Article

A Distributed Cybersecurity Monitoring Architecture for Smart Microgrids

Yinglian Zuo[1,*]

[1] Sichuan Jiewuji Technology Co., Ltd., Chengdu, China
* Correspondence: Yinglian Zuo, Sichuan Jiewuji Technology Co., Ltd., Chengdu, China

Abstract: Smart microgrids rely on continuous communication among distributed generation units, energy storage systems, flexible loads, and advanced control systems. Consequently, any network anomaly can propagate into frequency instability, voltage deviation, power-allocation imbalance, or protection-system failure. Centralised monitoring approaches support unified governance but may lose critical evidence or miss real-time deadlines across heterogeneous nodes, rapidly changing operating modes, or intermittently connected upstream links. This paper develops a distributed cybersecurity monitoring architecture comprising edge monitoring agents, regional coordination nodes, and a global security-orchestration platform. Responsibilities are systematically organised through four functional layers: local observation, neighbourhood corroboration, global governance, and controlled response. Network behaviour is interpreted jointly with the electrical state, while topology association, auxiliary evidence channels, and graded interlocks collectively support threat detection, fault localisation, and containment. Edge agents retain high-frequency raw evidence and perform low-risk local checks. Regional nodes verify events against neighbouring communication and electrical relationships. The global platform governs rules, models, credentials, and high-impact actions without entering the fastest control loop. A published islanded-microgrid simulation demonstrates how an independent auxiliary network can reveal covert intrusions and support system recovery. The proposed architecture creates a continuous, auditable evidence chain and preserves monitoring capability even when portions of the communication infrastructure are degraded, thereby enhancing the overall resilience of smart microgrid operations against emerging cyber-physical threats.

Keywords: smart microgrid; cybersecurity monitoring; distributed architecture; cyber-physical systems; intrusion detection

Received: 07 June 2026
Revised: 28 July 2026
Accepted: 11 August 2026
Published: 16 August 2026

1. Introduction

A smart microgrid integrates photovoltaic generation, energy storage, charging infrastructure, controllable loads, protection equipment, and energy-management functions within a localized energy system. It supports both grid-connected and islanded operational modes, distributed secondary control, and market-informed scheduling—requiring continuous exchange of measurements and control signals among devices. When unauthorized access occurs along a communication path, anomalous data may propagate through control dependencies and impact frequency stability, voltage regulation, power distribution, or protective device response [1]. Recent updates to the cybersecurity framework for electric power monitoring systems in China extend coverage to monitoring systems associated with distributed generation, energy storage, and virtual power plants, mandating integrated monitoring and early-warning capabilities enabled by embedded sensing and related technical approaches. Cybersecurity monitoring in microgrids must therefore identify irregular communication patterns and assess their potential implications for physical system performance.

Current technical guidelines prescribe security zoning, boundary isolation, identity verification, monitoring, and audit requirements for power-monitoring systems [2]. These

safeguards remain essential; however, their deployment in microgrids faces distinctive challenges arising from distributed node topologies, resource-constrained edge devices, and dynamically reconfigurable operating structures. Centralized collection of all packets and measurements imposes excessive bandwidth demands and renders time-critical detection contingent upon upstream connectivity. Conversely, fully decentralized decision-making at each node introduces limitations: no single node possesses sufficient contextual awareness to identify coordinated anomalies or reliably differentiate local faults from inconsistencies originating elsewhere in the system. A robust distributed architecture should not eliminate centralized coordination but rather allocate observation, validation, and response responsibilities across layers according to temporal constraints and evidentiary requirements.

The proposed design traces the causal pathways linking cyber-level events to physical-system outcomes. Raw observational data remains proximal to the source device. Adjacent nodes jointly verify network topology, temporal alignment, and physical consistency. A centralized platform maintains authoritative models, policy rules, authentication credentials, and protocols for high-consequence interventions. What is exchanged across layers is not an undifferentiated data stream, but a structured security event containing its origin, operational context, temporal reference, supporting evidence, and confidence assessment. This layered approach enables rapid local evaluation without conflating preliminary indicators with definitive conclusions [2]. It also supports cross-node corroboration and retrospective analysis even when the central platform is temporarily inaccessible.

