asymmetric learning approach, where off-peak performance exceeds peak performance, proving successful knowledge transfer across temporal boundaries.
- We present a PureChain-based smart contract architecture that verifies ATKT models and validates predictions.

- We provide a comprehensive evaluation across multiple energy forecasting prediction metrics.

II. Related Works

The evolution of load disaggregation methodologies has progressed from Hart's foundational work on steady-state analysis to sophisticated deep learning approaches. Early methods utilized power signatures and transient detection [2], [10], while machine learning advances introduced hidden Markov models (HMMs) for state modeling [11] and CNNs for temporal pattern recognition [6]. Recent developments have focused on computational efficiency for smart grid applications, with Ahakonye et al. developing low-cost CNN architectures for frequency stability prediction [8] and time-efficient deep learning models for predicting industrial energy consumption [12]. However, these supervised approaches require extensive labeled datasets and treat all temporal periods uniformly, which limits their practical deployment scalability and misses opportunities to exploit natural information asymmetry in residential energy consumption patterns.

Semi-supervised learning has gained traction in energy disaggregation to reduce labeling efforts, with transfer learning techniques supporting household adaptation [13]. Recent frameworks that leverage temporal convolutional networks [1] and domain adversarial networks for cross-household generalization [4] have shown promising results. Temporal analysis reveals significant asymmetry in consumption data, with accuracy varying by up to 30% between peak and off-peak periods [7]. However, existing semi-supervised methods treat temporal periods uniformly, neglecting the distinct characteristics of peak (appliance interactions) and off-peak periods (cleaner individual patterns), a gap in current approaches.

Blockchain technology has gained prominence in energy systems for data integrity and secure multi-party computation [14], with recent advances addressing IoT-specific challenges. Proof-of-authority and association consensus mechanisms have been developed for IoT blockchain networks [15]. Meanwhile, blockchain cybersecurity trends in IoT environments have been analyzed [9], highlighting the security considerations essential for distributed energy systems [16]. However, blockchain applications in NILM remain limited, lacking integration with machine learning model verification and temporal performance validation.

The proposed framework integrates ATKT with a blockchain-backed trust layer to bridge the identified gaps. By leveraging temporal asymmetry, it establishes a semi-supervised NILM system that not only enhances disaggregation accuracy but also embeds cryptographic verification of model performance. This dual design enables secure and efficient deployment in distributed smart grid scenarios, effectively addressing challenges of temporal optimization and trust that remain unresolved in existing approaches.

III. Mathematical Formulation of Asymmetric Temporal Knowledge Transfer

ATKT is designed to capture the inherent differences in energy consumption dynamics between peak and off-peak periods. The central idea is to apply more expressive modeling during peak hours, where load patterns are complex and volatile, while using lighter sequential processing during off-peak periods, where patterns are smoother. A key advancement of ATKT lies in its asymmetric loss formulation, where peak period errors are weighted more heavily. This ensures that the model extracts richer representations from information-dense peak data and transfers this knowledge to enhance off-peak prediction.

A. Multi-Head Attention

To capture high-dimensional signature features during peak periods, ATKT employs a multi-head attention mechanism. Given input signatures in Equation 1.

$$S \in \mathbb{R}^{B \times T \times 128}, \tag{1}$$

the multi-head attention is defined in Equation 2.

$$\text{MultiHead}(S) = \text{Concat}(\text{head}_1, \dots, \text{head}_8) W^O, \tag{2}$$

where each attention head computes Equation 3.

$$\text{head}_i = \text{softmax}\left(\frac{S W_i^Q W_i^{K^T} S^T}{\sqrt{16}} \right) S W_i^V. \tag{3}$$

This formulation enables the model to attend selectively to different temporal dependencies, capturing both short-term fluctuations and longer-range correlations present in peak-hour data.

B. Asymmetric Temporal Processing

The processing strategy varies with temporal settings. During peak periods, multi-head attention is applied to exploit the richness of load fluctuations, while during off-peak periods, a lighter LSTM-based sequential model is sufficient, as shown in Equation 4.

$$\text{TemporalFeatures} = \begin{cases} \text{Mean}(\text{MultiHead}(S)), & \text{if peak period,} \\ \text{LSTM}(S), & \text{if off-peak period.} \end{cases} \tag{4}$$

This asymmetric design ensures that model complexity is adaptively allocated where it is most beneficial.

C. Forward Pass

In the ATKT forward pass, the input sequence X is transformed into feature signatures via stacked 1D convolutions $\text{Conv1D}(X^T)$. Meanwhile, contextual information is derived using a multilayer perceptron (MLP), yielding $\text{Context} = \text{MLP(features)}$. The temporal features and related embeddings are then concatenated to form a joint representation, $\text{Combined} = [\text{TemporalFeatures}; \text{Context}]$. This combined vector is finally passed through another MLP followed by a softmax layer, producing the model's prediction $\hat{y} = \text{Softmax}(\text{MLP}(\text{Combined}))$.

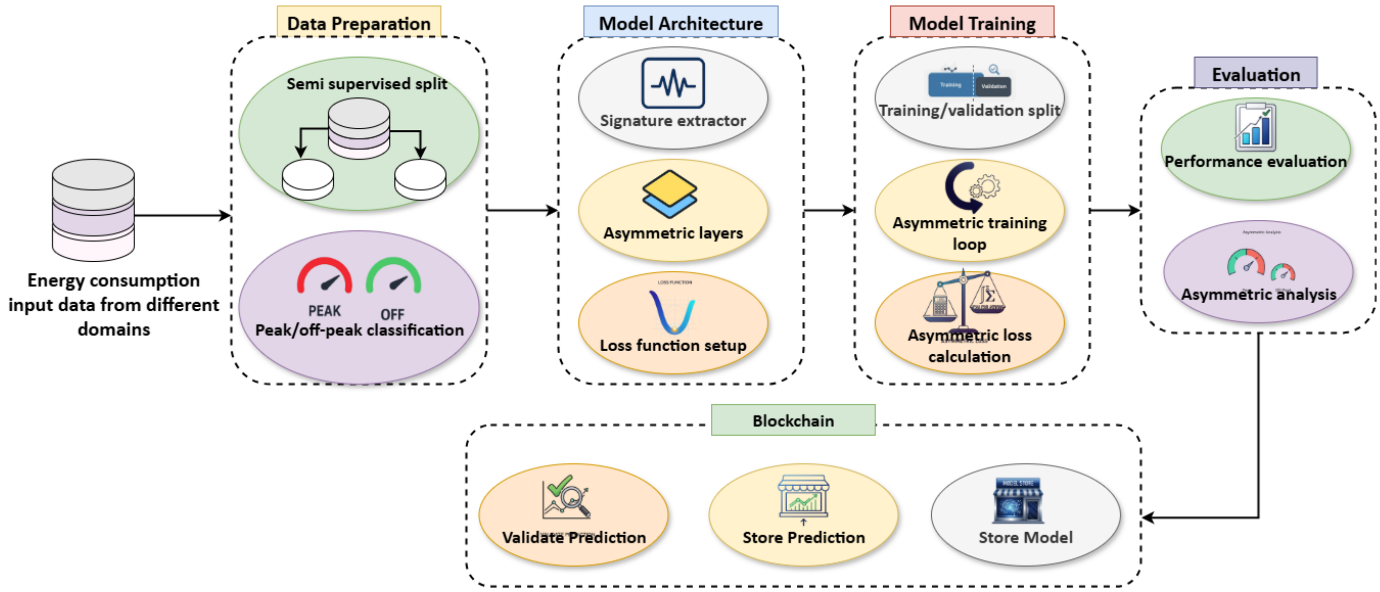

Fig. 1: Illustration of the proposed framework

D. Asymmetric Loss Function

The training objective explicitly encodes the asymmetric importance of peak and off-peak data in Equation 5.