2. Architectural Mismatches in Smart-Microgrid Security Monitoring

2.1. Single-Domain Evidence Cannot Explain Cyber-Physical Anomalies

Disruptions to service availability, message replay, unauthorized data modifications, and command alterations may leave observable traces in network connections, timing patterns, or message structures. These same phenomena can also occur within otherwise compliant communication exchanges. At the physical layer, variations in load demand, renewable generation output, sensor calibration drift, or switching device status can induce frequency and power deviations indistinguishable from those arising under abnormal operational conditions [3]. Traffic analysis alone may fail to detect subtle deviations that preserve protocol syntax and structure. Similarly, electrical-state monitoring alone may misclassify a planned islanding event as an anomalous condition. Consequently, effective security monitoring in cyber-physical power systems relies on the integrated use of network behavior indicators, physical model residuals, and available system recovery capabilities.

An alert originating from either domain should be interpreted as preliminary evidence rather than a definitive classification [4]. For instance, the appearance of a new communication endpoint may reflect routine system maintenance, configuration updates, or transient network reconfiguration. A high state-estimation residual may stem from equipment degradation, measurement uncertainty, or model-parameter mismatch with current operating conditions. Confidence in diagnostic conclusions increases when multiple independent indicators align—for example, an unverified control command, an atypical operational sequence, and an electrical response inconsistent with the intended control objective. This requirement tightly couples monitoring architecture with precise data semantics: network events must be unambiguously associated with equipment identity, geographical and topological location within the grid, and real-time operational context before enabling appropriate and safe system responses.

2.2. Central Aggregation Does Not Always Meet Edge Response Deadlines

A central platform is well suited to long-horizon statistics, cross-node correlation analysis and policy consistency [5]. It is poorly positioned to make every time-critical decision when edge-to-centre links have finite bandwidth, variable delay or intermittent availability. Waiting for centralized analysis may cause an urgent anomaly to miss the

window for rate limiting, command rejection or transfer to a safer control mode. A situational-awareness platform can combine multidimensional event correlation, baseline learning and cross-node behavioral analysis to form a system-wide view. That function is most effective when edge nodes have already extracted evidence, verified timestamps and assigned a preliminary scope.

The centre should enlarge the field of observation, not become a prerequisite for every local protection action. A device-facing agent must retain sufficient operational capability to record evidence and execute pre-authorized low-risk measures when upstream communication is unavailable. Regional coordination should remain available for events whose interpretation depends on neighbouring nodes. Only responses with broad operational consequences should require centralized authorization [6]. Such separation reduces a critical dependency: a monitoring system should not lose all detection and response capability precisely because the communication infrastructure it monitors has experienced temporary unavailability.

2.3. Fixed Baselines Fail Across Operating-Mode Changes

A transition from grid-connected to islanded operation alters traffic patterns, control gains, power flows, and permissible deviations [1]. Storage charging, maintenance activities, and demand response introduce additional legitimate variations from a static baseline. A model trained under one operational mode may generate numerous false alarms during mode transitions. An overly broad threshold designed to suppress such alerts risks overlooking sustained low-amplitude anomalies. Artificial-intelligence methods can extract features from terminal measurements, flow data, communication messages, and operational system logs, yet their outputs remain dependent on network topology, control configuration, and data integrity. Consequently, the baseline must be dynamically conditioned on the real-time operating state rather than defined as a fixed, universal representation of normal behavior.

Baseline updates also entail operational integrity considerations. If data collected during an unconfirmed or unresolved event are incorporated into model training, atypical behavior may inadvertently become embedded in the reference profile. Each model version must explicitly document its associated topology configuration, operational modes, data collection period, and any time intervals excluded due to ongoing investigations [7]. New versions should undergo validation using verified benign operational events and representative anomalous cases, be introduced incrementally, and remain co-deployed with a stable prior version to support controlled rollback. State-aware baselines effectively reduce false alarms only when model updates themselves follow rigorously controlled, traceable, and auditable procedures. The three mismatches described above collectively necessitate three architectural principles: distributed observation, dual-domain corroboration, and explicit representation of operating state.