$$\mathcal{L} = \mathcal{L}_{recon} + 0.1\mathcal{L}_{conf} + 0.5\mathcal{L}_{constraint}, \tag{5}$$

where the reconstruction loss is given in Equation 6.

$$\mathcal{L}_{recon} = \frac{1}{N}\sum_{i=1}^{N} w_i \|\hat{y}_i - y_i\|_2^2, \tag{6}$$

with temporal weighting in Equation 7.

$$w_i = \begin{cases} 2.0, & \text{if peak period,} \\ 1.0, & \text{if off-peak period.} \end{cases} \tag{7}$$

Additional terms regularize confidence estimation in Equation 9,

$$\mathcal{L}_{conf} = -\sum \log(c_i), \tag{8}$$

and enforce physical constraints on energy balance as in Equation 9.

$$\mathcal{L}_{constraint} = \sum \left(\sum_j \hat{y}_{ij} - 1 \right)^2. \tag{9}$$

This asymmetric weighting ensures that the model prioritizes learning from high-variability peak data while still maintaining generalization during off-peak intervals.

Our ATKT model incorporates the proof-of-authority and association (PoA^2) consensus mechanism [15], which enables model storage. The PureChain-based structure employs smart contracts for secure model storage and validation. It allows transparent attention pattern validation across a smart contract with key functions such as `storePredictionInput()` to record load disaggregation inputs with packed temporal context, including peak/off-peak indicators, `storePredictionOutput()` which stores multi-head attention predictions as packed appliance proportions, and `validatePrediction()/getPrediction()` for attention mechanism verification and asymmetric learning validation. The blockchain maintains records of attention head performance that inform our temporal context partitioning, with functions like `getModelStats()` tracking prediction accuracy across different temporal periods, `unpackAppliances()` decoding the 64-bit packed appliance data structure, and `hasDatasetAccess()` managing data sharing permissions.

Algorithm 1: ATKT Attention Pattern Validation

Input: Validators V, attention patterns A, contract C, temporal context T_c, threshold $\tau = 0.85$
Output: contractAddress$_{\text{ATKT}}$, $A_{\text{validated}}$
Initialize θ^0, buffer $B = 1024$;
Deploy $C_{\text{deployed}} \leftarrow \text{DEPLOY}(\theta^0)$;
Activate validators $V_{\text{active}} \leftarrow \text{REGISTER}(V, C_{\text{deployed}})$;
while *system* = *active* **do**
  $S_t \leftarrow \text{RANDOMSAMPLE}(A, 0.7|A|)$;
  foreach $a_i \in S_t$ ***in parallel*** **do**
    $(X, Y) \leftarrow \text{PACKTEMPORAL}(T_c)$;
    $w_{a_i}^{t+1} \leftarrow \text{PROCESSATKT}(\theta^t, X, Y)$;
    Submit $(\theta_{a_i}^{t+1}, |T_c|, \text{metrics}_{a_i})$;
  if *GroundTruth*$_{available}$ **then**
    $A^{t+1} \leftarrow \sum_{a_i} \frac{|T_c|}{n} w_{a_i}^{t+1}$;
    acc $\leftarrow \text{VALIDATEPREDICTION}(A^{t+1})$;
    if *acc* $\geq \tau$ **then**
      break
return *contractAddress*$_{ATKT}$, $A_{validated}$;

IV. Results Discussion And Analysis

A. Dataset Description and Experimentation

This study uses two data sets: Energy Forecasting [17] and Load forecasting [18].In the energy forecasting dataset has features like load, temperature, month, day, hour, weekend, season which were used while load forecasting dataset has load, temperature, hour, day, month as features used. The experiments were conducted on a Windows 11 Pro system with an Intel Core i5-8500 CPU at 3.00 GHz and 32GB of RAM. Visual Studio Code was used to carry out the experiments. The PureChain network was accessed through MetaMask with the smart contract written in Solidity and deployed using Remix IDE.

B. Results

This study evaluates the performance of the ATKT model and LSTM for energy consumption prediction using Loss, root mean squared error (R^2), and mean absolute error (MAE). Loss provides a general measure of model error during training, R^2 reflects how well the model captures data patterns, and MAE quantifies practical prediction accuracy by averaging error magnitudes. Together, these metrics provide a comprehensive evaluation of model performance, encompassing training efficiency and real-world predictive reliability. Figure 2 illustrates the ATKT model's learning process on the Load Forecast dataset. Over 50 epochs, both training and validation losses steadily decrease, with training loss dropping from 0.045 to 0.020. The validation loss shows a similar trend but remains slightly higher, indicating effective learning without overfitting. The MAE improves from 0.047 to 0.020, while the R^2 score increases from 0.70 to 0.95, reflecting significant gains in model accuracy and fit.

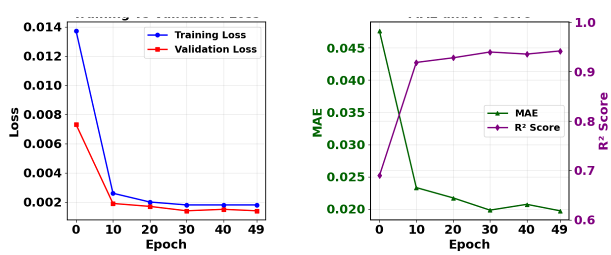

Fig. 2: Graph for ATKT Model using Load forecast dataset

Figure 3 shows the strong predictive performance of the ATKT model on the Energy Forecasting dataset, indicating regular underlying patterns in energy consumption. The model converges rapidly, with the training loss approaching zero within the first 10 epochs and remaining stable, while the validation loss remains consistently low. This suggests minimal risk of overfitting and that the multi-head attention mechanism effectively captures temporal correlations. The MAE stabilizes at approximately 0.005, and the R^2 score reaches 0.999, reflecting high prediction accuracy.

Table I compares model performance during peak and off-peak periods for two datasets, highlighting the effectiveness

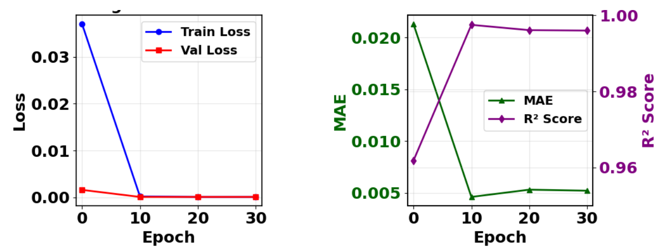

Fig. 3: Graph for ATKT model using Energy Forecasting dataset

TABLE I: ATKT Model Performance Results

ATKT	Metric	Load Forecast Data Scenario	Energy Forecasting Data Scenario
Peak	R^2	0.9324	0.9977
	MAE	0.0201	0.0043
Off-Peak	R^2	0.9647	0.9980
	MAE	0.0176	0.0042

of the ATKT approach. In the Load Forecast dataset, off-peak performance (R^2 = 0.9647, MAE = 0.0176) surpasses peak-period results (R^2 = 0.9324, MAE = 0.0201), suggesting that peak-period patterns enhance off-peak predictions. In the Energy Forecasting data scenario, both periods show similar high performance (R^2 = 0.997, MAE = 0.005), reflecting the ATKT framework's ability to capture regular consumption patterns. The low MAE values indicate practical prediction accuracy, while the high R^2 scores demonstrate the model's capability in capturing key consumption trends.