3. Functional Layers and the Monitoring Responsibility Chain

The architecture is structured around four interlinked functional responsibilities: local observation, neighbourhood corroboration, global governance, and controlled response. Edge agents are deployed adjacent to distributed generators, energy storage controllers, protection devices, measurement units, and energy management gateways. Regional nodes are established based on electrical topology or network access domains. The global platform integrates monitoring functions with security administration and audit capabilities. These layers do not simply relay alarms upward; each contributes contextual interpretation, validates distinct relational logic, and operates strictly within its assigned authority scope [8].

3.1. Local Observation and Preservation of Raw Evidence

The responsibility chain begins with evidence available when an anomaly occurs. An edge agent preferably observes traffic passively and records communication endpoints, protocol, direction, call frequency, authentication result, and inter-message timing. It also reads electrical quantities relevant to the attached device, including frequency, voltage,

active power, state of charge, and control residuals. Rule checks can identify unauthorised access, commands outside an approved sequence, and unexpected parameter changes [9]. A state observer can test whether the measured response is consistent with the device model and the current command.

A distributed detection scheme for DC microgrids has demonstrated that subsystem observers and unknown-input observer banks can use distinct residuals to enhance anomaly detectability. The architectural insight is not that a single observer suffices for all cases, but that detection evidence can be generated close to the physical process without transmitting all raw data [3]. Each local event should include the affected object, operating mode, time source, evidence summary, model version, and confidence level. High-frequency packets and measurements remain locally available for forensic review. During communication interruption, the agent can continue recording and apply only pre-authorised reversible measures.

3.2. Neighbourhood Corroboration and Event Localisation

A local alert identifies where suspicious evidence appeared, but it does not establish the origin or scope of the event. The regional node compares observations according to communication adjacency and electrical coupling. It can reconcile records from the transmitting and receiving ends, verify whether neighbouring power responses are consistent, and confirm whether a command originated from a legitimate control relationship. Research on security monitoring in distributed frequency control demonstrates that anomalies affecting links and nodes can be identified and localised using neighbouring information and model-based indicators [10].

Regional monitoring therefore maintains a dynamic map of node identity, network routes, electrical topology and control dependence [5]. The map must adapt when the microgrid islands, reconnects or reconfigures. Correlation distinguishes among equipment failure, path abnormality and coordinated deviations. A mismatch found on one communication path but not reflected in the physical response may indicate a network-level anomaly. Simultaneous inconsistencies across several related nodes support higher-confidence assessment and broader containment measures. Only the event summary and evidence index need to be sent globally; raw high-rate data can remain at the originating nodes until requested.

3.3. Global Governance without a Real-Time Single Point

Following regional validation, the global platform analyzes events across multiple regions to identify shared vulnerabilities and reconstruct cross-regional propagation pathways. It centrally manages detection rules, model packages, digital certificates, access policies, and response playbooks. Each deployed model must specify its compatible equipment, supported operational modes, training duration, and conditions for deprecation. All rules and models are cryptographically signed, rolled out incrementally, and validated post-deployment. The platform maintains an auditable log recording rule modifications, distribution status across nodes, and verification of resulting alert behavior against expected outcomes.

3.4. Controlled Response under Safety Interlocks

Response authority should scale with assessment confidence and potential operational impact. A low-confidence event triggers additional sampling, evidence preservation, or challenge-response verification. A medium-confidence event may warrant session rate limiting, redirection to a redundant path, or re-authentication. Isolation, controlled functional reduction, or safe shutdown is reserved for high-confidence events posing significant operational risk. This tiered approach prevents both underreaction—delaying action until damage occurs—and overreaction—disrupting healthy resources due to transient statistical fluctuations.

Actions involving disconnection of primary equipment, modification of protection parameters, or substantial load shedding must comply with defined operational requirements and require explicit human authorization. Autonomous security

mechanisms must never exceed these boundaries. Each response procedure specifies the triggering evidence, permissible scope, maximum execution duration, conditions for automatic rollback, and the designated operator. Under degraded communication, local automation is restricted to bounded, fully reversible interventions. The evidentiary chain thus enables timely response while ensuring the monitoring layer remains subordinate to engineered safeguards essential for microgrid stability.