TABLE II: LSTM Model Performance Results

ATKT	Metric	Load Forecast Data Scenario	Energy Forecasting Data Scenario
Peak	R^2	0.8553	0.9965
	MAE	0.0364	0.0279
Off-Peak	R^2	0.6636	0.9835
	MAE	0.0542	0.0347

Table II illustrates the LSTM's variable performance on the Load Forecast data scenario, with apparent differences between peak and off-peak periods. During peak times, the model achieves an R^2 of 0.8553 and an MAE of 0.0364, indicating decent performance. However, its accuracy drops during off-peak periods, with R^2 falling to 0.6636 and MAE increasing to 0.0542. On the Energy Forecasting Dataset, the LSTM performs much better overall. Peak period performance is excellent, with an R^2 of 0.9965 and MAE of 0.0279, while off-peak performance remains strong, with an R^2 of 0.9835 and MAE of 0.0347. This consistent performance suggests that the dataset contains strong temporal patterns that the LSTM can effectively capture.

Table III compares previous studies [3], [7]. They reported high error rates, with MAE values of 2.4123 and 13.271, but did not provide R^2 scores or specify the testing conditions. Our study demonstrates the ATKT model, achieving a significant R^2 score of 0.9965 (Peak) and 0.9835 (Off-Peak) periods, and MAE values ranging from 0.0279 to 0.0347.

TABLE III: Comparison of different existing works

Author	Season	R^2	MAE
[7]	Not Specified	Not Specified	2.4123
[3]	Not Specified	Not Specified	13.271
Our Work	**Peak**	**0.9965**	**0.0279**
	Off Peak	**0.9835**	**0.0347**

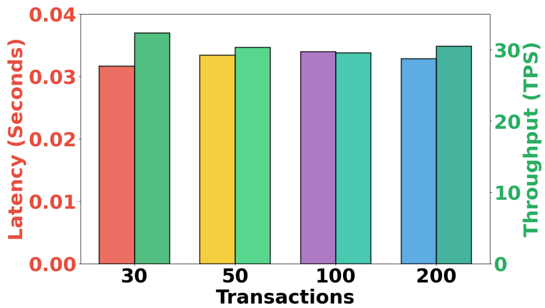

Fig. 4: PureChain Performance

Figure 4 illustrates the PureChain blockchain's performance with the ATKT framework for energy forecasting, demonstrating its effectiveness for reliable predictions. The data shows strong scalability, with latency remaining consistently low (0.0317 to 0.0340 seconds) as transaction loads increase from 30 to 200, exhibiting less than 7% variation. Throughput is stable at 29.6 to 32.4 transactions per second (TPS), with only a 9% fluctuation. This highlights the PoA^2 consensus mechanism's ability to efficiently handle varying verification demands without performance degradation, maintaining response times under 40 milliseconds while processing over 30 transactions per second.

V. Conclusion

The ATKT framework successfully addresses the temporal asymmetry challenge in energy forecasting by leveraging the rich information content of peak periods to enhance off-peak predictions. The integration of blockchain technology ensures secure and verifiable deployment in distributed smart grid environments. This approach provides utilities with a comprehensive solution for intelligent energy management systems that enhance grid stability and facilitate renewable energy integration, representing a significant advancement in energy forecasting.

Acknowledgment

This work was partly supported by Innovative Human Resource Development for Local Intellectualization program through the IITP grant funded by the Korea government(MSIT) (IITP-2026-RS-2020-II201612, 33%) and by Priority Research Centers Program through the NRF funded by the MEST(2018R1A6A1A03024003, 33%) and by the MSIT, Korea, under the ITRC support program(IITP-2026-RS-2024-00438430, 34%).

References

[1] Z. Zhu, M. Zhou, F. Hu, K. Wang, G. Zhou, W. Kong, Y. Hu, and E. Cui, "TSILNet: A Novel Hybrid Model for Energy Disaggregation Based on Two-Stage Improved TCN Combined with IECA-LSTM," *Building Simulation*, vol. 17, no. 11, pp. 2083–2095, 2024.

[2] L. Yan, R. Xu, M. Sheikholeslami, Y. Li, and Z. Li, "State Identification of Home Appliance with Transient Features in Residential Buildings," *Frontiers in Energy*, vol. 16, no. 1, pp. 130–143, 2022.