4. Cross-Layer Mechanisms for Continuous Evidence

Functional layering assigns responsibility, but cross-layer mechanisms determine whether evidence remains usable from detection to review. Incompatible event semantics, scoring rules, or model versions across layers fragment the architecture into three isolated alert systems. Continuity requires shared event semantics, synchronized temporal references, dual-domain evidence, communication diversity, and auditable learning.

4.1. Common Event Semantics and Time References

Cross-node analysis depends on comparable records. An event should include node identity, collection point, operating mode, source and reception times, communication direction, physical units, data-quality flags, and model version. Discrepancies in temporal alignment may indicate integrity anomalies, as timing inconsistencies can arise from unauthorized interference with data transmission or processing. When time synchronization is impaired, the system should reduce reliance on strict chronological ordering rather than maintain unwarranted confidence in sequence fidelity. A common event identifier and an evidence hash enable regional and global nodes to validate data integrity and support reconstruction of the observation, escalation, and response timeline.

4.2. Cyber-Physical Evidence in Risk Scoring

A risk score may integrate network anomaly indicators, physical residual measurements, topological significance, and data reliability, but the numerical value must not be interpreted as a definitive determination of malicious activity. Elevated network anomaly alongside stable electrical conditions warrants investigation into scanning behavior, unsuccessful access attempts, or configuration inconsistencies. Substantial physical residual in the presence of normal communication patterns necessitates verification of equipment malfunction or sensor calibration drift. When both network and physical indicators co-occur on electrically interconnected nodes, the response priority may be elevated. Weighting parameters vary across grid-connected, islanded, and maintenance operational modes and should be refined using verified incident records rather than unconfirmed alerts.

An interpretable risk score must retain transparency of its constituent elements. Operators require clarity on whether severity stems from critical device involvement, significant physical impact, repeated authentication failures, or diminished data credibility. Maintaining component-level records also facilitates subsequent refinement. If an event is determined to reflect a system fault rather than adversarial behavior, only the corresponding rule or model component should be updated. Treating the score as a black-box classification obscures such distinctions and promotes indiscriminate threshold adjustments that may mask underlying performance degradation.

4.3. Auxiliary Evidence Channels for Covert Attacks

A system operator managing the primary communication path may inadvertently introduce inconsistencies in both measurements and the associated consistency verification information. To enhance reliability, critical nodes can employ a logically independent auxiliary evidence channel that transmits low-bandwidth state summaries, challenge-response data, or neighbourhood validation values. This auxiliary path must avoid redundancy with operational traffic and must be isolated from the primary path in terms of credentials, routing infrastructure, and management systems. Its core function is to provide diversified evidence, not to augment data throughput.

A study on an islanded AC microgrid employed coordinated interaction between a control layer and an auxiliary layer to detect anomalies and support restoration of frequency regulation and active-power distribution. Table 1 summarises the published simulation results. These findings are not intended as universal performance guarantees but illustrate the architectural rationale for incorporating an independent evidence layer: when the primary consensus mechanism exhibits deviation, cross-validation against a separate information pathway enables inconsistency detection and triggers resilient control responses.

Table 1. Microgrid Response Before and After Activation of an Auxiliary Network in a Published Simulation

Simulation event	Without auxiliary network	With auxiliary network	Architectural implication
Frequency attack injected at 10 s	Frequency deviated by about 19 rad/s and settled at 295 rad/s	Recovered to approximately 314 rad/s in about 5 s	An independent evidence layer can support detection and recovery
Additional power attack injected at 30 s	Common steady-state frequency moved to 298 rad/s	Recovered to approximately 314 rad/s in about 5 s	Frequency and power information should be monitored jointly
Active-power allocation after the attacks	Power sharing was disrupted	Deviation lasted about 1.5 s before returning near the rated value	Monitoring results should be linked to resilient control

Source: Compiled from the islanded AC-microgrid simulation reported by Vaishnav et al.[8]

In the absence of the auxiliary network, simulated anomalies propagated through the distributed control information stream [11]. With the integration of supplementary evidence and a resilient control strategy, frequency stabilised near its nominal value and active-power sharing recovered following a brief transient deviation. Recovery duration varies depending on controller parameters, equipment characteristics, and network topology. The key architectural insight is more focused: corroboration must remain feasible even when the principal communication pathway cannot be fully trusted. Therefore, auxiliary evidence mechanisms, neighbourhood verification protocols, and response interfaces should be co-designed as an integrated functional loop rather than as disjointed security components.

4.4. Auditable Model Updating and Evidence Retention

Nodes may update anomaly baselines locally and exchange parameter summaries or feature statistics. Windows containing unresolved events are excluded to mitigate model contamination risks. Prior to deployment, a new version undergoes replay-based testing and is compared against the stable version for differences in missed detections, false alarms, and decision latency. The prior version remains available for rollback. Confirmed outcomes may adjust rules, thresholds, and topology configuration, but underlying records remain unaltered. Model accuracy, version history, operator decisions, and response results collectively constitute the audit chain.

Retention policy balances evidential utility with operational limitations at edge nodes. High-rate raw data may be retained locally for a defined duration, whereas event summaries and cryptographic hashes are preserved for extended periods at regional and global levels. A serious event triggers preservation of the associated local data window prior to routine deletion. Access to forensic data is logged and restricted by role-based authorization. This design avoids centralized, persistent storage of all network packets while ensuring availability of essential evidence needed to validate consequential decisions.

5. Engineering Implementation and Performance Evaluation

5.1. Staged Deployment Around a Clear Control Boundary

Deployment should begin with a part of the microgrid whose devices, data, and control responsibilities are well understood. Engineers first inventory communication dependencies, operating modes, external dispatch links, non-interruptible functions, and actions that can be reversed safely [12]. Edge agents are then installed in passive mode to establish baselines for grid-connected, islanded, switching, and maintenance conditions. Historical anomalies, simulated disturbances, and controlled exercises test whether alerts can be explained. Only after monitoring is stable should the system enable rate limiting, backup-channel switching, or limited local isolation.

Regional coordination can be added after node identities and topology updates are reliable. Global model distribution and high-impact operational procedures follow last. This sequence makes the failure location visible and prevents broad automation from being introduced before evidence quality is confirmed. Cross-zone traffic should contain only the event information required by the upper layer and must respect existing security zoning. A distributed design is not a justification for creating unrestricted new communication paths.

5.2. Evaluation Across the Responsibility Chain

Evaluation should follow the same responsibility chain as system operation. At the edge, relevant metrics include detection rate, false-alarm rate, evidence-formation delay, and resource consumption. At the regional layer, assessments examine localization range, correlation accuracy, and performance under partial communication conditions. At the global layer, evaluation covers policy consistency, model-distribution integrity, audit completeness, and sustained operation during central disconnection. Electrical performance indicators include maximum frequency or voltage deviation, power-sharing error, unsafe control actions, and recovery time.

Tests must cover routine operational variations, equipment faults, isolated and distributed anomalies, timing deviations, operational mode transitions, and network segmentation. Accuracy measured on static labelled datasets is insufficient, as it omits dynamic operating transitions that frequently trigger false positives. Response validation must also confirm system rollback capability. When an automated action exceeds safety thresholds, the system reverts to alert-only mode and requires authorized human oversight. Performance is deemed acceptable only when enhanced cyber visibility does not compromise microgrid stability or availability.

5.3. Operational Roles and Maintenance

Asset personnel maintain device identity, firmware, and local monitoring configuration. Microgrid operators assess electrical impacts and approve high-risk operational actions. Security teams update threat intelligence, correlation logic, and evidence handling procedures. Suppliers assist with sensor integration, system modeling, and protocol adaptation while operating under defined authority boundaries. Significant modifications to detection logic undergo coordinated review alongside protection configurations and control algorithms. Joint incident response exercises clarify evidentiary responsibilities and delineate operational authority across functional roles [13].

5.4. Boundaries of the Architecture

The architecture enhances observability and continuity, yet it does not ensure identification of all security incidents. Adversarial actions may affect multiple evidence sources, stem from modelling limitations, or fall below current detection thresholds. Distributed components also increase the number of software instances requiring timely updates and authentication. Consequently, this design serves as a complement—not a substitute—for network segmentation, secure development practices, access control, and coordinated protective measures [5]. Its efficacy relies on diversity across evidence pathways and rigorous governance of interpretation mechanisms.