[3] Z. Shan, G. Si, K. Qu, Q. Wang, X. Kong, Y. Tang, and C. Yang, "Multiscale Self-Attention Architecture in Temporal Neural Network for Nonintrusive Load Monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[4] T. Tang, K. Li, C. Su, and Z. Liu, "Semi-Supervised Learning with Flexible Threshold for Non-Intrusive Load Monitoring," *Heliyon*, vol. 10, no. 14, p. e34457, 2024.

[5] M. Ayub and E.-S. M. El-Alfy, "Contextual Sequence-to-Point Deep Learning for Household Energy Disaggregation," *IEEE Access*, vol. 11, pp. 75 599–75 616, 2023.

[6] R. Zhang, W. Luan, B. Liu, and M. Zhong, "A Lightweight Neural Network for Energy Disaggregation Employing Depthwise Separable Convolution," in *2021 IEEE Sustainable Power and Energy Conference (iSPEC)*, 2021, pp. 4109–4114.

[7] A. N. Khan, Q. Waqas Khan, J. Lee, R. Ahmad, M. Bibi, D. Ho Kim, J.-S. Sung, and D.-H. Kim, "Dynamic Temporal Analysis and Modeling of Residential Lighting Consumption for Energy Efficiency and Sustainability," *IEEE Access*, vol. 12, pp. 154 365–154 380, 2024.

[8] L. A. C. Ahakonye, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Low Computational Cost Convolutional Neural Network for Smart Grid Frequency Stability Prediction," *Internet of Things*, vol. 25, p. 101086, 2024.

[9] L. A. C. Ahakonye, C. I. Nwakanma, and D.-S. Kim, "Tides of Blockchain in IoT Cybersecurity," *Sensors*, vol. 24, no. 10, p. 3111, 2024.

[10] L. Yan, W. Tian, J. Han, and Z. Li, "eFHMM: Event-Based Factorial Hidden Markov Model for Real-Time Load Disaggregation," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3844–3847, 2022.

[11] Z. Zhou, Y. Xiang, H. Xu, Y. Wang, and D. Shi, "Unsupervised Learning for Non-intrusive Load Monitoring in Smart Grid Based on Spiking Deep Neural Network," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 606–616, 2022.

[12] L. A. C. Ahakonye, C. I. Nwakanma, J. M. Lee, and D.-S. Kim, "Time-Efficient Deep Learning-Based Energy Consumption Prediction for Smart Factory," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 879–882.

[13] M. Kaselimi, A. Voulodimos, E. Protopapadakis, N. Doulamis, and A. Doulamis, "EnerGAN: A Generative Adversarial Network for Energy Disaggregation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1578–1582.

[14] X. Zhao, V. Kumar, and R. Smith, "Blockchain technology for data integrity in distributed energy systems," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 12, pp. 12 456–12 467, 2022.

[15] D.-S. Kim, I. S. Igboanusi, L. A. Chijioke Ahakonye, and G. O. Anyanwu, "Proof-of-Authority-and-Association Consensus Algorithm for IoT Blockchain Networks," in *2025 IEEE International Conference on Consumer Electronics (ICCE)*, 2025, pp. 1–6.

[16] O. Jogunola, B. Adebisi, T. Shongwe, and A. Yunusa-Kaltungo, "Security of Blockchain-Based Applications: A Case of Distributed Energy Systems," in *Key Themes in Energy Management: A Compilation of Current Practices, Research Advances, and Future Opportunities*. Springer, 2024, pp. 397–414.

[17] D. Lajic, "Energy forecasting files dataset," 2024, accessed: August 25, 2025. [Online]. Available: dlaj/energy-forecasting-files

[18] M. Voss, "Load forecasting dataset," GitHub Gist, 2021. [Online]. Available: https://gist.githubusercontent.com/marcus-voss/load data.csv