6. Conclusion

Cybersecurity monitoring in a smart microgrid requires a deliberate allocation of observability and response responsibility. A single centralised system cannot always provide timely, continuous detection, while isolated edge probes cannot fully explain anomalies spanning multiple nodes and physical processes. Local observation, neighbourhood corroboration, global governance, and controlled response enable fast evidence collection near equipment, topology-based localisation at regional nodes, and unified management of rules, models, and audit logs at the global platform. Common event semantics, cyber-physical evidence, auxiliary communication channels, and controlled rollback preserve system functionality during intermittent connectivity, while safety interlocks ensure that monitoring outcomes never compromise electrical safety limits. Published simulation data illustrate the value of independent evidence; however, realistic topology, propagation delays, equipment heterogeneity, and multi-stage adversarial scenarios still require hardware-in-the-loop validation and controlled operational trials before broad deployment of automated containment mechanisms.

References

1. W. Lei, H. Wen, J. Wu, and W. Hou, "MADDPG-based security situational awareness for smart grid with intelligent edge," *Appl. Sci.*, vol. 11, no. 7, p. 3101, 2021.
2. D. Du, M. Zhu, X. Li, M. Fei, S. Bu, L. Wu, and K. Li, "A review on cybersecurity analysis, attack detection, and attack defense methods in cyber-physical power systems," *J. Mod. Power Syst. Clean Energy*, vol. 11, no. 3, pp. 727–743, 2022.
3. T. Bhuiyan, "AI in smart grid cybersecurity: A systematic review of machine learning and deep learning approaches against false data injection and other emerging attacks," *J. Comput. Sci. Technol. Stud.*, vol. 7, no. 8, pp. 1207–1295, 2025.
4. A. J. Gallo, M. S. Turan, F. Boem, T. Parisini, and G. Ferrari-Trecate, "A distributed cyber-attack detection scheme with application to DC microgrids," *IEEE Trans. Autom. Control*, vol. 65, no. 9, pp. 3800–3815, 2020.
5. L. Y. Lu, H. J. Liu, H. Zhu, and C. C. Chu, "Intrusion detection in distributed frequency control of isolated microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6502–6515, 2019.
6. V. Vaishnav, A. Jain, and D. Sharma, "Auxiliary network-enabled attack detection and resilient control of islanded AC microgrid," *IEEE Trans. Smart Grid*, vol. 16, no. 2, pp. 1641–1653, 2024.
7. M. Karanfil, D. E. Rebbah, M. Debbabi, M. Kassouf, M. Ghafouri, E. N. S. Youssef, and A. Hanna, "Detection of microgrid cyberattacks using network and system management," *IEEE Trans. Smart Grid*, vol. 14, no. 3, pp. 2390–2405, 2022.
8. A. Alshammari, "Securing smart microgrids with a novel multi-layer cybersecurity framework for Industry 4.0 renewable energy systems," *Discover Comput.*, vol. 28, no. 1, p. 80, 2025.
9. Y. Li and J. Yan, "Cybersecurity of smart inverters in the smart grid: A survey," *IEEE Trans. Power Electron.*, vol. 38, no. 2, pp. 2364–2383, 2022.
10. M. Ghiasi, M. Dehghani, T. Niknam, A. Kavousi-Fard, P. Siano, and H. H. Alhelou, "Cyber-attack detection and cyber-security enhancement in smart DC-microgrid based on blockchain technology and Hilbert Huang transform," *IEEE Access*, vol. 9, pp. 29429–29440, 2021.
11. M. Khaleel, Z. Yusupov, H. J. El-Khozondar, and A. Alsharif, "Cyber-resilience strategies for smart microgrids: Classification, construction, recent trends, and policy framework," *Int. J. Electr. Eng. Sustain.*, vol. 3, no. 3, pp. 31–47, 2025.
12. G. B. Gaggero, P. Girdinio, and M. Marchese, "Advancements and research trends in microgrids cybersecurity," *Appl. Sci.*, vol. 11, no. 16, p. 7363, 2021.
13. P. Thulasiraman, M. Hackett, P. Musgrave, A. Edmond, and J. Seville, "Anomaly Detection in a Smart Microgrid System Using Cyber-Analytics: A Case Study," *Energies*, vol. 16, no. 20, p. 7151, 2023.